



EDDATA II

Snapshot of School Management Effectiveness

Peru Pilot Study



July 2008

**EdData II Technical and Managerial Assistance, Task Number 1
Contract Number EHC-E-01-04-00004-00
Strategic Objective 3**

This publication was produced for review by the United States Agency for International Development. It was prepared by RTI International and Fundación para el Desarrollo Agrario.

Snapshot of School Management Effectiveness

Peru Pilot Study

Prepared for

Bureau for Economic Growth, Agriculture and Trade (EGAT/ED)
United States Agency for International Development

Consultant team, Fundación para el Desarrollo Agrario (FDA)

Fernando Rosas, General Coordinator; Jorge Luis Bazán, Data Analysis Specialist; Oscar Millones, Specialist in Instruments; Jesus Salinas, Data Processing Specialist

Prepared on behalf of

RTI International
3040 Cornwallis Road
Post Office Box 12194
Research Triangle Park, NC 27709-2194

RTI International is a trade name of Research Triangle Institute.

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Table of Contents

	Page
Figures	v
Tables	v
I. INTRODUCTION	1
II. PILOT STUDY	1
2.1 Instruments applied	2
Stage 1: Preliminary adaptation	2
Stage 2: The pre-pilot application	2
Stage 3: Definitive adaptation	3
Observations	5
2.2 Sample design	6
Target population	6
Sample	6
2.3 Training activities	7
2.4 Organizing field work	9
Guidelines for interviewers during fieldwork	10
2.5 Collecting and handling data	11
Getting to the IEs selected for the pilot study	11
Control of the quality of information	11
Building the database	12
III. RESULTS	13
3.1 Results of the field study	13
Department of La Libertad	13
Department of Junín	14
Department of Ancash	15
Department of Lima	16
3.2 Results of the application time	17
3.3 SSME results	18
3.4 EGRA results	19
Systems for EGRA coding	19
Results of the EGRA sections	20
Exploring the global EGRA score	22
Formulating an EGRA for count and speed	23
Putting together an EGRA dichotomic	30
3.5 SSME and EGRA relationships	34
4. DISCUSSION	35
4.1 Fieldwork results	35
Positive aspects	35
Aspects that need to be improved in future applications	36
4.2 Regarding the application time for instruments	37
4.3 The SSME results	37
General modifications	38
Specific modifications	38
4.4 EGRA results	42
4.5 Relationships between SSME and EGRA	42

V. RECOMMENDATIONS	43
5.1 Sampling recommendations	43
5.2 Recommendations to modify the instruments used in the national application	44
5.3 Recommendation for training for interviewers and materials	45
5.4 Recommendations for fieldwork organization	45
5.5 Recommendations for gathering and handling data	46

Annexes

Annex 1	Instruments adapted to the Peruvian education system
Annex 2	Users' guide for SSME and EGRA
Annex 3	Instruments applied during the pilot study
Annex 4	List of IEs participating in the pilot study
Annex 5	Basic statistics on SSME instruments
Annex 6	Basic statistics on EGRA instruments
Annex 7	Database for the pilot study
Annex 8	SSME responses

[NOTE: Annexes are available only in the Spanish version of this report, *Instantánea de efectividad en la administración de centros educativos: Estudio piloto in Perú*]

Figures

Figure 1.	Administrative network for the implementation of the Peru pilot	8
Figure 2.	Members of the operations team for the implementation of the Peru pilot	8
Figure 3.	Training the operations team for the application of the pilot in Peru	9
Figure 4.	Education institution participating in the pilot	11
Figure 5.	Effect of the number of elements read in the EGRA	27
Figure 6.	Percentage of incorrect responses in section I	31
Figure 7.	Percentage of incorrect responses in section V	32
Figure 8.	Percentage of nonresponses (unable to read) in section V	32
Figure 9.	Percentages of incorrect responses in section VI, listening comprehension	32
Figure 10.	Percentages of non-response (not read) in section VI	33
Figure 11.	Percentages of incorrect responses for questions in section VII	33
Figure 12.	Percentage of nonresponse (not read) in section VII	33

