



## RESULTS BASED FINANCING (RBF) IN EDUCATION

### QUICK-AND-DIRTY OPERATIONAL NOTES

## HOW CAN DATA SUPPORT RESULTS-BASED FINANCING?

**Small group discussion on the identification, collection, and use of data to support results-based financing**

## KEY CONCLUSIONS

This series consists of summarized conclusions and key discussion points from operational clinics and other events organized by the REACH Trust Fund.

These events bring together experts working on RBF to share their experiences and operational “trade secrets.”

Note 6, October 2017

The Results in Education for All Children (REACH) program in the World Bank’s Education Global Practice supports efforts in RBF. It is currently funded by the Governments of Norway, the United States of America, and Germany. See more at [www.worldbank.org/en/programs/reach](http://www.worldbank.org/en/programs/reach)

## Introduction

The REACH Trust Fund’s new learning series *RBF for Breakfast* brings together relevant World Bank Group (WBG) staff and external specialists working on selected topics related to results-based financing (RBF) to identify the latest operational lessons and existing gaps in knowledge. These events are also a way to draw out tacit knowledge from REACH grantees. The main conclusions of these events are captured in these Quick and Dirty Operational Notes.

The second *RBF for Breakfast* event focused on how to use data to advance RBF. The event aimed to identify the indicators and information systems required to support RBF within national education systems.

Project teams working on education in different countries have identified limited data as a key bottleneck to linking financing to results. In many instances, countries fail to measure the results that matter most and/or do not have the capacity to capture data that is relevant and of sufficient quality. RBF can only work if systems are in place to provide robust and reliable information about results. If it is not possible to accurately measure the agreed-upon results, then it is not possible to provide countries with the financing that is conditional on those results. Unless this problem is solved, the Bank’s conversation with governments on RBF in country dialogues cannot continue.

In recognition of this issue, REACH has funded grants to encourage country teams to explore how to identify the most appropriate indicators and to create data monitoring systems to enable the successful operation of RBF. These grants are also designed to yield lessons about the type of indicators that reflect the most important results (such as learning) and to avoid some common pitfalls in designing or adapting information systems to accommodate RBF.

The task team leaders (TTLs) of three REACH-funded RBF projects shared their experiences at this small group discussion, and participants were invited to provide additional comments, feedback, and advice. The projects are briefly introduced in Box 1. The TTLs are also listed as resource people to contact for additional information.

This note lays out first how to get started in terms of understanding the existing data, what additional data are required, and what data literacy skills are lacking, and then goes on to discuss how to implement the collection, verification, and use of new data.

The following main points are elaborated below (Ctrl + Click to follow a specific link):

<b>How to Get Started</b> .....	<b>5</b>
Understand the Current Data Supply .....	5
Make Decisions According to the Intended Use of the Data and Systems .....	6
Identify the Right Indicators to Measure Results .....	6
Choose More Immediate Indicators As Well As Learning Outcomes .....	7
<b>How to Implement</b> .....	<b>8</b>

Use and Triangulate Self-reported Data.....	8
Understand Respondents' Incentives to Lie and Game the System .....	9
Ensure the Quality of the Collected Data and Encourage its Use .....	9
Ensure that the Collected Data Meet Needs at the System and Local Levels .....	10
Allow for Course Corrections While Recognizing Any Political and Technical constraints .....	11

## **Box 1: Country Examples**

### **Haiti: From Financing Access Results to Learning Results** (TTLs: Juan Baron, Melissa Adelman)

The Government of Haiti has successfully increased primary school enrollment through a results-based mechanism called the Tuition Waiver Program, which pays schools for enrolling poor children in non-public schools. It intends to build on this achievement by providing financial incentives for schools to improve conditions, instruction, and learning outcomes for poor children, while reducing grade repetition and dropouts. The REACH Knowledge, Learning and Innovation (KLI) **grant** is helping Haiti to develop the capacity and the information systems required to enable the effective operation of an RBF mechanism that can inform future government funding decisions and donor support. Haiti's experience in developing these systems will yield lessons about how to implement RBF in other low-income, fragile situations.

