SABER
Overview

April 2013

The What, Why, and How of the Systems Approach for Better Education Results (SABER)

(Conference Edition)

THE WORLD BANK

Contents

Acknowledgements

Executive summary

The new **SABER (Systems Approach for Better Education Results)** initiative is helping the World Bank and its development partners to collect and analyze information on policies and identify actionable priorities for strengthening education systems, so that all children and youth can be equipped with knowledge and skills for life. This paper describes SABER's rationale and approach.

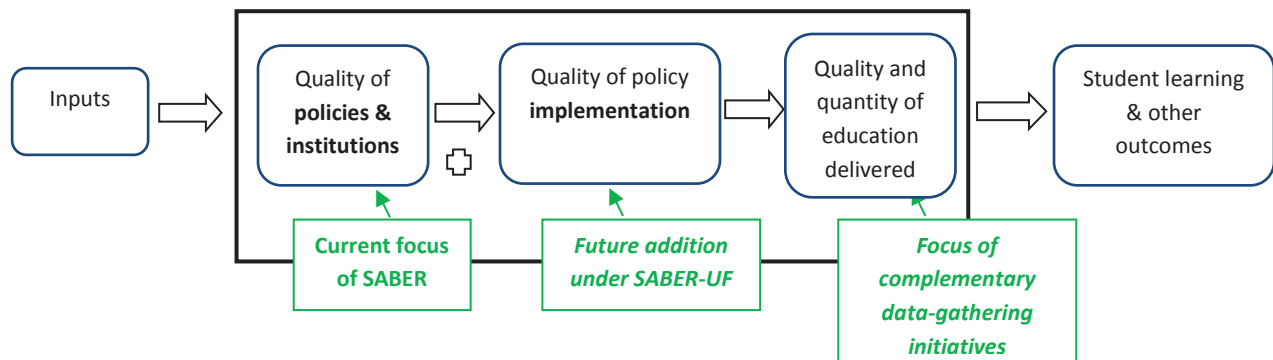## The challenge: Helping countries achieve Learning for All

Building strong education systems in the developing world is fundamental to development and growth. And since the agreements on Education for All in 1990 and the Millennium Development Goals in 2000, tens of millions of more children have gained access to schooling, a great achievement for developing countries.   Progress on student learning has been far less impressive, however. Many young people in low-income countries complete basic education without acquiring the fundamental literacy and numeracy skills, and even in middle-income countries, most students are unable to reach levels of learning that are expected in the OECD countries. Guaranteeing the human right to an education means little unless schooling leads to learning for all children and youth; growth, development, and poverty reduction depend on the knowledge and skills that young people acquire, rather than just the number of years that they sit in classrooms. For this reason, over the past few years **finding ways to promote learning for all children and youth** is now moving to the top of the education agenda, and learning and quality education have emerged as a top priority in the global consultations on the post-2015 development goals.

The need: A global knowledge base on education policies and institutions

Accelerating learning is no easy task, however. Learning depends not only on resources invested in the school system, but also on the policies and institutions that direct their use and on the quality of policy implementation. As the World Bank's *Education Sector Strategy 2020: Learning for All* argues, effective action to promote learning requires a more balanced analysis of the whole education system, aimed at identifying the binding constraints to learning, wherever they are. This **systems approach** recognizes that relaxing those constraints could mean increasing resources, but it may also require measuring student learning, or deploying better teachers to schools in poor areas, or tracking education expenditures better.

How do stakeholders know where to focus their efforts to improve quality? Making this choice requires actionable knowledge, but there is **far too little actionable, detailed knowledge about education policies and institutions available to policymakers and other education stakeholders**. What lies between an education system's inputs (the monetary and other resources that go into it) and its outcomes (such as years of education completed and learning acquired by students) is often a "black box".   But what is inside this box has a great deal of influence on what students learn. The results chain for learning includes the quality of policies and institutions and the quality of policy implementation; both of these are likely to be major influences on the quality of education delivered, which in turn affects student learning and other outcomes.

SABER and the Results Chain for Learning

| Inputs | → | Quality of **policies & institutions** | → | Quality of policy **implementation** | → | Quality and quantity of education delivered | → | Student learning & other outcomes |

Current focus of SABER ↑ (Quality of policies & institutions)

Future addition under SABER-UF ↑ (Quality of policy implementation)

Focus of complementary data-gathering initiatives ↑ (Quality and quantity of education delivered)

SABER's current focus is on the first box—policy and institutional quality. Policies have received some attention from data collection efforts in the past, but typically these efforts have covered a relatively small number of broad policies such as compulsory schooling or school fees. There is a lack of data and information on policies at a more granular level, of the type that policymakers and stakeholders need when they are making concrete choices on how to promote education and learning for all. SABER is working to fill that gap, as a complement to other initiatives that focus more intensively on the quality of policy implementation and the quality of education at school level.

The response: New data and system analyses from SABER

The World Bank and its partners are helping countries to fill this knowledge gap and guide education reforms. The new SABER program collects comparable data on the policies and institutions of education systems around the world and benchmarks them against good practice. SABER's aim is to give all parties with a stake in educational results **a detailed, objective, up-to-date, easy-to-understand snapshot of how well their country's education system is oriented toward delivering learning**, based on measures that can be easily compared across education systems around the world**.**   In each education policy area, SABER contributes in three ways:

- **Analytical:** *SABER highlights key policy choices for stakeholders.* The SABER data collection and analysis in each domain are built around an analytical framework that highlights for policymakers and other stakeholders the most important (and actionable) policy choices to spur learning. Each analytical framework is based on an in-depth survey of the most reliable global evidence, especially from impact evaluations and other rigorous research, and the experience of leading experts advising the domain team.

- **Descriptive:** *SABER provides new data on policies and institutions.*   In each national or subnational education system it covers, SABER collects comparable, well-defined, disaggregated data on policies and institutions. These data—on hundreds of system features—will be made publicly available to allow all governments, researchers, and other stakeholders to learn from them.

- **Evaluative***: SABER assesses and benchmarks education policies and institutions.* SABER also uses these new data to assess the level of development of policies and institutions in specific subsystems, or domains, of each education system. Each policy area is rated on a four-point scale, from "Latent" to "Emerging" to "Established" and "Advanced." These ratings, which are

summarized in country reports, highlight a system's areas of strength and weakness and also help to identify other systems to learn from.

SABER's systemwide coverage

The SABER initiative covers a range of **education policy domains that span the breadth of the education system**, as shown in this figure.

The domains of SABER
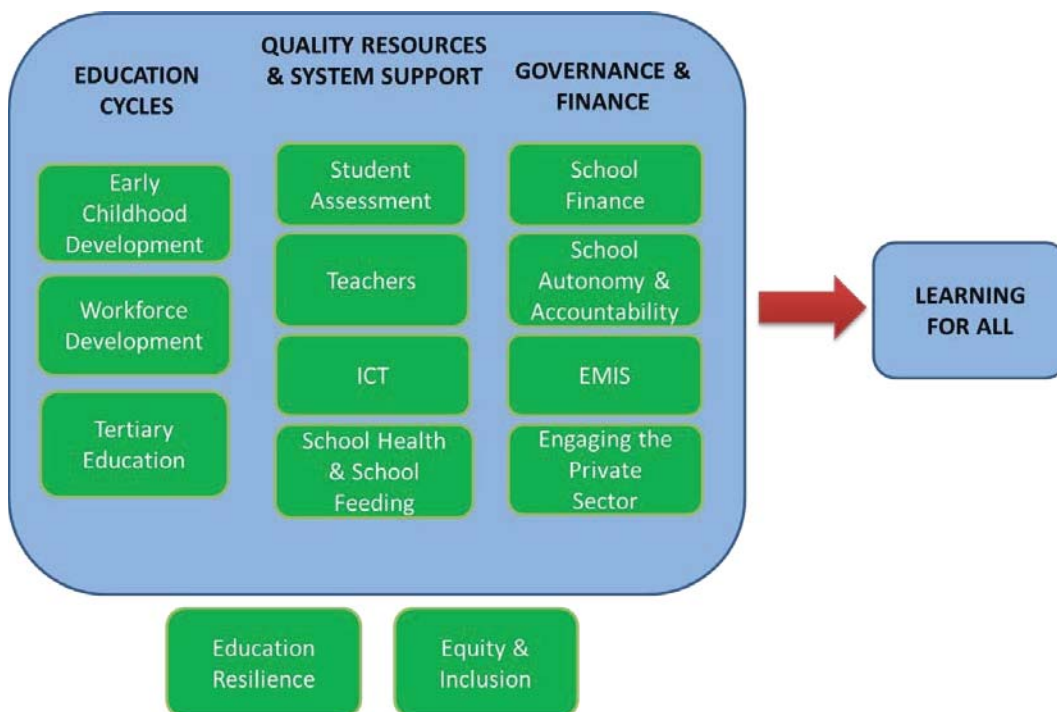
The new information provided by these domains will allow stakeholders to make better-informed decisions about education reforms and interventions by determining where the results chain from inputs to learning is breaking down.