Tables

Table 1.	Instruments applied in the pilot study	2
Table 2.	Modifications made to instruments following the pre-pilot study	4
Table 3.	Effective time utilized in pre-pilot activities	4
Table 4.	Sample of education institutions by department, according to strata	7
Table 5.	Instruments applied in the pilot study	9
Table 6.	Database for the instruments applied in the pilot study	12
Table 7.	Database structure of the instruments applied in the pilot study	12
Table 8.	Time of application: Averages and standard deviations (SDs) for fieldwork activities in the IEs, by department	17
Table 9.	Modifications made to the instruments following the pilot study	18
Table 10.	System to code the EGRA by section and format type	19
Table 11.	Descriptive statistics, reliability, and the EGRA normality test	20
Table 12.	Matrix of correlations of EGRA sections	21
Table 13.	Analysis of the EGRA items	22
Table 14a.	Descriptive statistics and correlation matrix for reading times	24

Table 14b.	Descriptive statistics and correlation matrix for elements read	24
Table 14c.	Descriptive statistics and correlation matrix for elements read incorrectly	25
Table 14d.	Descriptive statistics and correlation matrix for total elements read correctly.....	25
Table 15.	Cronbach's alpha reliability for a different number of elements read, according to the order in which they appear	28
Table 16.	Descriptive statistics, item analysis, and matrix of correlations for reading speed.....	29
Table 17.	Comparison of performance in reading speed for a text passage according to some student characteristics	35

I. INTRODUCTION

This document is the final report of the study to adapt the Snapshot of School Management Effectiveness (SSME) and Early Grade Reading Assessment (EGRA) instruments—provided by the Research Triangle Institute (RTI) to the Foundation for Agricultural Development (FDA)—to the Peruvian reality. The study was conducted from September 2007 to February 2008 and consisted of an initial stage for pre-pilot application in 4 education institutions (instituciones educativas, IEs) in Lima and a second stage for pilot application in 64 IEs in 4 departments (Ancash, La Libertad, Junín, Lima) in Peru.

The set of SSME and EGRA instruments includes a battery of questionnaires that are applicable to IEs for the initial years of education in primary school and that provide quick information regarding school management and “pre-reading” skills. This battery includes questionnaires for directors of IEs, teachers, parents, and students, and is directed to boys and girls of 2nd, 3rd and 4th grade of primary school. The idea, in general, is to prove the feasibility of using a series of documents that allow for a “quick snapshot” of an IE and, based on the collection of randomly selected snapshots, to develop a “portrait” of the education sector at the IE level.

It is very important to point out that the study does not attempt to generalize this portrait to the Peruvian educational reality but is instead geared to testing the validity and reliability of a series of instruments. RTI chose Peru and Jamaica as test countries for a trial run of the instruments and to study the behavior, feasibility, and utility of their application. The intent is that if the instruments and the processes utilized in fieldwork are useful and viable, it will be possible to offer the instrument to researchers who would like to use them (or certain aspects of them) in Peru or in other countries.

The document is organized in the following way: Section I is an introduction that contains the general aspects of the study; Section II provides details on several aspects of the pre-pilot and pilot studies; Section III provides the results obtained in the pilot study; and Section IV presents a series of recommendations for future work. Finally, in the annexes, the reader will find information on various complementary aspects.

II. PILOT STUDY

This section presents the methodology utilized by the FDA technical team to obtain the final version of the instruments applied in the pilot. Next, it provides details on the instruments applied in the pilot as well as the characteristics of the sample design. Details also are supplied regarding the training of interviewers and the training materials used, as well as the organization of the fieldwork for the pilot study. Data collection and management, fieldwork observations gathered in each department during the pilot study, quality control for fieldwork data, and actions taken to develop a database are also described in detail.

2.1 Instruments applied

The list of the instruments applied in the study appears in Table 1.

