



Summative Report

Advisory Assistance to Apply Policymakers' Workbench in Dialogue with Policymakers in Peru and Jamaica

Noel McGinn¹ and Massoud Moussavi²

Causal Links, LLC
www.causallinks.com

October 2008

¹ mcginn@causallinks.com

² moussavi@causallinks.com

Table of Contents:	Page
I. Introduction	3
II. The conceptual framework of our approach.....	3
III. A description of the Policymakers' Workbench model.....	5
IV. The utility of the Snapshot data in its operation.....	10
V. The results of discussions of that operation with policy makers and researchers in Peru and Jamaica.	12
VI. Use of Published research literature in construction of the model.....	15
VII. Sensitivity analysis of the model.....	16
VIII. Conventional policy alternatives and their likely effect..... on learning outcomes	17
IX. Likely Impact of Interventions that Change Intermediate Variables.....	21
X. Conclusion.....	22
XI. Contribution to Human and Institutional Capacity Building.....	23
XII. Next Steps.....	23
Acknowledgement.....	24
Appendix 1: List of Participants in Peru.....	25
Appendix 2: List of Participants in Jamaica.....	26
Appendix 3: Technical and Advisory assistance Program.....	27
Qualitative Participant information	
Appendix 4: Notes for Advisory Assistance.....	29
References.....	30

I. Introduction

This project has been funded by Higher Education for Development (HED). It builds upon the initiative of USAID to provide education policy makers with tools for the rapid and inexpensive collection of reliable and valid data about (factors affecting) the performance of schools. As a first step, the EdDataII initiative sponsored the development by the Research Triangle Institute of a data assessment tool called Snapshot for School Management Effectiveness (SSME). The set of data collection instruments was tested in a pilot application in primary schools in Peru and Jamaica.

The SSME generates data on a large number of variables identified by experts in education policy as the most important questions to be asked about school performance. In addition, RTI applied a brief test of reading ability and comprehension (the Early Grade Rapid Assessment device or EGRA). Statistical analysis can establish the association between answers to the SSME questions and scores on the EGRA, as a means to demonstrate the validity of the SSME instruments.

In addition, however, USAID seeks to demonstrate the utility of the Snapshot data in the policy formulation process. To that end USAID supported and facilitated HED's contract with Noel McGinn and Massoud Moussavi (representing the firm Causal Links) to apply their newly developed engine, *Policymakers' Workbench*, to the Snapshot data collected in Peru and Jamaica.

II. The conceptual framework – models in education

Education is a complex process in which relatively isolated teachers seek to interest and instruct large numbers of students who vary in previous knowledge, ability to learn and motivation. The typical effort to improve quality, on the other hand, affects only a few of the many variables in the education process.

For example, the ultimate objective might be to improve students' learning. Some policy proposals offer incentives intended to increase teachers' level of effort, but take no action to increase teachers' knowledge or skill level, may ignore organizational constraints to changes in teaching practices, and may overlook factors that limit students' ability to respond to new methods of teaching. The history of education reforms is filled with examples of failures to recognize that outcomes are influenced by a number of different factors that can frustrate the intentions of a single intervention.

The typical policy intervention affects inputs to the education system. For example policies may be intended to affect teachers, students, school management, physical facilities, and materials. Most policies impact one or another but not all inputs at once. Missing are policies intended to directly affect the process of instruction.

Most policies have been aimed at inputs rather than instructional process, we believe, because policy analysis has been relatively successful in describing the association between educational inputs and learning outcomes. Inputs and outputs are easy to

observe and can be measured reliably. When countries were expanding access to primary education, research was critical to learn how to get and keep students in school. Not much attention was paid to how much students were learning. In addition research on teaching yielded few insights. The teaching process is complex and changes in the process of learning. The effects of specific practices are difficult to observe and to measure reliably. While policy analysis has made great advances in statistical methods to estimate relationships between input and outcome variables, its methods have been less fruitful for explaining the teaching and learning process. Instead, understanding of how students learn and what kinds of teaching are most effective has been gained using anthropological and experimental methods.

The focus of our work has been to model the teaching and learning process as it takes place in schools. The work requires synthesizing results of empirical research and field experience, consistent with advances in learning and organizational theory. We seek to present in a coherent fashion the accumulated knowledge and understanding of many educators and researchers. We have attempted to represent their beliefs, and our own, in a systematic way that describes how inputs into schools are transformed into learning outputs.

The educational content of our model begins with the insights of John Carroll (Carroll, 1963) who suggested that amount of learning is a function of time spent on the learning task and the rate at which learning takes place. Various conditions must be met for students to learn the curriculum, including some form of instruction or instructional materials that present the curriculum and the presence and the attention of the students. A number of factors contribute to each of these conditions. Other conditions affect the rate at which students learn the curriculum. The conditions are in turn influenced by external factors such as physical conditions of the school, level and quality of training of teachers, health of students, management practices and others. In addition, learning can take place outside of schools, and therefore is affected by family and community factors.

Learning outcomes are the result of a number of connected factors. For example, some of the external factors can be affected directly by government policy; these typically are called inputs. The inputs affect the organization and operation of schools (for example, teaching practices, class size). Changes in organization and operation in turn affect learning outcomes. While inputs can be changed independently of each other, the intermediate factors of organization and operation are linked; changes in one generate changes in others. For example, changing textbooks can change teaching practices; changes in teaching practices affect student engagement in the learning task; but engagement is also influenced by instructional materials. The maximization of learning outcomes requires setting all factors at the level where they work together to generate learning.

In addition to the large volume of research stimulated by Carroll's insights (and summarized in part by Squires, Huit and Segars, 1983), there is a large volume of research that focuses on teaching practices [American Educational Research Association, 2001; Borich, 2006; Gagne, 1979; Joyce, 2003]. Missing, however, is a synthesis of

research across disciplines and linking classroom practices with characteristics of school organization and external influences including system-wide policy. The model we have developed links empirical research and field experience, but also advances in learning and organizational theory. As far as we know, however, our model is the first operative model that combines input and process variables that cause learning.

III. Description of Policymakers Workbench

Building upon previous work, we have developed the *Policymakers' Workbench*, a knowledge-based software engine that can assist policymakers in the process of thinking through a policy. This tool presents users with different scenarios and their consequences based on *heuristics*, *empirical cases*, and *information from stakeholders*. The tool enables decision makers to iteratively observe and analyze the effects of various actions and scenarios or consider different permutations of options. For example, users might want to analyze the impact of additional teacher training or private tutoring on the test scores of students. Alternatively, they might want to determine the major reasons for low English and Math scores for a given school given certain observations.

As described in the following sub-sections, there are two components of the Policymakers' Workbench: (i) an "inference engine" which enables inferences based on a Bayesian model representing the underlying relations or functions, such as "teacher motivation and discipline both affect time spent on instruction"; and (ii) a computer user interface.