How SABER adds value for education stakeholders

SABER strengthens the knowledge base for evidence-based policymaking in education. Until now, there has been a lack of comprehensive, systematic, detailed actionable information on developing-country education policies. Access to SABER information will benefit governments, World Bank staff, and other partners and education stakeholders in several ways:

- **Helping countries identify priorities for reform:** The analytical frameworks and evidence base on what policies and institutions matter most can serve as a foundation for policy dialogue about the reforms and programs that are most likely to improve learning and other outcomes, overall and within specific domains.

- **Providing progress metrics:** The SABER data and benchmarking help stakeholders to evaluate system development and track progress over time, and can be incorporated into projects and other operational work.

- **Promoting cross-country learning:** The detailed SABER database on policies will make it possible for stakeholders to identify comparator countries, both developing and developed, that can offer the most applicable lessons for a country facing a given set of challenges.

- **Enabling policy-focused analytical work:** Once there is a critical mass of data available, SABER will make possible a new area of rigorous research and other analytical work on the effects of education policies and institutions, not only within a given domain and country but across domains and countries.

SABER's approach

SABER's approach is designed to make it as useful a tool as possible for those who are trying to strengthen education systems, whether they are government officials, operational staff from the World Bank and development partners, or other education stakeholders.

- **Complementary with other knowledge tools:** SABER is meant to complement, rather than replace, other key sources of information, such as student assessments, impact evaluations, and school surveys. Knowledge generation and dissemination should therefore continue on those fronts as well.

- **Focused on providing actionable guidance:** Since the goal of SABER is to equip motivated education policymakers and stakeholders with the knowledge to improve learning results, the focus is on providing the framework, descriptive data, and evaluative information cited above. Unlike some other multicountry policy assessment tools, SABER does not produce rankings of national or state education systems—even within education policy domains—since the goal is not to "name and shame" less developed systems.

- **Publicly accessible:** The SABER data will be publicly available, and users who want to use the information to carry out their own analyses and rating exercises are encouraged to do so. While the evaluative ratings of policies are useful to many stakeholders, SABER's intent is to empower stakeholders with information, not to require them to accept a particular rating system.

SABER's current status and future directions

Since SABER's launch in early 2011, the SABER team and their partners within and outside the World Bank have moved quickly to take advantage of these new tools. SABER has collected data for country reports in over 100 countries, in one or more domains each, as well as in quite a few subnational jurisdictions. By mid-2013, SABER will have produced over 100 country reports.

Building the SABER knowledge base will involve a long-term process of continuous improvement and expansion. As the final section of this paper indicates, there is much still be done to develop and extend it—especially by applying SABER to new countries and domains, extending it from policies and institutions into policy implementation, and analyzing the data that SABER is producing. Nevertheless, the new frameworks, data, and benchmarking are already being used to help improve the quality of education policies and institutions in countries. This paper describes in detail how SABER works, how countries and partners may benefit from what it has to offer, what impacts it is having so far, and where it is going.

Introduction

This paper explains what the Systems Approach for Better Education Results (SABER) is, how it works, and how it can help countries gauge their level of educational development and identify priority areas for strengthening education systems. Building strong education systems is fundamental to development and growth. Providing access to quality education not only meets a basic human right, but also serves as a strategic development investment. Expanding educational attainment and achievement leads to healthier families, more involved citizens, and a more productive and creative workforce, among other benefits.

For education systems to deliver these benefits, they need to be able to **promote both schooling <u>and</u> learning**. This distinction is a crucial one. Growth, development, and poverty reduction depend on the knowledge and skills that people acquire, rather than just the number of years that they sit in classrooms. At the individual level, while a diploma may open doors to employment, it is a worker's skills that determine his or her productivity and ability to adapt to new technologies and opportunities. Knowledge and skills also contribute to an individual's ability to have a healthy and educated family and engage in civic life. And at the societal level, the level of skills in a workforce is a better predictor of economic growth rates than are average schooling levels (Hanushek and Woessmann 2007; Hanushek and Woessmann 2009).

In just the past few years, **the development education community has come together around the vision of promoting <u>learning for all</u> children and youth**. Getting tens of millions more children into school since the agreements on the Education for All (EFA, in 1990) goals and the education Millennium Development Goals (MDGs, in 2000) has been a great achievement for the developing countries, and for their partners (UNESCO 2011). But emerging evidence shows a pressing need to accelerate learning. Many young people are completing basic education without acquiring the most basic literacy and numeracy skills, and even in the more developed middle-income countries, the large majority of students are unable to reach levels of learning that are expected in the OECD countries. For this reason, there is now wide recognition—embodied in the recent strategies of numerous donor partners, for example—that finding ways to accelerate student learning must now move to the top of the education agenda.

The World Bank's **Education Sector Strategy 2020,** *Learning for All*, has promoted this shift with its strategy to advance learning in the developing world over the next decade. The strategy is based around two main thrusts: **encouraging country-level reforms of education systems, and building a global knowledge base that can guide those reforms**. Increasing our knowledge of what works and what system reforms to promote is crucial, because increasing the quantity or even quality of inputs will not be sufficient to fulfill students' learning potential. Resources do make a difference in getting children into school, particularly at early stages of educational development (Duflo 2001). But more funding is not enough. Especially when it comes to promoting learning—rather than just getting children into the classroom—research has found that what matter is not just the quantity and quality of resources, but perhaps more importantly a system's ability to transform those resources efficiently into better learning outcomes (Hanushek 1986; Hanushek 2003; World Bank 2003; Kremer and Holla 2009). Whether a system can do that depends on its capacity to formulate policy, set standards, implement quality assurance, assess student performance, manage human and financial resources, and take advantage of intergovernmental and external partnerships (Fullan 2005; Mourshed, Chijioke, and Barber 2010).

For these reasons, **the Bank is developing new knowledge approaches to help evaluate these system capacities and guide education reforms. The SABER program is a flagship in this effort.** SABER provides detailed system frameworks and benchmarking of country capacities ("system tools") in a wide array of education policy domains, from early childhood development (ECD), student assessment, and teacher policy to equity and inclusion, tertiary education, and skills development, among others. These SABER tools will analyze the "missing middle" of intermediate outcomes, illuminating the part of the learning results chain that lies between inputs and learning outcomes. Or to put it another way, imagine an education system as an iceberg, in which the visible, above-water portion is the resources going into the system and the number and skills of students coming out of it; then SABER will map out the vast bulk of the iceberg, all the system policies and institutions and implementation that currently lurk invisibly under the water. This information will allow policymakers and civil society organizations to make better-informed decisions about education reforms and interventions by determining where the results chain is breaking down. And by benchmarking progress against international best practices, the tool will highlight areas of strength and weakness, as well as identify successful reformers whose experience can inform education policy and practices in other countries.

SABER's aim is to give all parties with a stake in educational results—from students and teachers and parents to policymakers, business people, and political leaders—**an accessible, detailed, objective snapshot of how well their country's education system is oriented toward delivering learning**, based on measures that can be easily compared across education systems around the world. This knowledge will help policymakers and stakeholders understand their own education systems better and make better informed choices on actions to improve their systems. Over the next several years, as SABER repeats its data collection and produces updated ratings, it will reveal the progress that education systems are making in improving their education policies and capabilities. Finally, as SABER is extended across more education systems, it will facilitate efforts by policymakers and researchers to identify relationships between policies and results, carry out well-structured research to diagnose constraints and infer policy opportunities, and identify actions that will strengthen system performance and advance learning for all.

To fulfill this vision, SABER is being built around several key principles:

- **Quality and reliability:** generate data, indicators, and analysis that can serve as a reliable basis for making education policy
- **Scope and coverage:** achieve the broad scope within countries needed for systemwide analysis, and achieve the global coverage needed for cross-country learning and benchmarking
- **Timeliness and relevance:** provide data and analyses on important, actionable areas of policy, and in time to make a difference
- **Transparency and openness:** be transparent about data, methodologies, and ratings criteria, share the knowledge base freely, and be open to contributions by partners
- **Alignment and coherence:** pay attention not only to the quality of the individual parts of the education system, but also whether they work together to achieve learning for all
- **Efficiency and sustainability:** do all this cost-effectively, and in ways that make long-term sustainability possible

To ensure that SABER embodies these principles, the team is taking a **phased development approach** to implementation. In each SABER area, the team creates a strong evidence base and analytical framework, pilots the instruments and data collection, tests out the rubric for rating system development, and discusses its data and analyses with governments and other stakeholders before scaling up. Because SABER is intended to be a long-term, **regularly updated** source of data and analysis,

the indicators and benchmarks will be revised periodically to take advantage of the best new information available about the important policies and institutions in each policy area. SABER's open structure and its frequent use in policy dialogue and project work will subject it to constant market-testing, which will help keep it timely and relevant. Reliance on **partnerships**—in raising resources, collecting data, and interpreting and applying the results—will also help ensure that SABER remains relevant and useful to the broader education development community. As more data become available, SABER's indicators and the analytical framework underlying them will be **tested analytically**, to inform the continuous improvement. Finally, lessons from implementing each set of SABER instruments in one group of countries can be applied to improve the data collection and analysis in other countries and with other SABER instruments, as the program increases its scope and country coverage.