Table 1. Instruments applied in the pilot study

<i>Instruments</i>		<i>Name</i>
SSME	1	Classroom observation
	2a	Student questionnaire
	3	Matrix of class activities
	4	Parent questionnaire
	5	Director's questionnaire
	6	Observation of the education institution
	7	Teacher questionnaire
EGRA	2b	Pre-reading

The work methodology that FDA followed to obtain the final version of the instruments to be applied in the pilot study was divided into the following phases or stages:

Stage 1: Preliminary adaptation

During this stage, the language in the English-to-Spanish translation was adapted to the daily language used in Peru. The following terms were modified: "colegio" to "institución educativa" (education institution); "asociación/comité de padres" to "APAFA" (asociación de padres de familia [parent association]); "planes de lección" to "programas de lección" (lesson plans); "observaciones formales" to "observaciones programadas" (formal classroom observations); etc. We introduced synonyms for specific words such as salón/aula (classroom), retrete/inodoro/water (latrine), etc.

Stage 2: The pre-pilot application

The application dates for the pre-pilot were the 17th and 18th of October, 2007. The location was the city of Lima and 4 educational institutions were selected: two located in Local Education Management Unit (Unidad de Gestión Educativa Local, UGEL) 07¹ and two in UGEL 05² in the province of Lima. The IEs selected fulfilled the criteria of inclusion in the study given that they are state-run institutions that are full grade³ at the primary school level. The participants in the pre-pilot study included 4 directors, 16 students, 4 teachers, and 6 parents.

The FDA technical team—in coordination with the Ministry of Education (Ministerio de Educación, MINEDU)—put together an operations team to conduct the fieldwork. This team consisted of a coordinator and two supervisors.

¹ IE Virgen Milagrosa de San Borja and IE Albert Einstein de Surquillo.

² IE 1186 Santa Rosa de Lima Milagrosa and IE 1187 San Cayetano de El Agustino.

³ "Full grade" means there are no teachers who teach several grades in one classroom.

RESULTS OF THE PRE-PILOT STUDY

- We found that the application time for the instruments exceeded the school day⁴; as such, we proposed using two days to apply the instrument in each IE.
- We found that the directors of the IEs were not always present; as such, we suggested that the application team be in charge of directly delivering authorizations from the Ministry of Education (in addition to the direct route established for the Ministry to deliver authorization to the UGELs).
- We found that parents are not always present in the IE, which led us to propose that parents be given one day's notice of an interview appointment for the next day.
- The fact that the consent forms must be signed was a condition that led some teachers to refuse to participate due to the growing controversy between teachers and government stemming from the last teacher evaluation. For this reason, we proposed not requiring a signature on the document, which would reinforce the anonymous nature of participation in the interview process.
- The application time for the instrument administered to the director of the UGEL constitutes a critical point due to the overlap that exists with the Ministry of Education's own fieldwork applications, which require the direct participation of UGEL officials and their personnel. We proposed not applying this instrument because it is impossible to guarantee the participation of UGEL officials.

Stage 3: Definitive adaptation

As a result of the application of the pre-pilot, the FDA technical team considered it a good idea to divide instrument modification into two points: general and specific.

GENERAL MODIFICATIONS

The format for the original instrument was modified to facilitate the interviewer's work to collect information and to provide fluidity to the coding work and the process to gather information. Given that the initial questions on the original instruments did not permit an adequate identification of the sample, location of application, education institution, and the grade for application, we added different identification categories such as: questionnaire data (questionnaire code, name and code of interviewer), general data (department code, province, district, and zone), data on the education institution (UGEL code, name and school identification code) and classroom data (grade code, classroom code and name of the teacher), which vary according to the instrument to be utilized.

⁴ In Peru, the school day for primary is generally 8 a.m. to 1 p.m. and many schools do not work in the afternoon. The IEs that work at the primary and secondary level have a more restricted schedule because they work during two shifts. In this case, secondary is from 1 p.m. to 6 pm.