(i) The Bayesian Model

The underlying model for this engine is a belief (or Bayesian) network that specifies dependencies among different variables. A Bayesian network is a directed graph that represents dependencies among variables. The full specification of such a network is as follows [Russell, 2003]:

- A set of variables makes up the nodes.
- A set of links or arrows connects pairs of nodes. If there is an arrow from node X to node Y, X is said to be a *parent* of Y. The intuitive meaning of an arrow between nodes X and Y is usually that X has a direct influence on Y.
- Each node has a conditional probability distribution that quantifies the effect of parents on the node.
- The graph has no cycles.

Consider the following simple example: *teacher motivation and classroom discipline influence the amount of classroom time spent on instruction, and small classes have a better level of classroom discipline*. The Bayesian network for this example is shown in Figure 1.

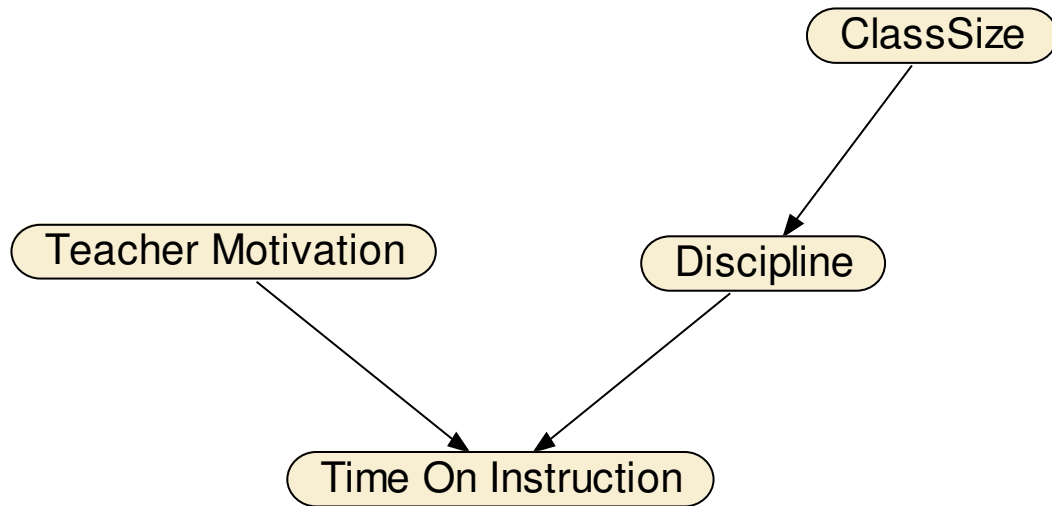


Figure 1. An Example of a Belief Network

A variable can take on a value from a domain of values. For example, consider the variable *ClassSize* in the above example. We can define the domain of values for this variable as {small, moderate, large}. The probability distribution associated with each variable then represents the chance that the variable takes on a respective value. For instance we can specify $P(\text{ClassSize}) = (0.3, 0.5, 0.2)$ which means that we believe the chance of class size being small, moderate or large is 30%, 50%, or 20% respectively.

For the sake of simplicity, all the variables in our model have been defined as binary variables. That is, they take one of two values only. A richer representation should improve the accuracy of the model. However, it also increases the complexity of the model and the effort to construct and maintain it.

Because a belief network represents variables and their relationships of interdependence, it can be used to answer queries about those relationships using probabilistic inference. For example, given the evidence that class sizes are small, we can query the model about the probability of having good discipline in classes or the probability that a high portion of the class time is spent on instruction. Now if we additionally observe that teachers are highly motivated, this latter probability would go up. Alternatively, it would go down if the observation was that the teachers were not highly motivated.

The model we have constructed can predict expected educational system performance and explain learning outcomes in terms of the choice and implementation of practices that affect those outcomes. At the level of individual schools the tool explains performance as a joint function of national policies and practices as well as local contextual conditions and characteristics (including teachers, students and community).

In addition, the tool can be used to diagnose non-observable aspects of school operation that require improvement.

The Policymakers' Workbench model includes 66 variables of three kinds. The first kind is nodes without parents. That is, they are nodes or variables that are not directly influenced by any other variable. In this model, they are mostly the variables that in education research are commonly referred to as input variables. Examples of such variables are textbooks, library, student attendance rate, etc. A second kind of variables is nodes that are not parents to any other nodes. That is, variables that do not directly influence other variables. Examples of such variables are reading test score, math test score, etc. These nodes represent learning outcomes. A third kind of variables, intermediate between the other two sets of variables, represent internal conditions in the teaching and learning process. Examples of such variables are curriculum coverage, time on task, time learning outside school, etc.

(ii) The User Interface

The user interface enables users to analyze the impact of various interventions or policies. It also permits users to examine the trade-offs between different factors, taken individually and in combinations. Figures 2 and 3 shown in the following pages provide sample screen shots of the user interface of the system. Figure 1 shows how a user can select a country and the focus of analysis (school level or national level). Figure 3 displays the situation in the selected country. As can be seen in the top section of Figure 3, we have categorized the input variables in four general categories: **School Organization** (Principal, Textbooks, Library, etc.), **Family** (family SES, Family Involvement, etc.), **Student** (Academic History, Health, Nutrition, etc.), and **Teachers** (Academic Education, Variety in Methods, training, etc.) The bottom section of the screen in Figure 2, displays the computed or inferred variables (e.g., Total Amount of Learning, Expected Reading Score, etc.) Here again we have grouped variables in four categories.