SABER is still in the early stages of that long-term process of continuous improvement and expansion. There is much still to be done, but the new frameworks, data, and benchmarking are already helping to improve the quality of education policies and institutions in countries. This paper describes in detail how SABER works, how countries and partners can benefit from what it has to offer, what impacts it is already having, and where it is going.

SABER goals and scope

What is SABER?

The **Systems Approach for Better Education Results (SABER)** is an ambitious initiative to collect and disseminate comparative data and knowledge on education policies, to help countries systematically evaluate and strengthen their education systems. The World Bank is working with partners around the world not only to collect the information, but also to develop and apply diagnostic tools that evaluate the quality of education policies and benchmark them against evidence-based global standards. SABER fills a gap in the availability of policy data, information, and knowledge on what matters most to improve the quality of education and help countries progress toward learning for all.[1]

Why is SABER needed?

SABER allows education leaders and stakeholders to see inside the black box between education inputs and outcomes, so that they can use evidence to strengthen education systems. SABER is part of a broader movement to improve the measurement of the quality of services provided or overseen by government.[2] Education systems are historically most effective at tracking *inputs*, such as the amount of money spent on education, and the quantities of schools built and teachers supplied. In recent years, they have become more effective at tracking the education system's *outcomes*—most notably enrolments, but increasingly also school completion and even learning in some cases. But knowing inputs and outcomes is not enough: research shows that the amount of spending on education explains only a small part of any gains in learning outcomes. What matters even more is *how efficiently resources are used to improve learning, which depends fundamentally on education policies and their implementation*.

---

[1] For a list of Frequently Asked Questions (FAQs) about SABER, please see Annex 1.

[2] For a summary of World Bank initiatives that seek to do this, see the materials from the World Bank seminar on "Measuring Government: Can It Be Done and Is It Helpful?" (October 11, 2011).

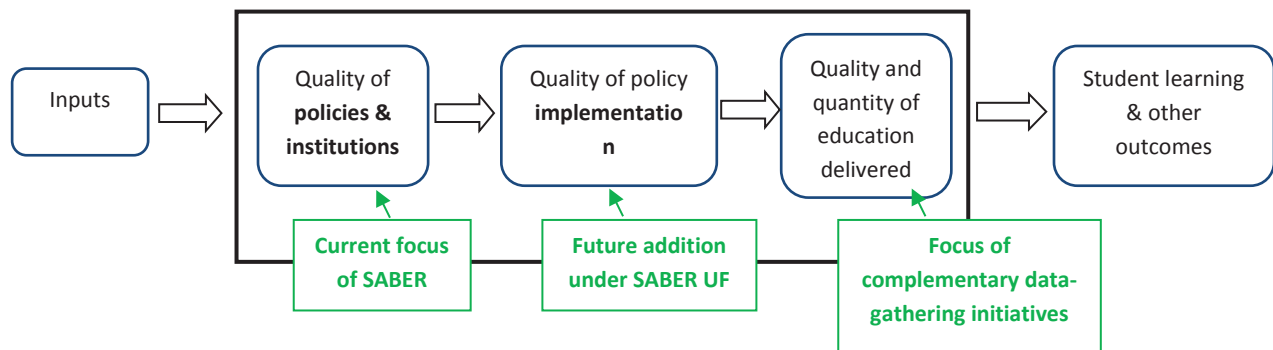*Figure 1: SABER and the Results Chain for Learning*

Figure 1 illustrates the three key links in that "education results chain"[3], as well as the role of SABER in illuminating it:

- **Policies and institutions:** At the core of SABER's current effort is an initiative to gather new data and information on the policy choices that countries have made in education. One key element in an effective education system is likely to be a strong framework of education policies and institutions[4]. ***SABER is now assembling the first detailed, disaggregated database of education policies and institutions in developing countries, and indeed globally***. The database will allow countries to evaluate and track the development of their policy frameworks, and also to compare with and learn from top and rising performers internationally. The policy indicators in each domain are selected on the basis of a domain-specific analytical framework paper that reviews "What Matters" most (in terms of policy and institutional choices) for promoting learning.

- **Policy implementation:** Of course, education outcomes depend not only on the quality of the policy and institutional framework but also on whether those policies and institutions are implemented effectively at the local and school level. In its next stage of development, SABER will also pull together tools and data for evaluating the quality of implementation at these levels, which includes aspects of governance of service delivery (see, for example, Amin, Das, and Goldstein 2008; Fiszbein, Ringold, and Rogers 2011). Although primary data collection through SABER will continue to focus primarily on cataloging and understanding policy and institutional frameworks, numerous studies over the past decade have begun pulling together data on implementation, and SABER will catalog these data in ways that are useful to education stakeholders.   Teams are also developing and testing out new ways to measure policy implementation (see discussion below).

- **Service delivery:**   A third link in the causal chain is the quality of education delivered in the classroom, as experienced by the students. The quality of policies and institutions and of policy

---

[3] The results chain illustrated here is simplified, of course. Policies and their implementation also affect the quantity and quality of inputs, and not just the efficiency of conversion of inputs into outputs and outcomes, but the simplified representation is used to emphasize the latter point.

[4] Because of the lack of detailed education policy data, it has been difficult to test whether this is in fact true. However, if policies are bad, good implementation seems unlikely to promote learning for all; indeed, if policies are misguided enough, better implementation may lead to poorer outcomes. Thus at a minimum, countries are likely to need policies and institutions of reasonable quality.

implementation both influence education quality, but they are conceptually distinct. While systematic classroom observation that generates comparable data on education quality remains rare, in recent years it has become more common. SABER will not generally collect these data, but where they are available, SABER will be an excellent tool for analyzing how the policies and institutions affect the quality of education actually delivered in classrooms, as well as student learning and other outcomes.

This paper focuses primarily on the first of these areas—how SABER captures and measures formal education policies and institutions—because that has been and will continue to be the main focus of the SABER initiative. At the same time, SABER should be recognized a tool for illuminating part of the results chain, meant to be used in tandem with other knowledge tools. These complementary tools include not only surveys for measuring the quality of service delivery, mentioned above, but also student learning assessments (which track one of the key outcomes) and impact evaluations (which analyze rigorously how changes in policies, programs, and institutions affect quality and outcomes).

What areas of education policy does SABER cover?

SABER's aim is to produce a detailed catalog and benchmarking of policies and institutions in all the major areas of education policy, as a tool for improving education outcomes. Initially, the team has focused on a set of education policy areas – or "domains" – that seem most important to the largest number of country education systems at various levels (see Figure 2, where each green box represents one domain). These domains have been selected based both on demand from regional teams and government counterparts and on an assessment of which policy areas are most important for promoting Learning for All.