SPECIFIC MODIFICATIONS

The magnitude of the modifications made to the SSME is shown in Table 2.

Table 2. Modifications made to instruments following the pre-pilot study

<i>Description</i>	<i>1</i>	<i>2a</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>Total</i>
Total of original questions that were not modified	8	15	0	18	47	13	35	136
Total of questions containing changes to words	1	9	0	5	22	6	12	55
Total of questions to which alternative responses were added	2	4	0	0	4	0	1	11
Total of questions with changes to words and added response alternatives	0	0	0	1	0	0	0	1
Total of questions added	0	1	0	0	0	0	0	1
Total of questions eliminated	0	0	0	0	0	0	0	0
Total of original questions (other than general data)	11	29	0	24	73	19	48	204

We found that 67% of the questions in all the questionnaires were not modified while 27% were subjected to changes in words.

STUDY TIME

Details on the effective time used in each of the activities in the pre-pilot can be found in Table 3. It does not include the time used to find the interview subject or the time that transpired between one activity and the next.

Table 3. Effective time utilized in pre-pilot activities

<i>Activity</i>	<i>1186 Santa Rosa de Lima Milagrosa</i>	<i>1187 San Cayetano</i>	<i>Virgen Milagrosa</i>	<i>Albert Einstein</i>
Previous coordination with the director and IE observation (Instrument 6)	40	30	40	30
Interview with the director (Instrument 5)	50	50	50	50
Observation 1 (Instrument 1)	15	15	15	15
Observation 1 classroom 1 (Instrument 3)	10	10	10	10
Observation 1 classroom 2 (Instrument 3)	10	10	10	10
Interview of classroom teacher 1 (Instrument 7)	40	40	40	40

<i>Activity</i>	<i>1186 Santa Rosa de Lima Milagrosa</i>	<i>1187 San Cayetano</i>	<i>Virgen Milagrosa</i>	<i>Albert Einstein</i>
Student questionnaire classroom 1 (Instrument 2)	120	120	110	115
Observation 2 (Instrument 1)	10	10	10	10
Observation 2 classroom 1 (Instrument 3)	10	10	10	10
Observation 2 classroom 2 (Instrument 3)	10	10	10	10
Interview of classroom teacher 2 (Instrument 7)	40	40	40	40
Student questionnaire classroom 2 (Instrument 2)	120	110	105	110
Observation 3 (Instrument 1)	10	10	10	10
Observation 3 classroom 1 (Instrument 3)	10	10	10	10
Observation 3 classroom 2 (Instrument 3)	10	10	10	10
Parent Interview (Instrument 4)	50	45	50	40
Time without the parent interview	505	485	480	480
Total time with parent interview	555	530	530	520
Total time with parent interview in hours	9 h 15 m	8 h 50 m	8 h 50 m	8 h 40 m

As can be seen, the times used in each IE exceed an eight-hour work day.

Observations

- The quantity of instruments is excessive with relation to the time available for application (1 day⁵). The corrective action would be to schedule application in each education institution over two days. We propose that the first day be used to apply the following instruments: teacher questionnaire, student questionnaire with children in a classroom, classroom observation instruments, matrix of class activities, director questionnaire, and instrument for the education institution. On the second day, it would be feasible to apply the same instruments to the second classroom plus administer the parent questionnaire.
- The presence of the director of the institution cannot be guaranteed given that he/she participates in activities outside of the educational institution. The preventive action would be to give the directors a letter from MINEDU in timely manner. The technical team should also take steps to gain a prior commitment from the director to participate.

⁵ As noted in the Introduction, in Peru, the school day for primary is generally 8 a.m. to 1 p.m. and many schools do not work in the afternoon. The IEs that work at the primary and secondary level have a more restricted schedule because they work during two shifts. In this case, secondary is from 1 p.m. to 6 pm.