### **Vietnam: Are School Characteristics and Teaching Practices Reliable Proxies for Learning Gains?** (TTL: Michael Crawford)

Vietnam's Ministry of Education and Training is in the process of improving the teacher performance evaluation system for general education. The KLI grant is funding research using existing data to establish the underlying factors that affect school quality. The results will be used to redesign the teacher performance evaluation system. It will also yield information about Vietnam, and will establish a model with the potential to be adapted by other countries to conduct their own evaluations of the factors that influence learning.

### **Colombia: Results-Based Monitoring System** (TTL: Pedro Cerdan-Infantes)

The Colombian government is making efforts to manage its education system based on results. Since 2014, Colombia has had a "synthetic index of education quality" for all basic education schools, with yearly targets for every school. However, this index is limited, and policymakers need more information to be able to make well-informed decisions. The KLI grant will support the development of a monitoring system that will make it possible to track multiple dimensions of education quality with a view to introducing results-based fiscal transfers to encourage schools to improve any areas where they are lagging.

Other countries with experience with such information systems that were discussed were:

**Peru** (TTL: Renata Lemos)

**Pakistan** (TTL: Dhushyanth Raju)

**Dominican Republic** (TTL: Juan Baron)

**Nigeria** (TTL: Kirsten Majgaard)

## How to Get Started

The discussion first focused on the steps to be taken when aiming to establish a new information system. The TTLs of the REACH grantee projects described the data situation and priorities in each of the grant countries. The participants then discussed how to decide which indicators are best at capturing results and the need to continually review and revise them.

### Understand the Current Data Supply

The first step in developing systems to support RBF is to assess what data are currently available, how they were collected, and what mechanisms exist to verify their accuracy. All of the countries with REACH grants started out by doing an informal diagnostic of what data were already available as follows:

- **Haiti** has a very weak learning environment, and there are substantial challenges involved in collecting and analyzing data. However, one of the main preconditions for an information system is a government that is ready and willing to make a change, as well as an existing program to build on. These preconditions are both met in Haiti.
- **The Dominican Republic** collects data from schools but they are fragmented and are, therefore, not analyzed or used to inform policy. However, the government is currently strengthening the information system to enable the data to be used more efficiently.
- **Colombia** has a large but non-functional information system. It has a synthetic index on school quality, which charts the growth of each school and provides one number to capture it every year. The schools are given goals for learning outcomes and PISA scores every year. However, the index is limited and largely inactionable. In addition, Colombia currently has an excessive number of complementary information systems with different actors using different data to try to capture information beyond the synthetic index. The schools are at very different levels in terms of their ability to use the data and their interest in doing so. Many private schools hire companies to provide detailed diagnostics to help them to reach a higher number on the synthetic index, since this is used to determine the amount by which they may increase their yearly tuition rate. However, many schools use no data at all. Moreover, while information on schools is being reported by them to the central level, it is not provided to local ministries where many decisions regarding schools are made.
- **Vietnam** is in the unique situation of having a significant amount of micro level education data available as a result of its participation in a multi-country, multi-level, multiyear, externally funded and executed “Young Lives” survey. This extensive dataset covers children from right after birth throughout their school years. However, the Young Lives data are not linked to the national information system, and the government has little appetite for embracing a new approach.

## Make Decisions According to the Intended Use of the Data and Systems

The choices about what additional data to collect and what information systems to develop should be driven by the purpose that they are meant to serve. In **Colombia** and **Haiti**, the goal is for the information system to serve as a management tool to improve decision-making at different levels. In the other countries covered, the more general focus has been on coming up with indicators that will demonstrate results and data monitoring systems that will make that possible.