Figure 2: SABER Domains
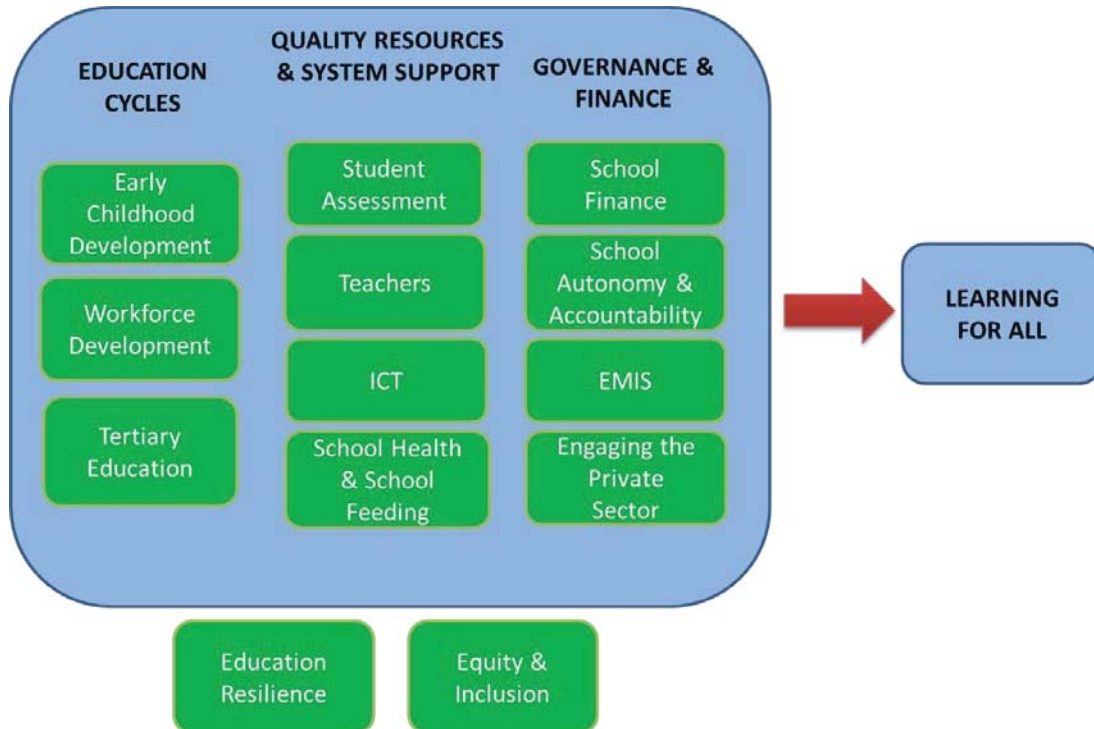
The first three domains represent **education cycles**—ECD, workforce development, and tertiary education. Because the other domains focus primarily on primary and secondary education, these domains provide comprehensive, more specific analyses of policies and institutions at these levels. The next four domains analyze **resources and support** for the provision of quality education, from student assessment (which allows the system to assess whether it is reaching its goals) through the core resource of teachers to ICT and school health and school feeding. The third group of domains includes those related to **governance and finance**: the system of school financing, the policies governing engagement with school-level governance and private-sector schools, and the EMIS system, which provides data and information necessary for effective system management. Finally, there are two **cross-cutting domains**: education resilience, which focuses on how the system works for children and communities that are most at risk from fragility or conflict, and equity and inclusion, which places a spotlight on whether all the elements of the system are sufficiently focused on the "all" of "learning for all." A majority of these domains have completed their analytical frameworks and are producing country reports, while the others are under development. Detail on each domain is available through the SABER website, at www.worldbank.org/education/SABER.[5]

What core principles guide SABER?

To fulfill the vision of empowering users with actionable new policy and institutional knowledge, SABER is designed around several core principles:

- **Quality and reliability:** generate data, indicators, and analysis that can serve as a reliable basis for making education policy

- **Scope and coverage:** achieve the broad scope within countries needed for systemwide analysis, and achieve the global coverage needed for cross-country learning and benchmarking

- **Timeliness and relevance:** provide data and analyses on important, actionable areas of policy, and in time to make a difference

- **Transparency and openness:** be transparent about data, methodologies, and ratings criteria, share the knowledge base freely, and be open to contributions by partners

- **Alignment and coherence:** pay attention not only to the quality of the individual parts of the education system, but also whether they work together to achieve learning for all

- **Efficiency and sustainability:** do all this cost-effectively, and in ways that make long-term sustainability possible

The methodology outlined in the next section is aimed at ensuring that SABER embodies these principles.

---

[5] For Fiscal Year 2013, the Education Sector Board has focused the Bank's efforts on developing and rolling out six of these domains: Early Childhood Development, Student Assessment, Teachers, School Finance, Tertiary Education, and EMIS. But the full set of thirteen domains represents the medium-term vision for SABER, and all are under development already.

The methodology of SABER

How are SABER data collected and analyzed?

The SABER policy data-gathering and evaluations are designed to be carried out relatively quickly and cost-effectively. Because the goal is to analyze the quality of the policy and institutional framework, rather than do a detailed assessment of school-level conditions, most of the data-gathering can usually be done centrally by knowledgeable experts, at either the national or provincial level. While there is some variation across domains, the typical model for data collection and analysis is as follows:

1. **"What Matters" paper:** The SABER domain team writes a "What Matters" paper—such as "What Matters Most for Teacher Policy"—that surveys the best evidence and experience in the domain and uses that survey to identify the elements of the policy and institutional framework that matter most for improving education outcomes, including learning.[6]

2. **Indicators and scoring rubric:** The team uses the "What Matters" paper to identify indicators of policy and institutional development, as well as the data source for each indicator.   It then develops a rubric for combining the indicators to come up with ratings of the country's progress within each domain. These indicators and scoring rubrics will generally be objective, rather than subjective, to ensure cross-country comparability and replicability.

3. **Data-collection instrument:** The team develops a data-collection instrument appropriate for collecting data for the policy and institutional indicators. As noted above, this is not an instrument for a large-sample survey, but essentially a survey for one respondent—an experienced principal investigator in the country—to fill out using information from key informants, documents, and other sources.

4. **Data collection:**   In the typical model, an experienced principal investigator will collect the policy information and data necessary to fill out the data-collection instrument, by drawing on his or her knowledge of the system and on government contacts. Data collection can usually be completed within a few weeks in this way. An alternative approach used in some domains is to convene a workshop of experts, including government officials, and use that group process to collect the evidence and code data. In either case, data sources are clearly identified and made public when the data are posted.

5. **Analysis:** The domain team uses the data to analyze how developed the country's education policies and institutions are, from the perspective of achieving key education goals. In the process, it will also generate benchmarks of progress in those specific areas against other countries or provinces. In some domains, these evaluations will be embedded in a more in-depth report discussing policy options and relevant experiences from other countries. While the principal investigator in each country may carry out the initial analysis, the central SABER team for each domain is responsible for completing the analysis and ensuring cross-country comparability.

6. **Validation and discussion:** As noted above, the team presents the data to World Bank regional team leaders for validation with government officials, to ensure that SABER reports the country's policies and institutions correctly. Any corrections to the information on which the

---

[6] The first of these papers to be published was "What Matters Most for Student Assessment Systems: A Framework Paper" (2012).

analysis is based will be incorporated before publication. In addition, the team or, preferably, the Bank regional staff member) will discuss the resulting report with government counterparts. Finally, the report is reviewed and approved by the World Bank's education sector board before it is finalized and made public.

7. **Publication of analyses and data:** Both the country report and (in the near future) the data underlying it are made public on the SABER website. This ensures that users can benefit from expert analyses, if they choose to, but that they can also carry out their own analyses and benchmarking using the SABER data.

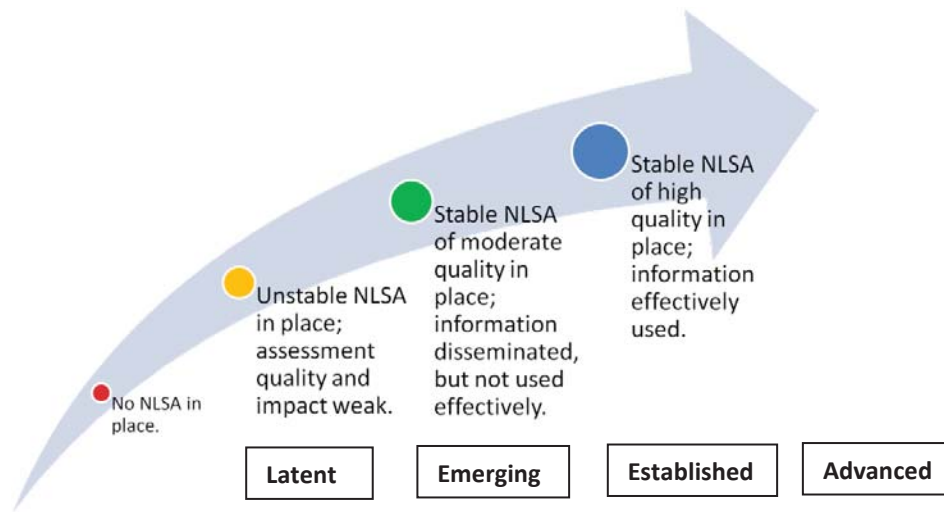## How does SABER data illuminate policy choices?

SABER shines a light on education policy frameworks by providing three types of information: analytical, descriptive, and evaluative.

Analytical: Even before data collection takes place, ***SABER highlights key policy choices*** for stakeholders. The SABER data collection and analysis in each domain are built around an analytical framework that highlights for policymakers and other stakeholders the most important (and actionable) policy choices to spur learning. The analytical framework is based on a survey of the best global evidence, including results from rigorous impact evaluations, other research, and professional standards. The framework can be used to structure conversations about system development and potential reforms, drawing on the wealth of information that policymakers and others already have about the education system.