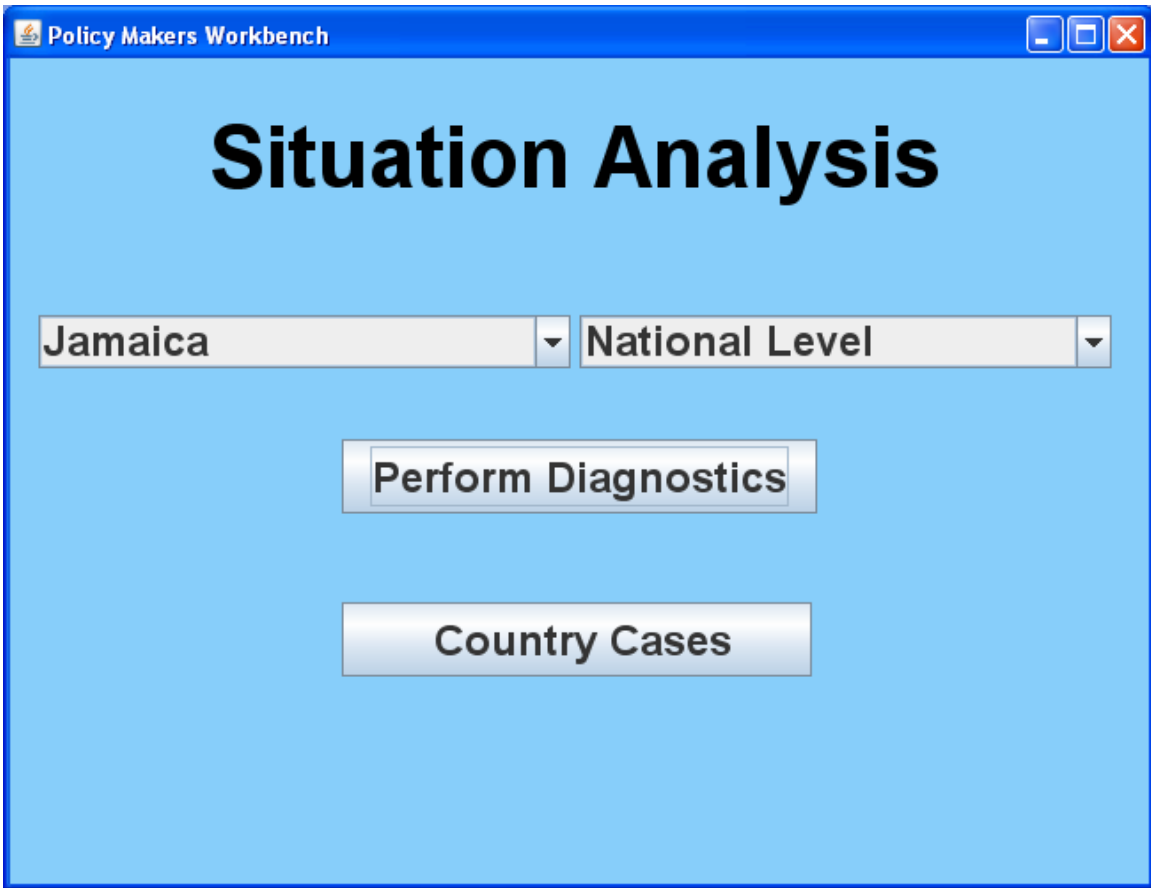


Figure 2: Situation Assessment Screen

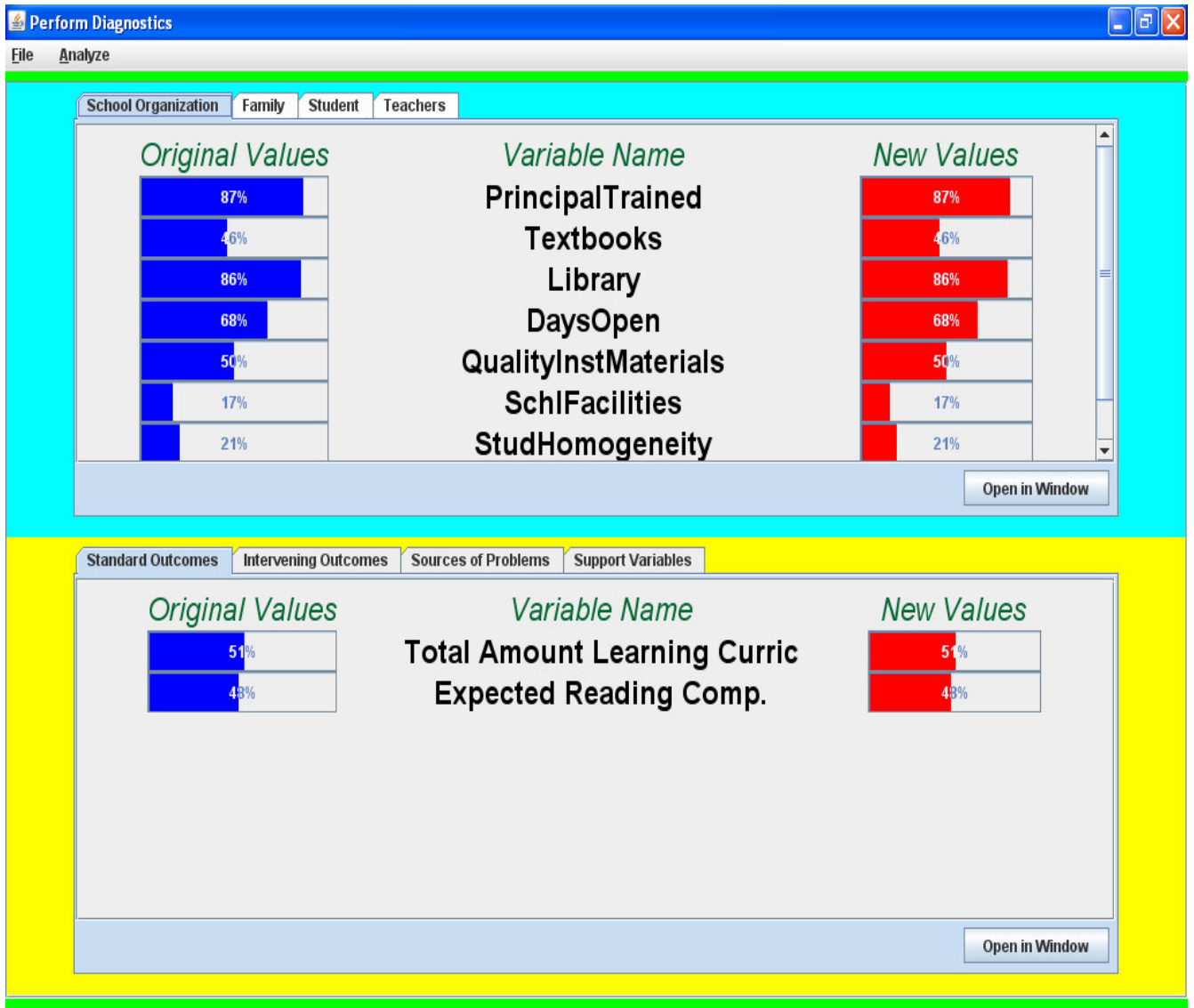


Figure 3: The input and Output variables

The user interface can be used to:

- 1) assess the likely impact of changes in specific observable input variables (such as supply of textbooks, or levels of training of teachers) on learning outcomes;
- 2) identify the likely values of non-observed variables that mediate the effect of specific inputs (for example, how changes in supply of textbooks affects teacher use of time);
- 3) assess the likely impact of changes in mediating variables on other variables including learning outcomes;
- 4) estimate the changes that would have to be made in both mediating and input variables to maximize learning outcomes; and
- 5) describe differences in inputs and mediating variables of schools with different levels of learning outcomes.

IV. Using the SSME Data to Model Schools in Peru and Jamaica

Copies of the SSME data collection instruments and the data sets for Peru and for Jamaica were obtained courtesy of Jenny Spratt and Amy Mulcahy-Dunn of the Research Triangle Institute. The SSME includes 7 instruments: Student, Teacher, Director, School Facilities, Classroom Facilities, Classroom Activities, and Parent. Given the pressure of time, only data from the first 5 instruments was used. The Jamaican data came with a number of pre-defined variables, reducing the time required for cleaning and construction of variables. The Peruvian data required more attention.