- In **Colombia**, the World Bank has been working with the government to develop education quality metrics that would complement the synthetic index and would work as a management tool. The team has developed a comprehensive system and a matrix with a large number of indicators. The intention was to base the chosen indicators only on proven measures. However, the evidence is very limited, especially on what works in education in Colombia or in comparable countries. The goal is to collect information from every school and systematize the information so that actors at all levels can use it for decision making. The team's challenge currently is to cut down the number of indicators to four or five for each of the six quality dimensions and to identify which indicator is relevant for each actor, which is proving difficult since all of the actors have different information needs.
- In **Haiti**, similar to Colombia, the team supported the government in building a “quality assurance system” (a set of dimensions with clear standards of achievements and indicators). The data in the system will help policymakers to identify the strengths and weaknesses of all schools and guide them in targeting interventions and allocating resources.
- In **Vietnam**, the team is using the extensive “Young Lives” dataset to identify classroom interventions that reliably improve learning outcomes. They started by narrowing over 2,000 variables down to 200 and then into 15 clusters, focusing on those that affect learning outcomes. Rather than rewarding schools for improved test scores, which can lead to distorted effects, the goal is to encourage the government to finance the inputs and outputs that lead to positive learning gains, thus eliminating the idea of requiring students with different academic abilities to all reach a single target.

## Identify the Right Indicators to Measure Results

In the absence of a proven set of indicators to measure results in the education sector, various countries are using different approaches to measuring progress.

In many cases, student test scores are used to measure learning outcomes, but evidence has shown that they are a particularly “noisy” target. According to [Goodhart's law](#): “When a measure becomes a target, it ceases to be a good measure.” This relates to the situation when individuals try to anticipate the effect of a policy and then take actions that alter its outcome.

The goal of the current REACH grants has therefore been to come up with indicators that correlate to improved learning but are not test scores and are relatively easy to measure. For example, evidence shows that teacher effort is important for learning. If teachers are mostly absent, one can conclude

that students are probably not learning. Therefore, policymakers may choose to give teachers an incentive to show up in the classroom as a first step towards achieving better outcomes. However, it is not easy to measure a factor like consistent teacher attendance in the classroom especially if rewards are introduced as this requires impartial monitoring of the attendance more than once a day. Identifying indicators other than test scores that can be measured cost effectively is what the teams are attempting to do.

## Choose More Immediate Indicators as well as Learning Outcomes

Although the ultimate goal of an education program is to improve final outcomes such as learning and attainment, linking financing to outcomes that are evident earlier in the results chain is often more productive. The following points were made during the small group discussion:

- The participating project teams indicated that it is important to choose objective and realistic indicators that are faster and easier to verify than learning outcomes to complement the indicators that measure learning outcomes.
- Deciding on what other indicators to use is a process, and several different combinations may need to be tried before the final list is determined.
- Focusing on intermediate results that will ultimately contribute to better learning is especially useful in countries in fragile circumstances where a top priority for education systems is to increase access to school and improve conditions. Often these education systems have little accurate data on other results indicators such as progress in learning. Helping country clients to gather data on the immediate needs of the schools (such as infrastructure) can motivate them to aspire to reach longer-term learning goals.
- In **Haiti**, the approach has been to focus on gathering data on simple indicators, one at a time, with the goal of expanding later. This has been strategic and necessary because most of the actors in the education sector in Haiti consider the situation hopeless and have no expectations that it will improve. Therefore, even small successes are welcomed with enthusiasm, which can be used to motivate further engagement. Building on these smaller successes, the team is working to develop 3<sup>rd</sup> grade learning standards and assessment methods so that information on learning outcomes can begin to be collected more systematically.
- It is not clear what the balance should be between focusing on learning outcomes and focusing on other indicators. While World Bank teams need to stress the importance of gathering data on learning outcomes, it is equally important to encourage countries to choose other indicators on which it is easier and faster to collect data.
- Local solutions can be a good starting point in situations where it's not clear what the key indicators of success should be. RBF is a particularly appropriate tool in such cases. It gives an opportunity for the homegrown solutions to be tested, adjusting them as necessary, and potentially implementing them on a larger scale if the indicators gathered as part of the pilot test help to achieve the specified outcomes.

- The interventions can start small and with low stakes but with enough incentives for performance. This solicits innovation, flexibility, and autonomy.