Descriptive: Another major contribution of SABER is its ***systematic collection and cataloguing of data and information on the policy and institutional choices*** that countries have made in each education domain. The choice of data to collect is based on thorough background research aimed at identifying the policies and institutions that are most likely to be related systematically to student learning. These data will allow policymakers or concerned citizens in any country to see at a glance what their official policies and institutions are in the area of teacher policy. This ability to make policy choices more transparent, in a detailed way, should be a major contribution of SABER. Very few people (even within an education ministry) know all of their country's policies in a given domain, let alone across all domains; and far fewer will know other countries' policies that well. SABER will empower them by making the invisible parts of the education sector more visible.

Evaluative: A third contribution of SABER is to use these data to ***evaluate the level of development of education policies and institutions***. In each domain, the team uses extensive literature surveys and practitioner input to identify policy and institutional choices most often seen in the best-performing and fastest-improving education systems. (See below for a description of this process.) The team codifies these good-practice choices in a set of indicators and then uses the data collected for those indicators to evaluate how advanced each education system in the given area. SABER has been standardized so that progress along each dimension is summarized in a 4-point scale, ranging from Latent (the lowest level) through Emerging and Established to Advanced (the highest level), based on international benchmarks.   To make concrete what these levels signify, consider the following example from SABER-Student Assessment.

Progress toward developing national large-scale student assessments is captured in four levels, each based on a number of underlying indicators. This is just one of many different aspects of the assessment system in which SABER-Student Assessment provides a rating of the level of development. By surveying these ratings and the indicators on which they are based, stakeholders in each country can identify priorities for policy reform, or at a minimum for further investigation. The World Bank has already found this evaluative information to be useful for structuring conversations about education policy with government officials and civil society.

It is important to stress that while these ratings of levels of development represent judgments by experts in the field about what policies matter, based on the best evidence available, that evidence is usually rather limited. As SABER enriches the database on education policies, it will be possible to carry out new analyses of whether in fact the policies identified by the SABER team and the experts it has consulted do in fact contribute to improvements in student learning. SABER is an ongoing process of data collection, review, and revision, and ***the accumulation of data and evidence should allow continuous improvement of the SABER approach over the next several years.***

How do we know that SABER policy data and analysis are valid?

The impact of the SABER policy data initiative ultimately depends on the accuracy of its data and evaluations. The team has therefore put in place numerous **quality-control measures**:

- **Analytical framework:** As noted above, all of the work in each domain is built on the framework paper summarizing evidence on "What Matters" in that domain—for example, "What Matters in Learning Assessments" or "What Matters in Workforce Development". Each paper draws not only on the extensive experience of the task team for that domain, but also on input from leading academics and practitioners who serve on domain-specific advisory committees. In addition, before it is published in the "What Matters" series, each paper will undergo a high-level internal review. The same is true for the ratings rubrics and data-collection instruments.

- **Data collection:** Before being approved, data-collection instruments are reviewed to ensure that they meet common guidelines to ensure accuracy and comparability. To aid in the

preparation process, the team has prepared guidelines for SABER domain teams to consult as they develop their instruments and work with consultants to collect data.

- **Data validation and updating:** To validate their responses to particular questions, consultants are required to append documents detailing the policy in question and corroborating the responses. The data are discussed with government counterparts, to ensure that policies and institutions are represented correctly. And even after publication, the open data nature of the SABER initiative will provide an opportunity for longer-term crowd-sourcing of data validation. The website will invite users to notify the team of any apparent errors in coding and to submit their own evidence. In this way, it should be possible to update information even between rounds of data collection (since data collection in a given domain/country combination will likely take place only every few years). While a full ratings update will likely have to wait until the team's next round of data collection in the country, in the interim the SABER website will allow users to submit information on any policy and institutional changes, so that the public can be made aware of policy improvements since the previous ratings were published.

- **Benchmarking:** After the data are gathered and analyzed in a consistent way, the assessment process is also harmonized to ensure that a rating of "emerging" or "established" on a particular dimension has the same meaning in one country or state as in another. To ensure consistency, it is the core team—and not the consultant recruited for in-country data-gathering—that determines the final ratings, in accordance with a scoring rubric that specifies detailed ratings criteria. Bank staff and government counterparts then discuss the ratings and report in-country, so that any necessary adjustments (based on the evidence) can be made before the report is finalized.

Box 1 provides an example of how this process operates in one domain, SABER-Workforce Development. Over the longer run, the team will also assess the validity of the data through *statistical analysis* of which policy variables and ratings actually help predict outcomes. As noted earlier, the selected indicators are based on the best research and expert judgments currently available, as summarized in the *What Matters* papers. But until now there has been no comprehensive policy database in education to allow testing of different indicators against each other and in combination with other factors affecting outcomes (such as policy implementation quality and household characteristics). Once SABER has collected enough policy data for enough countries, we can begin to fill this gap.

How do we know whether these education policies are actually implemented?

This paper has focused largely on SABER's policy-measurement frameworks and tools, because they are SABER's main contribution—in an area where very little information is currently available—and the major focus of the SABER initiative. These tools will offer a new look at the quality of formal policies and institutions, as well as indicators the government's seriousness about making them a reality. Without strong policies and formal institutions, the education system is likely to be rudderless at best, and badly off-course at worst. High-quality implementation of poorly designed policies is not likely to improve student learning.

At the same time, the most well-designed policy and institutional framework will not succeed in improving learning outcomes unless it is implemented effectively down to the school and classroom level. This raises a key question: how will users of SABER data know whether or not the policies catalogued in by SABER are actually being implemented? Or to use terminology that is frequently used in the field, how can we assess whether the *de facto* policies match the *de jure* policies?

Box 1: Quality control at the domain level

Like all SABER domains, the SABER-Workforce Development (WfD) team developed a Data Collection Instrument (DCI) to gather data in a systematic and comparable manner across countries. The team consulted with a large number of experts from around the world through the development and piloting phases of the DCI to incorporate a wide range of perspectives and experiences. Since a key goal of the SABER-WfD is to provide information for analysis, benchmarking and comparison across countries, the World Bank team developed standard implementation protocols, as well as a training package for Principal Investigators (PI) to guide them through the implementation of the SABER-WfD tool.

The quality of the data collected is critical for the reliability of the benchmarking exercise. During the data collection phase, there are work-in-progress meetings mainly involving the PI, the Task Team Leader (TTL) and the Bank SABER-WfD team. Members of the national team should also be engaged for data validation at critical junctures in the process to ensure the answers selected in the DCI are accurate. The DCI is structured so that the PI provides notes and a reference(s) that supports the answer for each question. This allows for any discrepancies to be verified against the evidence provided and amendments to be made as necessary. The PI then submits the completed DCI to the SABER-WfD team allowing for information to be analyzed and scored following established scoring criteria and decision rules (rubric) corresponding to four levels of maturity: latent, emerging, established and advanced. The DCI must be fully completed by answering all the questions on every worksheet. The final sheet of the workbook contains a list of the documents and sources consulted, including the informants and the basis of their credibility as informants. This documentary evidence base forms an integral part of the output from the data collection effort.

The SABER-WfD country reports summarize the benchmarking results succinctly, following a standard format. Each report includes contextual information on the WfD system and detailed information on each of the nine SABER-WfD Policy Goals. The country report is reviewed with key in-country stakeholders as part of the quality assurance arrangements. Following this meeting, the PI revises the report as needed. The TTL uses the completed report as appropriate in policy dialogue with the government. Once the country report is approved for dissemination by the World Bank's education sector board, the report, DCI, and supporting evidence are made available through the SABER database at www.worldbank.org/education/saber.

One response is that **the core SABER policy indicators being collected in each domain incorporate indicators that are useful for evaluating policy implementation** (or probability of implementation), to the extent that these can be collected cost-effectively and quickly at the central level. Some of the institutional and even policy data will serve as indicators of whether a given policy is likely to have an effect on the ground. They will ask whether countries have in place institutions and mechanisms to implement policies. For example, suppose that a country legally requires that early childhood centers meet certain safety standards, but has issued no implementing regulations and has set aside no budget for inspections. In this case, even without a survey of ECD centers, it is reasonable to conclude that the government is not very committed to the safety-standards policy. Each SABER policy domain includes

indicators of this sort, and the evaluative ratings of SABER generally incorporate some of those indicators; this means that the core SABER tools do partially pick up whether it should be possible to begin mapping out implementation with these core tools.