The SSME does include questions about some of the internal process variables but, given constraints of time, we used the SSME data to provide initial values for 23 of the input variables.

The criteria for selection of the SSME questions to be used were 1) the conceptual fit between the SSME definition of the variable and that of the model; 2) relatively low skew in the distribution of values of the variable; and 3) responses for at least 90% of the cases. The SSME data did not include measures of Health, Family Size, Student Time on Entertainment, Teaching Experience, or Quality of Instructional Materials. Data on frequency of homework assignments and the activity of other students did not meet our criteria. Most of the SSME data appeared to be of high quality and, as the RTI analyses show, account for variance in reading scores in Jamaica and in Peru almost as well as larger and more expensive data sets.

The SSME data was then converted from the form required for analysis using measures of association, to the binary values required for our Bayesian model used in this project. Conventional analytic tools such as regression analysis use data in the form of nominal, ordinal or interval scales. Each value represents a category or amount of what the variable measures. Bayesian models use probability estimates for a variable taking on a particular value. These can range from 0 to 1 (or, 0 to 100%). Some of the SSME

questions provided dichotomous variables in which values could easily be converted into a Yes/No format. Other variables required a definition of the value level that would be considered as “True” or “Not True”. Each of these variables was structured to meet the requirements of our model, that is, a binary format.

Most of the SSME questions for teachers, schools and directors required Yes/No answers; a few others required us to make subjective judgments about how to split the variables. For example, answers to the question about informal classroom visits by the director were recorded by the interviewers on a 7-point ordinal scale from Never to Daily. For the purposes of our model, Supervision was defined as informal visits every two weeks or more often.

Subjective judgments were required because while prior research has demonstrated which variables (e.g., teacher training) have strong correlations with other variables (e.g., teacher skill), it has not established the level of each variable that makes it sufficient as a cause of or condition for another variable (e.g., what constitutes a skilled teacher).

There are three levels of variation in the variables: across students, across teachers, and across schools. A few of the SSME variables for students (and their families) provide data in ordinal or interval form, more are dichotomous. For the dichotomous variables, we defined the “positive” response as indicating a sufficient value. Ordinal or interval scale variables required a subjective judgment as to the values that define presence of the state or condition defined by the variable. For example, Family SES was estimated by presence in the home of 10 household facilities (running water, washing machine, TV). High SES was defined as having 7 or more of these facilities in the home. These judgments were based on prior experience and qualitative research.

The values of the re-coded variables were used to determine the “prior probabilities” of 23 input variables in the model. Probability estimates associated with the student-level variables encode the probability that a given student in the country meets some conditions (e.g., an acceptable level of nutrition). Similarly, probability estimates associated with Teacher variables encode the probability that a teacher will meet a particular condition (e.g., in Jamaica have had at least 1 day of in-service training, in Peru 60 days). School-level variables represent the probability that any school in the country will meet a condition (e.g., have received textbooks on time).

In addition to the instruments, the SSME applied a test of reading comprehension (the Early Grade Reading Assessment). This test has four parts: recognition of letters, pronunciation of words, pronunciation of nonsense words, and reading a story. The first three sections are coded for speed, the last for speed and comprehension. Reading comprehension is the standard in most schools for promotion from 1st to 2nd grade. We chose this as our measure of learning outcomes because comprehension is regarded as essential for learning through reading. The variable for reading comprehension is based on answers to 5 questions about the selection read. As a consequence, scores are not highly reliable; the scores do not correlate highly with those on the other sections of the EGRA.

Some SSME variables have missing values. In conventional analysis this requires either elimination of the case or estimation of the missing value. This is not necessary in Bayesian analysis.

V. **Presentation of the SSME-Based Results in Peru and in Jamaica**

Peru

The USAID Mission in Peru arranged two formal sessions in which the model was presented. The first, on September 9, attracted about 40 participants from the Mission, and from the Ministry of Education's National Council of Education, the national research and policy analysis organization GRADE, the Foro Educativo (which promotes discussion of educational issues), the Pontifical Catholic University, the University Cayetano Heredia and other groups. After an introduction by mission director Paul Weisenfeld, Noel McGinn gave an introduction to the objectives and methods of the SSME project, and its use in the detection and analysis of problems in the education system. He then introduced a conceptual framework for analysis of the process of teaching and learning in schools. Massoud Moussavi introduced the logic of probability analysis using Bayesian models, and introduced the framework of the model used in analysis of the SSME data.

Using SSME data, Moussavi and McGinn showed how changes in input variables taken individually have only a slight impact on learning outcomes, but when combined add up to more significant changes. The audience noted that, consistent with research using regression analysis, changes in school inputs such as levels of teacher training, presence of libraries, and training of principals contribute to small but significant increases in test scores. This finding is consistent with Scheerens' (2005) review of a large number of regression studies.

Participants also asked if the model could indicate the changes that would have to be made in input variables in order to have all students achieve the highest possible test scores. This is one of the powerful aspects of our tool, the fact that it can be run in a predictive mode as well as a diagnostic mode. When the tool is run in the diagnostic fashion, it calculates the probability values for all linked variables in order to produce the specified outcome value. Unlike changes in specific input variables, which generate effects through the specific branches connected with the output variable, all the links in the model are involved in the recalculation of probabilities.

The discussion then turned to the intermediate variables in the model, such as Curriculum Coverage and Student Engagement. The analysis showed that changes in these variables (more closely linked to learning outcomes) have a larger impact on test scores than do changes in the input variables. This suggests that learning outcomes could be improved more dramatically by acting directly on these variables. For example policies could be designed to increase the proportion of the curriculum teachers cover in a year. Teacher

behavior could be changed by specific training, closer supervision by principals or by incentives.

Participants noted that some of the variables, such as Teacher Training, measure only amount of training and do not indicate the quality of training received or the specific content or emphasis of the training. The SSME could provide this information by addition of questions assessing quality and content. Participants asked about EGRA, the measure of learning outcomes that was used. The model results using this measure were similar in direction (but not in strength of relationship) to those obtained using data from another country (Botswana) that employed a 50-item language test. A longer EGRA test would have greater reliability and more predictive power but would also require more time for administration and scoring.

Some participants pointed out that the Early Grade Reading Assessment does not capture all the dimensions of learning in primary school. No doubt it would be possible to develop similar quick tests for mathematics and perhaps other subjects. The participants asked whether the model would work with other measures of learning. Individual student test scores vary by subject matter, as a function of differences in student ability, teaching practices and instructional materials. The model is sensitive to these differences but most of the variables that explain learning are common across subject matter.

A more critical question is the extent to which varieties of teaching practices are described by the SSME data and represented in the model. Descriptions of teaching practices appear in the Teacher and Student questionnaires, and in the Classroom Observation and Classroom Activities observation instruments. These data were not used in the current version of the model (which included only input variables).

A second meeting was held with staff and officers of the National Council of Education. About 35 people attended. The discussion in this meeting focused more on the logic and limitations of the model, and less on the SSME data. Participants were interested in the conceptual development of the model and its foundation in research and accepted theory. Both meetings were two hours in length.