- Parents are not always present at the education institution; a preventive action would be to make an interview appointment with parents from one day to the next through the IE director.
- Signing of consent forms is a condition in our political reality that impedes teacher participation. This reluctance is due to recent conflicts between teachers and the government over a recent evaluation process. For this reason, the preventive action would be to eliminate the signature requirement, which would also reinforce the anonymous nature of the interview process.
- Applying the instrument to the UGEL director constitutes a critical point due to overlapping with field applications conducted by MINEDU officials; the aforementioned instruments require the direct participation of UGEL officials and their personnel. The corrective action would be to eliminate this instrument's application because it is impossible to guarantee the participation of UGEL officials.

2.2 Sample design

Target population

Given the fact that this was a pilot study that focused more on the instrument itself, the results cannot be generalized to all education IEs at the primary level in Peru. We define the target population as public IES that are full grade and that have a minimum 3 sections of grade 2, 3 and 4 of primary in a regular basic education format. The sample framework consisted of 6,720 IEs. This population was stratified considering two points: geographic location and IE size. In the first stratification criterion, two types were considered: rural and urban. The urban category could be further stratified as urban marginal, urban, shantytown, or residential urban. The second criterion differentiated the education institutions into small IEs that had exactly 3 sections in the aforementioned grades and large IEs that had more than 3 sections of said grades.

Sample

Given that the study's objective basically was to test the instrument without achieving countrywide representation, we chose a judgment sample that included the following departments: Ancash, La Libertad, Junín, and Lima.⁶

Based on the stratification criteria used and keeping in mind the objective to maximize the representation of each department, we chose one UGEL from each department that supervised predominantly rural IEs and one UGEL in an area that was preponderantly urban.

⁶ Ancash and La Libertad are departments in the northern part of Peru that have IEs in both the sierra and coastal areas. Junín is a department in central Peru that has IEs in the sierra and the Peruvian coast. Lima is the country's capital and its schools are located along the coast.

Next, a drawing was held for each UGEL that participated in the pilot study so that the IEs would correspond to the strata defined for the target population. The total of IEs for each UGEL defined in the terms of reference was 8.⁷

Finally, in each IE sampled, we interviewed the director and one or two parents selected in a semi-random manner. Additionally, we sampled 2 classrooms with students from 2nd, 3rd, and/or 4th grade of primary. In each classroom selected, we interviewed a classroom teacher and randomly selected 4 students (2 boys and 2 girls).

The distribution of IEs according to departments and the strata considered in the design sample can be found in Table 4. Details of the participating IEs are shown in Annex 4.

Table 4. Sample of education institutions by department, according to strata

<i>Strata</i>	<i>Ancash</i>	<i>Junín</i>	<i>La Libertad</i>	<i>Lima</i>	<i>Subtotal of IEs</i>	<i>Total</i>
Small rural	5	4	3	1	13	20
Large rural	2	3	1	1	7	
Small urban	3	3	2	1	9	44
Large urban	4	5	7	6	22	
Small marginal urban	0	0	0	1	1	
Large marginal urban	2	1	1	3	7	
Large urban shantytown	0	0	2	3	5	
Total departments	16	16	16	16		64

2.3 Training activities

The following administrative network (Figure 1) was developed for the pilot study.

The FDA technical team participated in the planning, implementation and control of all the study's activities. The operating team comprised a coordinator who was in charge of two supervisors. The latter coordinated and supervised the interviewers' work.

⁷ With the exception of the Department of Junín, where 7 IEs were considered for the UGEL Chanchamayo and 9 IEs for the UGEL Huancayo due to the relative importance of the latter.