## Key Messages on How to Get Started

- Understand the supply, quality, and future uses of data
- Understand the supply of resources and incentives related to the current data situation
- Identify how new data could be used and make allies in decision-making bodies to encourage the data to be used
- Create information management systems and think about how all of the data connects to the big picture
- Continuously evaluate, refine, and, where necessary, replace the chosen indicators
- Focus not only on final learning outcomes but also on outcomes earlier in the results chain to keep the data collection process simple.

## How to Implement

The discussion about how to collect, use, and verify data and build information systems focused on the following topics:

- Using and triangulating self-reported data
- Understanding respondents' incentives to lie
- Ensuring the quality of the data collected by building countries' capacity and encouraging the proper use of data
- Ensuring that the data that are collected meet needs at both the system and local levels
- Using current education management information systems (EMIS) to verify the data while taking into consideration any system constraints
- Learning from gaming rather than over monitoring for it
- Allowing for course corrections, while recognizing political and technical constraints.

## Use and Triangulate Self-reported Data

In order to help schools to improve their performance, data need to be collected annually from each school. However, it is often unrealistic and expensive to conduct surveys or externally run assessments to gather these data. Therefore, country education systems are increasingly relying on self-reported data, especially to collect information on the perceptions and attitudes of students, teachers, or parents. This is the case in **Colombia** and **Haiti**.

Some guidelines were mentioned in the small group discussion as to how to collect this type of data:

- Collect the same data from all of the relevant actors (students, parents, and teachers) to maximize the reliability of the data.
- Avoid self-assessment questions but instead ask respondents about the performance of others.
- Do not ask straightforward good/bad questions, but instead ask whether there has been a change in certain conditions or behavior.



- Ask respondents initial questions to understand their possible biases. In **Colombia**, every indicator used was mapped according to potential biases.
- Consider having respondents answer some of the questions in groups to generate discussion and consensus.
- Make it clear to respondents that the instrument is not an evaluation and that there are no negative consequences.

However, self-reporting does not work well for hard-to-measure indicators such as socio-emotional skills and classroom environment. Surveys or third-party assessments are necessary to gather such data, of adequate quality.

Moreover, in **Vietnam**, where the goal of the project's data collection effort was to identify a set of classroom interventions that can reliably improve learning outcomes rather than to provide management data at the school level, the benefits of third-party data were considered by the project team to be significant. The "Young Lives" data that were used in **Vietnam** were gathered by a third party as part of a multi-country survey of household data. While allowing a third-party group to collect the data had its challenges, the extensive time period and number of variables collected outweighed them. The "Young Lives" survey produced data on 2,000 variables that were identified as being influential to students' academic performance, which were then carefully disaggregated to define different levels of academic achievement.

Wherever possible, it is advisable to use data that have already been collected and reorganize the information rather than collect it again. It is also preferable to build on existing information systems rather than creating new ones.

### **Understand Respondents' Incentives to Lie and Game the System**

As soon as monetary or other significant incentives are linked to the achievement of the indicators, self-reporting becomes problematic. Trying to triangulate by collecting information from a variety of different respondents does not always yield more accurate information or lead to less gaming. When schools are rewarded for good performance and the incentives are high, collusion between different respondents can occur. More efforts should be made to identify respondents' incentives to lie. When the risk of lying is particularly high, it is usually not advisable to link monetary rewards to the achievement of outcomes.

### **Ensure the Quality of the Collected Data and Encourage its Use**

For any information system to work, the data need to be reliable. However, many countries lack the capacity to verify the data, so it has to be built. In addition to training workshops, on the job training for the people meant to do the verifications can be useful by engaging them in the data collection activities.

It is also important to build the data literacy skills of school and education administrators at all levels and create a local, friendly environment for them to use the data to inform changes in policy and

procedures. **Argentina** is an example of a country where work is being done to build capacity to use data on learning results. In addition to otherwise encouraging it, it may be useful to publicly reward those who use data actively and accurately. This helps to emphasize its importance.