In both of these meetings we sought participants' suggestions and feedback concerning the improvements to the model. The following is a brief summary of the discussion and suggestions for additional variables:

- Focus on interventions and their cost/benefits.
- Child-labor and employment.
- Use of ICT in schools.
- Family structure/one parent children.
- School culture/environment.
- Principal training, type of training, leadership, commitment, etc.
- Contract teachers/lack of stability.
- Teachers' attitude toward change.
- Class preparation, enrichment classes, teacher coaching/mentoring.

List of participants for the meetings in Peru is provided in Appendix 1.

The model was also informed by interviews with teachers. On September 10, 2008, We visited the Jorge Portocarrero School in Ventanilla, which is one of the sites where teachers have received training through the Centers for Excellence in Teacher Training project (CETT). The visit was arranged by Prof. Zoila Del Valle of the Universidad Cayetano Heredia, who participated in the first meeting to present the model. Like many other schools in marginal areas, the Portocarrero has a limited supply of material resources such as textbooks and library assumed to be important for good instruction. On the other hand, students in classroom with CETT-trained teachers were highly engaged in an active learning process of superior quality.

We interviewed the principal of the school and several teachers (some who were not CETT-trained) in an effort to understand the outstanding teaching and learning process we had observed in some classrooms. The differences across classroom could be attributed to the CETT training, rather than to levels of pre-service training, years of experiences, materials, or other input variables in the model. The picture shown below is representative of a group project in one of the classes in the school.



Jamaica

The meeting in Jamaica was introduced by mission director Karen Hilliard. In addition to Massoud Moussavi, the presenters included Amy Mulcahy-Dunn of RTI and representatives from the Ministry of Education and the University of the West Indies. The latter group was responsible for the administration of the SSME instruments and made extensive comments on that process and the quality of the data generated. The Jamaican education community clearly values the SSME and was eager to be associated with it.

Presentation of the use of the SSME in the Policy Maker's Workbench was necessarily brief, about 30 minutes. Massoud Moussavi described the conceptual framework of the tool, briefly discussed the Bayesian models, and introduced the framework of the model used in analysis of the SSME data. As the time allocated for the presentation was short, he focused mostly on showing to the participant how the tool can be used both as a predictive and a diagnostic tool.

The reaction and engagement level of the participants were similar to the ones in Peru. At the end of the presentation, Moussavi conducted a feedback session in which he asked for participants' suggestions for further improving the model. The following is a brief summary of participants' discussion and suggestions:

- Use of extra-curricular material.
- Gender gap and male students drop out rate.
- Special needs students.
- Child labor/employment and its consequence such as lack of attention, fatigue, etc. How to deal with such concerns?
- Violence in Schools, issues concerning discipline.
- How to take advantage of various studies? How to consolidate and use them?

Appendix 2 provides the list of participants in the meeting.

VI. Use of Published Research

In both Peru and Jamaica the primary concern of national politicians and the public alike is the apparent low quality of the educational system. This judgment is based, in the case of Peru, on very low scores on international tests of knowledge and communication skills. On a test comparing countries in Latin America, Peru ranks near the bottom on average student scores. In a comparison with industrialized countries, Peru ranks last.

Three factors have been identified as contributing to the low quality of educational institutions. Low student performance on international comparisons is seen as a result of poverty, especially among populations that do not speak the language of instruction (Spanish in Peru, English in Jamaica). Second, the quality of teaching is low, and especially in low income populations and those not speaking the language of instruction. Teachers are poorly motivated, are frequently absent and do not cover the curriculum adequately. Poor student performance in the lower grades is attributed to inadequate

instruction in reading. Third, low parental involvement contributes to low student motivation and lack of accountability of teachers. In Jamaica, a high rate of attrition in primary grades especially among boys contributes to a high unemployment rate.

In Jamaica USAID is supporting a program aimed to increase child involvement in the instructional process, and participating in a program to improve teachers' skill in teaching reading. In Peru, in addition to the program to improve teachers' skill in reading instruction, the mission is supporting a new school program, AprenDes, based on the Escuela Nueva model of Colombia that retrains teachers to be facilitators of learning and introduces self-instructional workbooks for students. In both countries these are relatively new initiatives that so far have reached only a small portion of the total student population.

Other policy responses to low education quality have sought to improve teacher performance by changes in school governance, or through incentives to teachers. Some governance policies focus on the role of principals. These policies give principals more responsibility (and authority) for the supervision of teachers, or alternatively train principals to act as instructional leaders who affect teachers' instructional practices. Another approach seeks to change teacher (and principal) behavior by increasing parental involvement and control. Some countries have adopted (in response to encouragement by donor agencies) policies that reward teachers for further training, or evidence of improved teaching in the form of higher student test scores.

The model includes the principal variables that are affected by each of these policy alternatives. For example, the model includes Principal Leadership, Teacher Training, Parent Involvement, Supervision of Teachers, Teacher Motivation, and Teacher Skill. It also includes those factors that contribute to low quality that are not affected directly by the policies, such as Family SES and match between home language and language of instruction. The relationship between each of these variables and learning outcomes is well-researched. Scheerens (2005) for example, has reviewed 1211 studies looking at how various factors affect learning outcomes. He carried out a meta-analysis of the best studies in order to estimate the average effect size (or impact) of these variables on the measure of learning outcomes used. This and other studies were consulted as part of the process of anticipating the likely effect of policies on the variables defined in the model (Cabrales & Eddy, 1992; Coe & Fitz-Gibbon, 1998; Duyar, 2006; Fuller, 1987; Glewwe & Kremer, 2005; Haertel, Walberg, & Weinstein, 1983; Hanushek, 2006; Helmsley-Brown & Sharp, 2003; Leclerq, 2005; Marzano, 2003; Mella & Ortiz, 1999; Nascimento, 2008; Raudenbush & Bryk, 1988-1989; Reynolds & Teddlie, 1999; Sammons & others, 1995; Scheerens, 2000, 2005; Simmons & Alexander, 1978; Subotnik & Walberg, 2006; Unidad de Medición de la Calidad Educativa, 2001; Wyatt, 1996).

VII. Sensitivity Analysis of the Model

We conducted a sensitivity analysis of the model to measure the sensitivity of final learning outcome (i.e., Reading Comprehension Score) to different variables in the model. The following is a summary of that analysis:

Most input variables have a minimal impact on the learning outcomes. For example, variables such as class size, student homogeneity, professional training of teachers, library have practically no impact on the learning outcome.

Of the input variables, the ones that have the largest impact are the ones that cannot be easily changed through policy interventions. These include the social status of the family and where they live (urban or rural).

Of the intermediate variables, student knowledge base has most impact on learning outcome. Other variables such as time learning curriculum outside school, time learning curriculum in school, time on homework, and curriculum coverage have the strongest impact on the learning outcome.