In the next phase of SABER, with support from the proposed SABER Umbrella Facility (SABER-UF), one goal is to further illuminate the quality of policy implementation through two mechanisms:

- **SABER teams will develop new policy implementation indicators and data-collection approaches.** Where feasible, some SABER teams will explore appropriate methods for assessing whether or not policies are being implemented. There may not always be cost-effective ways of getting reliable data on implementation within the SABER framework, but the teams are exploring the possibilities. In some cases, SABER has already been able to find opportunities and funding for extensive school- and household-level data-collection efforts focused on policy implementation. Two examples are: (1) the SABER-Teachers effort to collect data on policy implementation in a large representative sample of Indian schools, and (2) SABER-School Autonomy and Accountability's partnership with the OECD to collect data in schools in Thailand that are participating in PISA. In other cases, SABER teams will not be involved in the data collection but will provide survey instruments and methodologies for governments and other partners that want to collect data on implementation themselves.

- **SABER will link users to existing measures of quality of implementation and service delivery in education.** Many aspects of implementation quality—especially at the school level—can be measured reliably only through surveys of large representative samples of students, educators, households, employers, or others. For example, finding out whether budgeted funds are actually reaching schools and being spent on the intended line items (as specified by policy) requires something like a Public Expenditure Tracking Survey (PETS), a rigorous, time-intensive survey carried out by teams visiting a representative sample of hundreds of schools. And reaching definitive conclusions about whether teachers are teaching in the classroom as envisioned by policy requires that large sample of teachers be observed by trained evaluators. Such representative in-person surveys of education stakeholder populations typically cost tens of thousands of dollars and require much more time to design, pilot, implement, and analyze than the SABER policy surveys. Research of this type is generally going to be outside the scope of the typical core SABER policy data-collection effort, which involves a relatively low-cost, quick-turnaround study – implemented in a rigorous, comparable way in a large number of countries – to illuminate the education policy framework.

  Fortunately, the amount of data on the quality of governance in service delivery, as well as the quality of service delivery itself, has increased sharply over the past decade. For example, there are now PETS surveys and other quantitative service delivery surveys for quite a few countries, as well as an increasing number of countries with data based on classroom observation of teachers. In sub-Saharan Africa, the World Bank and the African Economic Research Consortium are expanding and systematizing these efforts with a major new Service Delivery Indicators Initiative. The initiative aims to collect school-level data on policy implementation and education quality regularly in 15 to 20 African countries (Bold and others 2010). SABER will capitalize on these initiatives, by linking from the policy framework to these existing indicators of policy implementation.

Under either of these approaches, it will likely take several years to collect data on school-level implementation in enough countries to allow rich cross-country comparisons. In the meantime, it may

be worthwhile to explore using off-the-shelf indicators of policy implementation that are already available for a large sample of countries. These indicators will provide only rough, temporary proxies for implementation quality, but they may still be useful in "discounting" the quality of policy frameworks in poor-implementation environments.

Uses and benefits of SABER

How do countries participate in SABER?

Countries choose to participate in SABER, and specifically in the domains that they find most valuable. The World Bank does not make participation in SABER mandatory as a condition for receiving Bank support, although a World Bank task team leader may encourage a country to participate, as a way of strengthening the knowledge base for the country program. As noted above, the rollout of SABER domains is typically financed out of the funds from the country programs of the World Bank or other partners, which ensures that SABER implementation has to be demand-driven. This means that some countries may participate in a suite of domains, to map out the overall policy framework in education, while others may choose to participate in only the one or two domains where they feel the greatest need.

Eventually, of course, the aim of the initiative is to have many countries participate in SABER, and in many domains each. This broad coverage will allow a more comprehensive, system-wide analysis of how policy design and implementation can better achieve Learning for All. The demand that has already emerged for SABER has been quite substantial, suggesting that this is not an unreasonable goal. The growth is encouraging: each additional country that participates in SABER not only gains information and insights for itself, but also increases the value of SABER to other countries by increasing the stock of knowledge on what policies are in place around the world.

Who will have access to the SABER data?

Everyone. SABER's goal is to empower governments, Bank staff, and the public with new data and insights that they can use in crafting and implementing education policies. As explained above, within each domain, the SABER team will be providing **benchmarking and ratings of the level of system development** toward achieving different policy goals. Within each goal area, the policies of the country being studied are implicitly benchmarked against the policies of top-performing systems and those that have improved most rapidly. SABER teams apply a common rubric to data gathered by consultants based in each country, to ensure global consistency of ratings criteria. Governments have already found these SABER policy analyses useful as indicators of where they may need to devote the most effort in improving their policy frameworks.

These evaluations are only one of the ways that SABER empowers stakeholders. Users who want to delve into the policy data themselves will be able—and in fact encouraged—to do so. In line with the World Bank's Open Data Initiative[7], all of the **disaggregated data on education policies and institutions will be made available on the SABER website**, so that stakeholders can make their own evaluations of their systems' level of development. Free access to data will allow stakeholders to choose their own comparator systems, since many education leaders have already identified systems that they judge to have succeeded in circumstances similar to their own and that they would like to learn from. These comparators may not be the countries regularly recognized as global leaders, but instead relatively high-performing systems in economies at the same development level or with the same structure. Analysis

---

[7] See data.worldbank.org for details.

of the detailed policy data from those systems will make it easier to investigate areas where there may be useful lessons.

Note that unlike some related efforts, SABER does not produce rankings of education systems, even within policy domains. Educational rankings do have their place: "naming and shaming" poor performers on student learning can catalyze reforms, for example. But SABER is predicated on the assumption that stakeholders are already motivated to improve learning results—perhaps because of results from student assessments or budgetary concerns—and what they need is the knowledge and tools to make those improvements. For this reason, SABER is envisaged as a tool to inform and align the efforts of the various stakeholders that are working to improve education systems—most notably, developing-country governments and civil society, but also the World Bank and other donor partners. Indeed, SABER will achieve long-term sustainability only as a partnership, because its financial sustainability and policy impact both depend on working together with others.

How does SABER bring value to governments and World Bank staff assisting them?

SABER is intended to be useful to the broader education community, but a key target audience is the development community in education, including especially governments and World Bank staff working on education issues. New SABER data on policies—disaggregated and actionable, and also integrated with existing data on the quality of policy implementation—will reinforce governments and operational staff in their work in three main ways:

- **SABER provides a knowledge base for analytical sector work, policy dialogue, and design of operations:** Not only does SABER provide data and analysis within each policy domain, but it also structures that information in a way that makes it easily comparable across countries, and then provides that comparator information from other countries. This helps Bank staff and government counterparts—not just in education ministries, but other ministries such as labor and finance—to identify areas where a country's policy framework should be strengthened, and where international experience shows that strengthening is possible. Even before any data are available, the analytical framework underlying each domain can serve as an excellent basis for policy dialogue—in that it provides a structured, comprehensive list of key policies and institutions that government will likely have to address. This dialogue can identify entry points for operational support to develop better policies and the capacity to implement them.

- **SABER makes possible targeted cross-country learning:** There has been a major push to promote South-South learning in the development world, and for good reason: in many cases, the most relevant recent experience is from other developing and emerging economies. SABER's policy database and analyses will directly serve as a vehicle for cross-country learning of all types, including South-South, by providing comparable information about what has been done in other countries and how it has worked. It will also provide a more comprehensive and systematic evidence base for identifying countries for more in-depth South-South engagement, for example through case studies or study tours.

Box 2: How SABER is being used operationally

This annex offers a non-exhaustive list of examples of SABER's early use in policy dialogue and design of operations in a variety of domains and countries.

- Through the READ Trust Fund, the Bank has provided technical assistance based on **SABER-Student Assessment** to **Angola, Armenia, Ethiopia, Kyrgyz Republic, Mozambique, Tajikistan, Vietnam,** and **Zambia**. Countries have used this support to identify strengths and weaknesses of their assessment systems, to launch new assessment centers, to train assessment staff, to implement national and international large-scale assessments, and to develop materials for classroom assessments, among others.

- In **Nigeria,** SABER analyses were used to help states identify priorities for reform in preparation for the $150 million State Education Program Investment Project (SEPIP). The SABER team worked with regional colleagues to apply four SABER instruments (**Student Assessment**, **Teachers**, **School Autonomy and Accountability**, and **EMIS**) in Anambra, Bauchi, and Ekiti states, and there are plans to carry out further SABER analyses as part of the M&E for the project.

- In the **Middle East and North Africa region**, the **SABER-Teachers** team worked with MENA staff to carry out a regionwide analysis of teacher policy in six countries, taking advantage of the SABER-Teachers framework's ability to create comparability of analyses across countries. MENA then pulled together the individual SABER–Teacher country reports to give a broader picture of patterns in the region ("MENA Regional Synthesis on the Teacher Policies Survey: Key Findings from Phase I", 2011).