Figure 1. Administrative network for the implementation of the Peru pilot

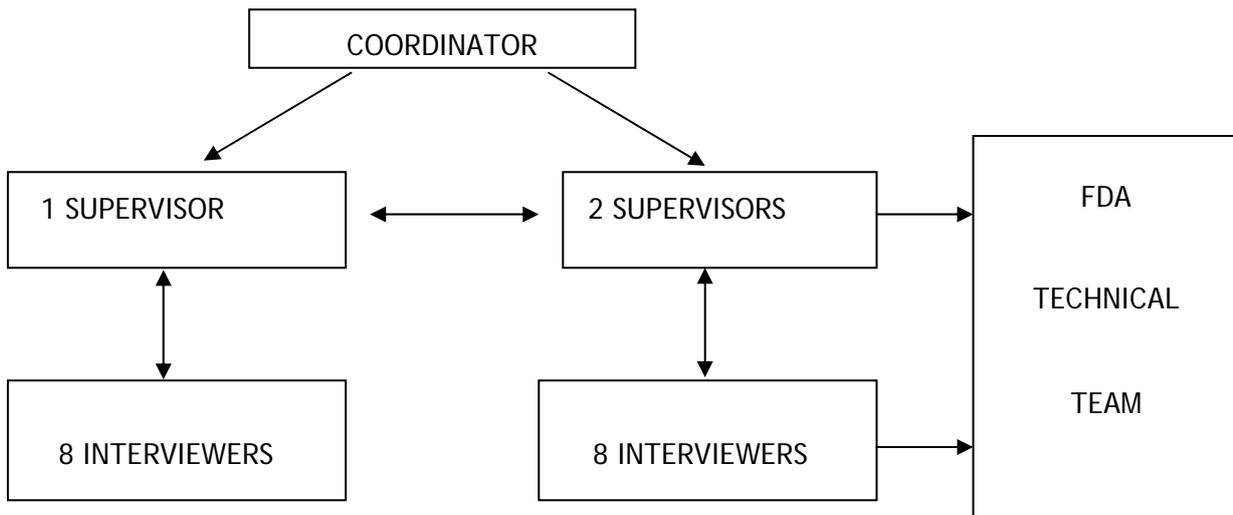


Figure 2. Members of the operations team for the implementation of the Peru pilot



Activities that were conducted prior to training:

- Contact and selection of interviewers.
- Requirements for materials to be used for the pre-pilot.
- Contact with the IE directors where the pre-pilot would be applied.
- Pick-up of MINEDU credentials for the IEs where the pre-pilot would be applied.
- Review of the access routes for the IEs and the instrument application sequence.

The operations team's training was conducted in the following stages:

a) Training for the pre-pilot

The objective of this training was to ensure that coordinators and supervisors would understand and adequately apply the study's instruments. This training was conducted

between October 11th and 12th of 2007. Training was held at the FDA headquarters and a specialist in instruments was in charge of the technical team.

b) Pilot training

The objective of this training was to ensure that interviewers would understand and adequately apply the studied instruments and to develop awareness of their responsibilities during the pilot stage. It was conducted between the 1st and 2nd of November 2007. The training was conducted at the headquarters at the Center for Pre-University Studies of La Molina National Agricultural University. The coordinator and the supervisors were in charge of the training.

During the training period, participants were made aware of: the objectives of the 2007 pilot evaluation; the training schedule; the responsibilities of each interviewer; and the tasks that must be conducted before, during, and after the application. Discussions took place regarding appropriate use of the manual and correct administration of the instruments. This training was complemented by exercises to simulate use of the instruments and resolution of situations that might arise during the fieldwork.

Figure 3. Training the operations team for the application of the pilot in Peru



2.4 Organizing field work

According to the results of the pre-pilot study, we agreed that the application at the IEs would be conducted over two days. In Table 5, we provide a schedule for the implementation of the pilot study.