It is important to note that our model is a work in progress and sometimes small modification to the model would change the sensitivity of the outcome to some of the variables drastically. For instance, we believe that the learning outcome (i.e., reading score) would be more sensitive to some of the input variables (e.g., SpeaksLanguage). But as we found the data regarding those variables unreliable, we eliminated or revised some of the original links that we had in the model. For example, we originally had envisioned that SpeaksLanguage would have a direct impact on the Reading Score. This link was eliminated when we applied the model in Peru and Jamaica.

VIII. Conventional Policy Alternatives and Their Impact on the Model's Outcomes

The model provides an explanation of why most efforts to improve education have met with only limited success. As the previous section showed, changes in the input variables taken singly have only a small impact on the probability of good learning outcomes. Intermediate variables, on the other hand, have a much larger impact. The reason is obvious—the effect of distant input variables is mediated by a number of other factors, some beyond the control of the school and most not observed by policy designers. Scheerens (2005), summarizing a number of reviews of research studies on school effectiveness, shows that the effect sizes (an estimation of the size of the impact of changes in one variable on values of the other) are relatively small for resource measures such as class size, teacher training, teacher experience, and negative for changes in teacher salaries. School organizational factors (such as leadership, supervision, parental involvement) have a greater impact. The largest impact on learning outcomes, however, comes with changes in the instructional process, for example through increasing opportunity to learn, time on task, feedback and reinforcement of learning.

Table 1 below shows the results, for Peru and Jamaica, of simulating the introduction of policies that would affect different variables in the model. The last column in Table 1 shows the impact on the reading score according to the following coding scheme:

VS: Very Small Change; less than 1%

S : Small Change; 1%-2%

M : Moderate change; 2%- 4%

L : Large Change; more than 4%

Table 1: Policy Alternatives and the Variables they impact in the Model

Focus of Policy	Where Impacts	Impact on Reading Score	
		Peru	Jamaica
1. Time in School		VS	VS
a. Lengthen school year one month	OfficialDays		
b. Lengthen school day 2 hours	HoursDay		
c. Impose sanctions on teachers for non-attendance	TeacherAttendance		
2. Student Health		VS	VS
d. Provide snack or lunch for all students	Health		
e. Provide snack or lunch for all students below poverty line	Health		
f. Test eyesight of all students	Health		
3. Instructional Materials		VS	VS
g. Provide all schools with 500 book library	Library		
h. Provide each student with required textbooks	Textbooks		
i. Provide teachers with worksheets for all students (200 pages each)	InstMaterials		
4. Teacher Preparation			
j. Raise academic degree requirements to next highest level (e.g., BA to MA)	AcadEducn	VS	VS

k. Double length of pre-service professional training	ProfTrng	VS	VS
l. Provide teachers with 4 weeks of in-service training in methods	InServiceTrng	VS	VS
5. Teacher Assignment			
m. Assign most experienced teachers to early grades	TeachingSkill	S	S
6. Classroom Management		VS	VS
n. Train teachers (and principal) in	TeachingSkill		
o. methods to maintain discipline	PrincipalSkill		
p. Sort students so that classes are homogeneous in ability.	StudentHomogeneity		
q. Reduce class size by hiring more teachers	ClassSize		
7. Teacher Knowledge of Students		VS	VS
r. Train teachers in methods of student assessment.	TchrsKnowStudents		
s. Administer system-wide achievement tests in 4 th and 8 th grades and give individual student results to teacher	TchrsKnowStudents		
t. Reduce teaching schedule to allow teachers to meet with parents	TchrsKnowStudents		
8. Teacher Salaries			
u. Offer a bonus of 1 year's salary to all teachers who complete 7 years of service.	TchrExperience	VS	VS
v. Raise annual salary by 10% of teachers whose students score in upper 10% of achievement test	(no effect)		
w. Raise annual salary by 10% of all Teachers in schools that improve test scores compared to last year.	TeacherMotivation	S	S
9. Parental Involvement		VS	VS
x. Establish school councils with authority to hire and fire teachers	FamilyInvolvement		

z. Offer free kindergarten to all children	PriorEducation		
aa. Run media campaign to encourage parents to read to their children and help with homework		Family Involvement	
10. Compensatory Programs		M	M
ab. Require students who fail a grade to spend 4 weeks in a remedial program during vacation	AbilitytoLearn		
ac. Identify students with low reading ability and provide 2 hours intensive coaching per week	AbilitytoLearn		

IX. Likely Impact of Interventions that Change Intermediate Variables

Learning outcomes are the direct result of behaviors of teachers and students, the participants in the learning process. As a consequence, the policies most likely to have a strong impact on learning outcomes are those that change teaching and learning directly. The model specifies that amount of learning (what should be learned) is a function of time spent trying to learn, and the rate at which learning occurs.

Students learn (the official curriculum primarily in school, but also can learn outside. Three variables are linked directly with time spent learning in school. These are: the (amount of) opportunity the student has to learn; the (amount of) coverage of the content to be learned; and the student’s engagement in or attention to the teaching/learning process. Rate of learning depends on the quality of the instructional process, the knowledge base that student brings to the learning situation, and learning ability (innately determined but also learned over time).

The model shows, for Peru and Jamaica, that changes in each of these variables has a much larger impact on the probability of attaining a good learning outcome than do the input variables discussed earlier. Opportunity to learn is primarily a function of attendance, less an issue in Peru than in Jamaica, and this is seen in the results of the model. What is particularly striking is that in both countries the model reports low levels of coverage of the curriculum, and improvements in coverage would have very large impact on learning outcomes. This finding for Peru is consistent with an analysis by the World Bank (authored by Luis Crouch) that states that the lack of standards and accountability in schools in Peru results in an inconsistent and incomplete coverage of the curriculum (World Bank, 2007).

The model further indicates that a dramatic improvement in learning outcomes could be achieved by increased time learning curriculum outside of school. This finding is consistent with high average achievement scores in countries that include homework as an important element in instruction, and in countries in which many students receive out-

of-school tutoring. Student performance has been shown, in Latin America and elsewhere, to increase when parents are more actively involved in their children's learning at home and/or in the instructional work of the school. We take these experiences as validation of the results of the model.

X. Conclusion

In each of the four presentations we have now made of the model (2 in Peru, one in Jamaica, one in USAID/Washington), participants asked us to use the model to test the likely impact of alternative policies. When the model showed that changes in resources or other inputs would have little impact on learning outcomes, participants tried other changes, using the model to identify critical variables. As participants had different ideas about which policies might have more impact, they began to discuss their differences among themselves (rather than with us). In effect, the participants accepted the validity of the model, and used it to visualize the likely effects of alternative actions.