- In **Nicaragua**, the World Bank country team used **SABER-ECD** framework to help the government to diagnose and take stock of existing ECD policies and programs and set a baseline for a new national ECD policy. SABER complemented other efforts, such as a study tour to learn how Chile had put in place a comprehensive ECD system, and strengthened the knowledge base for the new National Commission on ECD. By late 2011, Nicaragua's Congress had approved a new "Política Nacional de Primera Infancia" that SABER had helped inform. The Bank has also used the SABER-ECD findings in designing a major ECD component of a new education project.

- In **Jordan,** staff and government counterparts used the **SABER-EMIS** framework and tool to analyze Jordan's education management information system as part of the implementation of the Education Reform for Knowledge Economy II (ERfKE II) project, in one of the first such operational applications in MENA and globally. The team leader in MENA reported that the SABER analysis had helped spur progress in this important area of education policy.

- In **East Asia,** multiple SABER domains were applied across many countries, leading to a regionwide conference with government officials and other development partners and a joint UNESCO-World Bank report on *Strengthening Education Quality in East Asia*.

- **SABER provides indicators for operations:** A challenge in operations of all types, including especially results-based operations, is to identify the appropriate results indicators. Ultimately, what governments care about in most cases—or should care about—is promoting student learning, but numerous studies have shown that it can take a considerable amount of time for

even well-designed reforms and innovations to increase learning. Moreover, it is usually difficult to attribute improvements in learning outcomes to any given set of reforms. SABER analyses may be able to suggest evidence-based intermediate indicators of policy quality that can be used as disbursement-linked indicators (DLIs) for results-based lending, because they are under the control of the government, and are expected to contribute to accelerated learning. More generally, they can serve as an indication of whether the sector is moving in the right direction, especially when combined with other indicators of progress in implementing the reforms at school level.

Within the World Bank, SABER is already being used as a basis for regional or country-level AAA in a number of countries, and also in the design of operations. Box 2 provides an illustrative list of ways in which SABER is already being used operationally at the World Bank, while Box 3 offers further ideas for possible SABER uses in Bank operations.

How does SABER benefit other education partners and stakeholders?

SABER is not meant merely as a tool for governments and the World Bank to use in their own discussions about education systems. As noted above, the data and analyses will be made public, and over the long term SABER's greatest benefit may be in informing the efforts of wider education community and public. Some of these potential beneficiaries of SABER include:

- **Civil society:** SABER's data and analyses will give civil society a systematic look at the inner workings of the policies and institutions that drive an education system. Combined with the local- and school-level work that documents the quality of service delivery in education—in which many NGOs are involved—SABER will help civil society groups understand where public spending is going, where it is being used well to promote learning, and where the results chain is weak and needs attention. Equally importantly, the systematic analysis and cross-national (and in some cases cross-province) comparisons can help provide an impetus and opportunities for change.

- **Donor partners:** Not only have other donor partners provided financial support for SABER, but they are also serving as partners in other key ways—by building support for applying the analysis in countries, and by helping apply the findings to operational work and policy dialogue. For example, AusAID has supported efforts to apply adapted versions of the SABER frameworks and tools in the South Pacific, a region in which it is particularly active. This kind of involvement will be crucial going forward: to achieve its full potential, SABER needs to be viewed as a useful tool for the education development community as a whole, not just the World Bank and governments**.**

- **Researchers:** The research community should also benefit from SABER, especially over the long run. The new information provided by SABER will make possible new empirical analyses, by enabling researchers to explore which policies and institutions are associated with higher-quality education and improvements in student learning, and in what contexts. One analytical question, for example, is which of the SABER indicators are good predictors of better outcomes; another is whether progress on policies in one area (as measured by better ratings) can compensate for weaker development in another. Until now, researchers focusing on education in developing countries not been able to carry out such analyses, because of the lack of

---

Box 3: Other SABER entry points in World Bank operations

Box 1 offers examples of ways in which SABER can be used operationally by any partners; here, we offer some additional possibilities for intervention points where SABER can contribute to World Bank operations:

1. **Strategic dialogue (pre-identification).** When a country is considering new education sector policy reforms, or is in discussion with development partners about future policy options or investments, the availability of SABER reports would provide an excellent foundation. Even if these SABER reports are not available, it could be possible to collect data and prepare SABER country diagnostic reports in time to inform the discussion.

2. **Project concept (identification).** Once a decision is taken to prepare a project or a reform that is related to one or more SABER domains, SABER diagnostic reports in those domains would serve as invaluable knowledge resources to have a full picture of those domains and to identify where the leverage points are for a reform.

3. **Project targets (KPIs and results matrix).** SABER goals or policy levers can serve as key performance indicators or as indicators in the results matrix for projects or reforms, and the diagnostic reports can provide the information needed for the baseline in policy or institutional reform projects or specific investments. SABER indicators can be measured during project implementation to assess whether the targets have been attained. For policy reforms (development policy operations), progress on SABER-measured policy goals or "levers" could serve as triggers. SABER measures, down to the indicator level, can also be used in result-based lending, including as metrics for such new instruments as disbursement linked indicators or P4R. They could also serve as key performance indicators for a specific investment operation.

4. **Project implementation and monitoring (supervision).** SABER analyses and ratings exercises can be conducted periodically to inform project supervision.

5. **Project completion.** Once a project that used SABER in its design or result framework is completed, SABER can be administered again to evaluate whether the objective was achieved.

---

comparable, disaggregated data on education policies.   Initially, the empirical research will likely need to focus on comparing countries' policies within the domains that have data on the most countries, but as data accumulate it will become possible to carry out cross-country analyses that bridge multiple domains and parts of the education system.

How does SABER relate to the World Bank's "Learning for All" strategy?

SABER is one of the major initiatives for implementing the World Bank's strategy for the next decade, *Education Strategy 2020: Learning for All.* The strategy is built around the common-sense idea that for education to benefit children and youth, it must result in their acquiring knowledge and skills. This view of the centrality of learning for all is consistent with the aspirations emerging from the global consultations on the post-2015 development goals framework, which emphasizes education quality and equitable learning for all.

How to achieve that? The Strategy calls for building two pillars of Learning for All:

- **Recognize education as a system, not just a collection of inputs:** While building schools, training teachers, and providing textbooks are all important, merely increasing inputs is not enough to improve student learning very much. As discussed above, evidence shows that how schools and school systems use those resources matters a great deal in driving learning. Any education system is a complex network of providers and beneficiaries, institutions, and individuals whose individual efforts and interactions determine whether the country is using resources effectively to advance learning for all. The strategy calls for governments and donors to focus much more on improving the non-input factors that drive learning—the information flows, accountability relationships, incentives, financing structures, and behaviors, both within and outside the formal education system, in the public and private education sectors alike. By covering the system as a whole, rather than focusing on schooling inputs in the public sector, SABER is able to analyze what matters most in each of the core domains of the education system. This allows it to help countries identify potential leverage points improving system performance and results.

- **Build the knowledge base for better education systems:** The strategy recognizes that knowledge is the crucial ingredient that the World Bank can provide for strengthening education systems. Without knowledge of where financing can make the most difference, financial assistance will achieve minimal results. And ultimately, the aid provided by the World Bank and other donors is only a fraction of the total spent by governments and households on education. The key for aid effectiveness is to create and apply knowledge in a way that improves the efficacy of all education spending in the country.

SABER directly strengthens both of these strategic pillars. First, it makes the system approach concrete. It covers—or will cover—policies in all of the key parts of the education system. These include not only the domains focusing on inputs (such as Teachers and ICT) and on schooling levels (such as Tertiary and ECD), but also domains that address other elements such as information and accountability (such as Student Assessment, as well as Autonomy and Accountability). Even in the input- and level-focused domains, SABER's ambit extends well beyond the traditional concern with inputs to encompass issues such as incentives, support, financing, and autonomy.

A system approach requires not only covering the key areas of the system, but also analyzing how these areas work together most effectively to provide quality education. Over the long run, the SABER data should strengthen the Bank's and counterparts' ability to do this analysis. One important question, for

example, is whether strengths in one area can compensate for weaknesses in another—for example, whether systems that select the most motivated, knowledgeable candidates into the teaching corps can manage with less robust accountability systems at the school level. An alternative hypothesis is that the system is only as strong as its weakest domain. As discussed below, doing these system-level analyses is on SABER's agenda for the longer term, once we have enough data and information from individual domains in enough countries to allow analysis of how the domains interact.

In keeping with the second pillar, SABER represents a significant step toward improving the knowledge base for system strengthening. As emphasized above, until now, there has been a lack of comprehensive, systematic information on developing-country education policies. SABER provides that information in ways that are most useful—that is, information built around robust analytical frameworks and benchmarked against comparator systems in both developing and developed countries.

Collaboration with partners on SABER

How is SABER collaborating with other partners that collect policy data?