### **Ensure that the Collected Data Meet Needs at the System and Local Levels**

Especially in countries where the information system is meant to work as a management tool, it needs to respond to the needs of all stakeholders. It is therefore crucial for policymakers to understand all of the different information needs at the central government, local government, and school levels. Ensuring that the data in the information system is relevant to the actors involved is a strong incentive for them to provide the information in the first place. For example, there is an incentive for schools to provide information when the data generated by the information system helps them to make decisions and allocate resources within the school. The meso level, such as local ministry and district or province-level officials, is often the hardest to incentivize because their incentives are not as clear as those of the schools or central ministries who need the information for their decision-making.

### **Use Current Information systems to Verify Data Taking into Consideration Any Constraints**

It is important to signal that verification and data accuracy matter. In parallel with capacity building, data verification mechanisms should be integrated into the information systems such as EMIS and/or other quality assurance mechanisms should be put in place. **The Dominican Republic** has an EMIS, but the data are not verified and there are no signals to indicate the importance of verifying their accuracy. A system is now being developed to regularly check the quality of the data with a goal of embedding it into the EMIS system. In **Pakistan**, limited spot checks were built into the information system. These checks, conducted by third parties, focus only on certain key indicators that should remain the same regardless of how they are measured (such as enrollment rates or student-teacher ratios). System-level indicators are prone to interpretation and thus are harder to verify.

When using different systems to cross-check the information, it is important to ensure that the data are comparable. For example, household data and EMIS data are very different and should not be used to validate each other. Suggested interventions and systems should be introduced with the reality of their potential rigidity in mind.

### **Learn from Efforts to Game the System**

There is a risk that extensive efforts to eliminate cheating lead to inefficient use of resources and energy. It was suggested that some gaming is unavoidable, and instead of putting a lot of effort into trying to eliminate it entirely, teams should be prepared to go through two or three different reorganizations of the system after observing where the gaming happens. As any gaming of the original technical solutions and/or the creation of perverse incentives are detected, these can then be fixed in the next version of the system.

## Allow for Course Corrections While Recognizing Any Political and Technical constraints

Even though ideally, the identification of the indicators, the collection of data, and the building of the system are all parts of fluid process that make it possible to make adjustments when necessary, political and technical costs may make this impossible. The government may not be willing or able to make changes to the system after it is implemented, especially if it is in the public sector. In order for a results-based information system to be implemented in any given country, the government needs to be willing to make any necessary changes and have the technical capacity to do so. The government also needs to have a genuine interest in taking action in response to the results to improve the quality of education.

## Key Messages on How to implement the Collection, Verification, and Use of Data

- Use and triangulate self-reported data to ensure their quality
- Ensure that the collected data meet needs at both the system and local levels
- Use existing EMIS to verify the data while taking into consideration any system constraints
- Do not waste resources in trying to eliminate cheating but learn lessons from it to adapt the system
- Attempt to modify the system wherever possible in response to information about what is and is not working.

## List of Experts

In attendance:

1. Samer Al-Samarrai (chair)
2. Michael Crawford (Vietnam)
3. Pedro Cerdan-Infantes (Colombia)
4. Juan Baron (Haiti)
5. Lars M. Sondergaard (EAP)
6. Reema Nayar (Practice Manager, LCR Education)
7. David Evans (AFR)
8. Angela Demas (SABER)
9. Emily Gustafsson-Wright (Brookings Institution)
10. Francisco Haimovich Paz (Colombia)
11. Francisco Zavala (Colombia)
12. Luis Felipe Martinez (Colombia)
13. Alonso Sanchez (Colombia)
14. Kesha Lee (REACH team)

15. Minna S. Mattero (REACH team)
16. Koji Miyamoto (SABER)
17. Shwetlena Sabarwal (Nepal)
18. Renata Freitas Lemos (Peru)

Others:

19. Juan Manuel Moreno
20. Neda Bostani
21. Xiaoyan Liang
22. Tamar Atinc
23. Eduardo Velez Bustillo
24. Peter Holland
25. Rafael De Hoyos Navarro
26. Manal Bakur N Quota
27. Mohamed Yassine
28. Diego Angel-Urdinola
29. Samira Halabi
30. Elena Maria Roseo
31. Marie Evane Tamagnan