Participants were most interested, not unreasonably, in those variables that can be changed that have most impact on learning outcomes. They appeared to find it most easy to define policies to affect the resource/input variables, perhaps because these are most familiar to them. On the other hand, national discussions about projects with a strong conceptual definition, such as AprenDes, have prepared many to discuss policies to change instructional practices as well. The most difficult policies to represent in the model are those concerning variations in the quality, rather than quantity, of inputs. This difficulty reflects current disagreements in the education community about how best to prepare teachers and principals, and the absence of data assessing variations in quality of preparation. The model currently contains no information about the financial cost and difficulties of implementation of different policies, but this can be added in a future stage of development.

We attribute the lively discussions that resulted to several features of the model:

First, the logic of the underlying structure of the model is familiar to most educators and education policy makers. Perhaps no one has ever assembled and linked together all the variables the model contains, but all the components have been well studied and documented. At the same time, the use of a causal network linking variables together with probabilities makes it possible to see connections and influences in a fashion not previously possible for most audiences. This encourages dialogue.

Second, emphasis on the results of decisions, rather than on the statistical significance of relationships between variables allows practitioners and decision-makers alike to use their experiential knowledge.

Third, users can simulate the likely effect of a variety of different alternatives in a relatively short period of time. They can formulate hypotheses and test them out quickly. This allows them to develop a better understanding of how the specific education system (nation or school) operates. In this sense the model is an instructional tool.

XI. Contribution to Human and Institutional Capacity Building

The three presentations contributed to human and institutional capacity building in several ways. First, the presentations reinforced the importance of evidence-based decision making. Both the SSME data and the Policymakers' Workbench analysis challenged conventional policy interventions that affect only levels of resources (such as quantities of textbooks or levels of teacher training). The analyses presented evidence that these interventions while positive make only a small improvement in learning outcomes.

Second, the presentations demonstrated the importance of decision-makers acquiring and employing a systemic conceptual understanding of how schools and the teaching-learning process operate. Analyses done with the Policymakers' Workbench emphasized the complex nature of the school organization and the teaching process. Recognition of complexity leads to consideration of a wider range of alternative explanations of problems, and suggests a broader set of interventions to improve system performance.

Third, the presentation raised awareness of the limitations of current data collection in education. Decision-makers at present have little information about what actually happens in schools; development of a systemic conceptualization of the education system, based on previous research, makes it possible to identify the key variables on which to gather information. The ability of the model to identify key variables for which data is not now being collected prompted participants to discuss changes in data collection instruments.

Research and conceptualization are the prime missions of higher education institutions. The Policymakers' Workbench lends itself to both tasks: to identification of critical research questions and to the analysis of data for their resolution; and especially to teaching others how to think systemically and systematically about education and the organization and operation of schools.

In our opinion, the analyses generated using the Policymakers' Workbench are more easily understood by students and policy makers in developing countries than are reports based on the highly sophisticated statistical approaches currently used in technical assistance. Current approaches use an inductive thinking approach that is learned only by a small portion of the population. The logic of the Policymakers' Workbench is more closely linked to what we have learned from cognitive psychology and developments in neurology about how people learn from reasoning. Participants in our presentations had no difficulty in understanding the idea of the Workbench and in using it to answer questions, even when they did not understand its technical underpinning. We believe, therefore, that this approach is a more effective way to provide technical assistance in developing countries.

XI. Next Steps

The following provides a summary of possible next steps:

1. Collaborate with RTI and USAID to work out changes in and additions to the SSME, to be able to include more variables measuring internal processes in the model.
2. Improve the underlying model of the Policymakers' Workbench by incorporating additional variables based on a more detailed examination of the SSME, participants' suggestions in Peru and Jamaica, and also based on our own experience gained through the application of the model to the SSME data.
3. Conduct country level and school level comparative analysis using the model.
4. Add a decision component to the tool. That is, develop a mechanism for the user of the tool to evaluate various policy options through a cost/benefit analysis.
5. Add a knowledge component to the tool. This component would enable the user to draw upon relevant case studies and best practices in various countries.
6. Improve the user interface of the tool by incorporating the components discussed above and also by adding other functionalities such as an online glossary of terms and online help.
7. Improve the user interface by enabling the users to load individual school data and compare different schools.
8. Develop a web-based version of the tool which would enable users in various countries to run the tool directly using their own data.

Acknowledgement

We would like to thank Ms. Marilyn Crane and Ms. Meena Nabavi from HED for all their help in making this a smooth project, and RTI staff Ms. Amy Mulcahy-Dunn and Ms. Jenny Spratt for providing us the SSME data for Peru and Jamaica. We would also like to thank Ms. Sandra Bertoli (USAID) for her enthusiasm and unwavering support of our work. Finally, we owe a special thank to Dr. Louis Crouch (RTI) for his continual encouragement, advice and support since 2005 when we first started thinking about developing Policymakers Workbench.

Appendix 1: List of Participants in Peru

Participants, Meeting in USAID, 9 September, by Institution

Ministry of Education	10
USAID including mission director	5
Consejo Nacional de Educacion, including President	4
Proyecto Aprende	4
GRADE	3
Universidad Cayetano Heredia	3
Empresarios por la Educacion	2
Pontificia Universidad Catolica	1
Universidad Ruiz de Montoya	1
Preparate—IYF	1
UNICEF	1
UNESCO	1
Total	36

Participants, Meeting in Consejo Nacional de Educacion, 11 September

Consejo Nacional de Educacion	15
Alternativa	3
USAID	2
IDRC Canada	1
Instituto Nacional de Monterrico	1
Pontificia Universidad Catolica	1
Foro Educativo	1
Centro Peruano de Audicion, Lenguaje y Aprendizaje	1
Colegio San Felipe Neri	1
Total	26

Appendix 2: List of Participants in Jamaica

Participants, Meeting at the University of the West Indies, 15 September

About 35 participants from the following institutions attended the presentation:

Ministry of Education

USAID including the mission director

University of the West Indies and other universities

Prime Minister Office

Independent groups and Foundations

List of Attendees

1.	Mrs. Gertrude McKenzie	MOE (for Ms. Barbara Allen)
2.	Ms. Claire Spence	USAID
3.	Mr. Franklin Bennett	Mico
4.	Ms. Grace Munroe	
5.	Ms. Miranda Sutherland	
6.	Dr. Karen Hilliard	USAID
7.	Dr. Swithin Wilmot	
8.	Dr. Rose Davies	
9.	Ms. Jeanette Campbell	
10.	Mr. Frank Weeple	Transformation Team
11.	Ms. Mulcahy Dunn	RTI
12.	Dr. Massoud Mousavi	
13.	Prof. Stafford Griffith	
14.	Dr. Halden Morris	
15.	Ms. Jeanelle Babb	UNESCO
16.	Ms. Barbara Foster	
17.	Ms. Charmaine Montague	CASE
18.	Ms. Beverley Pullen	CASE
19.	Ms. Paulette Wickham	CASE
20.	Ms. Maxine Jones	CASE
21.	Mr. Joseph Thomas	CASE
22.	Ms. Sasha Parke	USAID
23.	Ms. Ruth Chisholm	USAID
24.	Dr. Loraine Cook	
25.	Ms. Dasmine Kennedy	MOE
26.	Ms. Joan Reid	MOE
27.	Mr. Lincoln Phipps	
28.	Ms. Dorothy Palmer	
29.	Dr. Susan Anderson	
30.	Dr. Disraeli Hutton	
31.	Mrs. Marceline Collins-Figueroa	
32.	Dr. Nadine Scott	
33.	Prof. Zellynne Jennings-Craig	