SABER is not alone in collecting cross-country data on policies and institutions in education, of course. It builds on and complements the efforts of other partners, for both system-wide and domain-specific collection of policy data. SABER typically goes into more detail on policy and institutional choices than do other data-gathering initiatives, so it is able to get further inside the black box. But there is some overlap, and the teams are drawing on those related efforts and communicating with partners to make sure that the initiatives complement each other. Box 4 provides some examples.

How is the Bank working with partners to finance and implement SABER?

The World Bank has relied on the funds of its Human Development Network and on generous support from partner institutions to fund the start-up costs of the domains developed to date. These up-front costs are significant: each domain requires the research and writing of a comprehensive "What Matters" paper, consultation with international experts in the field, development and piloting of a data-collection instrument, design and testing of data-entry templates, development of a valid rubric for analyzing and rating policy frameworks, and preparation of sample reports. Several partners have provided generous support for SABER's start-up and for country-level application of the framework and tools:

- AusAID
- Children's Investment Fund Foundation (CIFF)
- Japan International Cooperation Agency (JICA)
- Korea Ministry of Education, Science, and Technology
- The Kingdom of the Netherlands (through the BNPP program)
- The government of Russia (through the READ program)
- UK Department for International Development
- UNESCO East Asia (Bangkok)
- UNICEF

Box 4: SABER and other policy data initiatives

SABER complements numerous other partner initiatives that also have a policy-data component. Here is a non-exhaustive list of examples:

- **UNESCO Institute of Statistics (UIS) indicators:** Like SABER, UIS's data-gathering efforts aim to encompass all developing countries. UIS relies primarily on administrative systems as the source for its data, so most data are oriented toward the traditional concerns of those systems—measuring quantitatively the inputs and outputs of the system. Since SABER's focus is policies and institutions, these efforts largely complement each other. In areas where there is overlap, the SABER and UIS teams are working to take advantage of these overlaps (e.g., by peer-reviewing each others' survey instruments and exploring opportunities for collaboration in data collection), and SABER teams use UIS's existing data wherever possible.

- **OECD Education at a Glance[1]:** Education at a Glance offers comparable indicators on education systems that document, in the OECD's words, "who participates in education, how much is spent on it, and how education systems operate. They also illustrate a wide range of educational outcomes . . ." Compared with SABER, the OECD effort is focused more on measuring inputs and outcomes, although in a number of areas—such as assessment systems, school accountability systems, and tuition-fee reforms—the two overlap. Overall, SABER is centered around a different group of countries, because its aim is to inform developing countries' policy choices. It also goes into greater depth in most domain areas, so that it can provide more detailed guidance in particular domains. By contrast, the OECD effort is focused on providing an overall snapshot of the system, using a smaller number of indicators to generate a very useful bird's-eye view of how well a country is doing at devoting the necessary resources to education and achieving good outcomes.

- **Teaching and Learning International Survey (TALIS)[1]:** TALIS, managed by the OECD, collects information on teacher policies and practices in about 30 OECD member and non-member countries, as well as teachers' own views of how these policies and practices are working. SABER-Teachers is aimed at collecting data from developing and developed countries alike, with a concentration on meeting the needs of developing-country stakeholders, and so has a different focus. Where possible, it draws on TALIS data, but many of the indicators that SABER-Teachers is collecting are not available from the OECD. As TALIS is expanded to more non-OECD member countries, the two efforts may be able to complement each other well: SABER's detailed system-level policy variables can be matched with TALIS' rich survey data on the teachers' experience.

Once each domain is ready for rollout, SABER generally relies on partnerships with World Bank regional and country programs, as well as donor partners, to fund data-gathering and analysis in new countries. The financing for early rollouts has come from World Bank regional budgets for analytical (sector) work and projects. The "public good" provided centrally by the core SABER team—the framework, data collection instrument, and comparator data in each domain—is aimed at saving operational units (whether in the World Bank or other partners) the time and effort they would normally incur in trying to

devise an analytical framework and to assess international experience in a particular education domain. Since the fixed costs have already been paid, operational teams need only pay the marginal costs of applying SABER in a new country. Even these costs can be a challenge under current budget constraints, of course, and SABER will also depend on support from other development partners, including other donors operating in a country, that are willing to fund application of a domain tool as part of the collective effort to build an evidence base for country support.

Present status and planned evolution of SABER

What is the status of the SABER rollout?

Since SABER was formally launched in 2011, contacts with governments and World Bank regional colleagues have led to substantial take-up of its frameworks and tools. As of April 2013, SABER had collected data on policies in one or more domains for country reports in over 100 countries, covering all World Bank regions and including also a handful of systems in developed economies for benchmarking purposes. These data have been used to carry out country-level analyses in a majority of the SABER domains, and by mid-2013, SABER will have produced over 100 country reports.

How will SABER evolve over the medium term?

To ensure that it is as effective as possible in meeting the needs of policymakers and other stakeholders, the SABER program will continue to evolve along various dimensions:

- **Expanded country coverage:** SABER is adding new countries in most of the domains at a rapid clip. Each additional country analysis increases the value of SABER to other countries, because of the greater number and variety of comparator countries available in the database.
- **Analysis:** As data become available on enough countries and domains, the SABER team will carry out a variety of empirical analyses that will complement the country-specific reports that each domain will be generating. These new analyses will include, among others: (1) analyses aimed at refining the selection of indicators in each domain, based on empirical studies of how these indicators vary with student learning and other outcomes; and (2) multicountry, multidomain analyses aimed at analyzing how policy progress in different domains predicts progress in educational outcomes.
- **Development of policy-implementation indicators:** As described above, some domains will design and pilot new indicators of the quality of policy implementation, to complement the standard SABER policy-framework indicators.
- **Cross-domain alignment and indicator linkages:** Another methodological step that is needed is to carry out deeper analysis of how the domains interact with each other, including by identifying cases in which ratings of policies may conflict with each other because of differences in domain goals. The goal will be to improve alignment of domains that have developed independently, although in communication with each other.
- **System-wide overviews of relative strength in different policy areas:** Over the medium term, the SABER team hopes to develop an overview policy-analysis tool that can provide a system-wide snapshot of areas of policy strengths and weaknesses. The goal would be to help stakeholders identify parts of the education system that require the most in-depth SABER analysis, especially in cases where budget constraints don't allow application of all the SABER domains over the short term. As the domain databases become more populated and as data from complementary sources (such as service delivery surveys) become more available, it will become possible to do analysis necessary to identify core variables from each domain.

- **Expanded partnerships:** In addition to the partnerships that have supported the start-up of SABER, as the work expands across countries and domains and goes deeper into policy implementation, it will be important to bring new partners on board and to draw even more on the strengths and capacities of current partners. These partnerships can have a multiplier effect on SABER's impact, both in allowing more data collection on the SABER framework and in strengthening education systems using the resulting knowledge. The new SABER Umbrella Facility (SABER-UF) will serve as a vehicle for building these new and expanded partnerships.

References

Amin, Samia, Jishnu Das, and Markus Goldstein, Eds. 2008. *Are You Being Served? New Tools for Measuring Service Delivery*. Washington, D.C., World Bank.

Bold, Tessa, Bernard Gauthier, Jakob Svensson, and Waly Wane. 2010. Delivering service indicators in education and health in Africa : a proposal. The World Bank, Policy Research Working Paper Series, 5327.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* **91 4**: 795-813.

Fiszbein, Ariel, Dena Ringold, and F. Halsey Rogers. 2011. "Making services work: indicators, assessments, and benchmarking of the quality and governance of public service delivery in the human development sectors." The World Bank, Policy Research Working Paper Series, 5690.

Fullan, M. 2005. *Leadership & Sustainability: System Thinkers in Action*. Thousand Oaks, CA, Corwin Press.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* **24 3**: 1141-77.

Hanushek, Eric A. 2003. "The Failure of Input-Based Schooling Policies." *Economic Journal* **113**(485): F64-F98.

Hanushek, Eric A., and Ludger Woessmann. 2009. Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. National Bureau of Economic Research, Inc, NBER Working Papers, 14633.

Hanushek, Eric, and Ludger Woessmann. 2007. *Education Quality and Economic Growth*. Washington, DC: World Bank.

Kremer, Michael, and Alaka Holla. 2009. "Improving Education in the Developing World: What Have We Learned from Randomized Evaluations?" *Annual Review of Economics* **11**: 513-42.

Mourshed, M., C. Chijioke, and M. Barber. 2010. *How the World's Most Improved Systems Keep Getting Better*, McKinsey and Company.

UNESCO. 2011. *EFA Global Monitoring Report 2011*. Paris: UNESCO.

World Bank. 2003. *World Development Report 2004: Making services work for poor people*. Washington, DC, Oxford University Press for the World Bank.

World Bank. 2012. *World Governance Indicators Database, 1996-2011*. http://info.worldbank.org/governance/wgi/index.asp.