34. Dr. Paulette Feraria
35. Dr. Carol Gentles
36. Dr. Lorna Down
37. Prof. Hyacinth Evans

Appendix 3

TECHNICAL AND ADVISORY ASSISTANCE PROGRAM

Advisory Assistance to Apply Policymakers' Workbench in Dialogue with Policymakers in Peru and Jamaica

Quantitative Participant Information

ACTIVITY	NUMBER OF MEN						NUMBER OF WOMEN						DATES OF ACTIVITY	LOCATION	
	Faculty, Admin.		Students		Other		Faculty, Admin.		Students		Other				
	25 & under	26 & over	25 & under	26 & over	25 & under	26 & over	25 & under	26 & over	25 & under	26 & over	25 & under	26 & over		US	Host
Host Country Receiving non- training under this Training includes, limited to, seminars, special fee definition		10	5		5	15		10	5	5	5	16		11	70

training:

ation of the use of survey research data in diagnosis of problems in an education system, employing the PolicyMaker's Workbench tool.

Non-Degree Training:

A learning activity taking place in the U.S., a third country, or in-country in a setting predominantly intended for teaching or imparting knowledge and information to the participants with designated instructors or lead persons, learning objectives, and outcomes, conducted fulltime or intermittently.

The transfer of knowledge, skills, or attitudes (KSAs) through structured learning and follow-up activities, or through less structured means, to solve problems or fill identified performance gaps. Non-degree training can consist of short- or long-term technical courses in academic or in other settings, non-academic seminars, workshops, on-the-job learning experiences, observational study tours, or distance learning exercises or interventions.

Appendix 4

NOTES FOR ADVISORY ASSISTANCE

Higher Education for Development

Advisory Assistance to Apply Policymakers' Workbench in Dialogue with Policymakers
in Peru and Jamaica

Indicators of Institutional Capacity Strengthening

Indicator	How addressed (meetings, presentations, seminars, tutorials, written materials, other)
Curriculum development and products:	
Teaching methods:	
Research methods, topics, products:	Presentation
Publication plans:	Currently preparing an article for publication describing this experience in the development of the PolicyMakers' Workbench
New academic programs:	
Extension type work outside the institution:	
Internal management issues:	
Financial topics:	
Other:	

Indicators of Human Capacity Strengthening

Indicator	How Addressed (meetings, presentations, seminars, tutorials, written materials, other)
Technical/program training:	2 hour presentation to three audiences (81 persons)
Methods for teaching:	PowerPoint and participant engagement using an interactive model
Faculty consultations:	
Individual consultations:	
Other:	

References

- Cabrales, E. J., & Eddy, J. P. (1992). Developing a simulation model for the development of higher education administrators. *Journal of Research on Computing in Education*, 25(1), 105-112.
- Carroll, J. (1963). A model of school learning. *Teachers College Record*, 64, 723-733.
- Coe, R., & Fitz-Gibbon, C. T. (1998). School effectiveness research: criticism and recommendations. *Oxford Review of Education*, 24(4), 421-438.
- Duyar, I. (2006). Analyzing education productivity: A review. *Education Review*, 9(4), Retrieved 23 February 2008 from <http://edrev.asu.edu/essays/v9n4/>.
- Fuller, B. (1987). What school factors raise achievement in the Third World? *Review of Educational Research*, 57(3), 255-292.
- Glewwe, P., & Kremer, M. (2005). Schools, Teachers, and Education Outcomes in Developing Countries. In E. A. Hanushek & F. Welch (Eds.), *Handbook on the Economics of Education* (Vol. 2,). Amsterdam: Elsevier.
- Haertel, G. D., Walberg, H. J., & Weinstein, T. (1983). Psychological models of educational performance: A theoretical synthesis of constructs. *Review of Educational Research*, 53(1), 75-91.
- Hanushek, E. A. (2006). School resources. In E. A. Hanushek & F. Welch (Eds.), *Handbook of Economics of Education* . Amsterdam: Elsevier.
- Helmsley-Brown, J., & Sharp, C. (2003). The use of research to improve professional practice: a systematic review of the literature. *Oxford Review of Education*, 29(4), 449-470.
- Leclerq, F. (2005). *The Relationship between Educational Expenditures and Outcomes* (Working Document DT/2005-05). Paris: DIAL.
- Marzano, R. J. (2003). *What Works in Schools: Translating research into action*. Baltimore, MD: Association for Supervision and Curriculum Development.
- Mella, O., & Ortiz, I. (1999). Rendimiento escolar. Influencias diferenciales de factores externos e internos. *Revista Latinoamericana de Estudios Educativos*, XXIX(001), 69-92.
- Nascimento, P. A., & Meyer, M. (2008). School resources and student achievement: worldwide findings and methodological issues. *Educate, Special issue* (March 2008), 19-30.
- Raudenbush, S. W., & Bryk, A. S. (1988-1989). Methodological advances in analyzing the effects of schools and classrooms on student learning. *Review of Research in Education*, 15, 423-475.
- Reynolds, D., & Teddlie, C. (1999). *The International Handbook of School Effectiveness Research*. London: Routledge.
- Russell, S., & Norvig, P. (2003). *Artificial Intelligence: A modern approach*. Inglewood Cliffs, NJ: Prentice Hall.
- Sammons, P., & others, a. (1995). *Key Characteristics of Effective Schools: A review of school effectiveness research*. Ringwood, Hants, England: B & MBC Distribution Services.
- Scheerens, J. (2000). *Improving School Effectiveness*. (Vol. No. 68). Paris: UNESCO International Institute for Educational Planning.

- Scheerens, J. (2005). *Review and Meta-Analyses of School and Teaching Effectiveness Research*. Twente, Holland: University of Twente.
- Simmons, J., & Alexander, L. (1978). The Determinants of School Achievement in Developing Countries: A Review of the Research. *Economic Development and Cultural Change*, 26(2), 341-357.
- Subotnik, R., & Walberg, H. (Eds.). (2006). *The Scientific Basis of Educational Productivity*. Greenwich, CT: Information Age Publishing.
- Unidad de Medición de la Calidad Educativa. (2001). *Efecto de la escuela en el rendimiento en lógico-matemático en cuarto grado de primaria*. Lima: Ministerio de Educación.
- World Bank. (2007). *Toward High-quality Education in Peru: Standards, accountability and capacity building*. Washington, DC: The World Bank.
- Wyatt, T. (1996). School effectiveness research: Dead end, damp squib or smouldering fuse? *Issues in Educational Research*, 6(1), 79-112.