Table 5. Instruments applied in the pilot study

<i>Instruments</i>		<i>Name</i>	<i>Days</i>
SSME	1	Classroom observation	1 and 2
	2a	Student questionnaire	1 and 2
	3	Matrix of classroom activities	1 and 2

<i>Instruments</i>		<i>Name</i>	<i>Days</i>
	4	Parent questionnaire	1 or 2
	5	Director's questionnaire	1 or 2
	6	Observation of the education institution	1
	7	Teacher questionnaire	1 or 2
EGRA	2b	Pre-reading	1 and 2

Guidelines for interviewers during fieldwork

GUIDELINES FOR SELECTING THE STUDENT SAMPLE

1. The interviewers should ask the director how many classrooms of 2nd, 3rd, or 4th grade of primary exist in the IE.
2. The interviewer should choose two classrooms, preferably from the 2nd grade. If it is not possible to apply these instruments to the 2nd grade, interviewers should proceed to select classrooms from 3rd grade, or if needed to complete the sample, from the 4th grade.
3. In the classrooms selected, the interviewers should compile a random list of four students, including two boys and two girls.

GUIDELINES FOR THE FIRST DAY OF APPLICATION

1. The interviewers will enter classroom 1 and apply Instrument 1, classroom observation.
2. The interviewer will administer the instruments to the students selected from classroom 1 on an individual basis outside the classroom and preferably in a private setting (director's permission should be sought prior to beginning the process). This applies to Instrument 2a, student questionnaire, and the EGRA instrument (Instrument 2b). At the end of the instrument application process, students will be given a pencil case.
3. Instrument 3, matrix of classroom activities—The interviewers should apply this instrument at three points in time: at the beginning, in the middle, and at the end of class, and may determine the specific times for application.
4. The times for administering the teacher and director questionnaires (Instruments 7 and 5 respectively) will be coordinated with the respondents. The director questionnaire can be applied on this day or the next.
5. We suggest that the interviewer apply Instrument 6, observation of the education institution, at the end of the morning.
6. In case a parent is at school, even if that parent's child is not in the classroom being evaluated, apply the corresponding questionnaire (Instrument 4, parent questionnaire). If no parents are present in the school, the classroom teacher will be asked to make an appointment with the parents for the following day so that the questionnaire can be applied.

GUIDELINES FOR THE SECOND DAY OF APPLICATION

1. The interviewer will enter classroom 2 to apply Instrument 1, classroom observation.
2. The interviewer will apply instruments to the students selected in classroom 2. Students will be called one by one and interviewed (preferably) outside the classroom in a private environment (this should be requested of the director beforehand). This applies to Instrument 2, student questionnaire; and to the EGRA instrument. At the end of the questionnaire application, students will be given a pencil case.
3. Instrument 3, matrix of classroom activities—The interviewer will apply this at three points in time: at the beginning, in the middle, and at the end of class. The interviewer can determine application times.
4. The time for administering Instrument 7, teacher questionnaire, will be coordinated with the teacher.
5. The parent questionnaire is applied to the individual that has been given an appointment (Instrument 4, parent questionnaire). If this parent is not present at school, the questionnaire will be applied to any parent present at the school.

2.5 Collecting and handling data

Getting to the IEs selected for the pilot study

In the first week (November 5-8), two teams made up of a supervisor and 8 interviewers traveled to La Libertad and Junín and applied the instruments simultaneously in the UGELs selected in each of the departments. In the second week, the same teams conducted a simultaneous application in the UGELs selected in the department of Ancash (12-13 of November) and Lima (15-16 of November).

Figure 4. Education institution participating in the pilot



Control of the quality of information

Quality-control checks of information gathered in the pilot study were carried out in three phases:

- Phase 1: At the end of the application of each UGEL instrument; consisted of a joint review by the supervisor and interviewers of the quality of the information gathered by the interviewer.
- Phase 2: At the end of application of instruments in all IEs; consisted of coding of the open questions and a second review of the quality of the information gathered. This was conducted by supervisors.
- Phase 3: At the end of the process to digitize the instruments; consisted of the technical team's review of the incoherencies in the digitizing process. This process was led by a specialist in information generation processes.

Building the database

The process to digitize the SSME and scan the instruments applied was handled by a specialized company known as POLYSYSTEMS under the supervision of the technical team led by a specialist in information generation processes. The list of databases can be found in Table 6. The list of each of the data structures of each of these databases can be found in Table 7. The data are stored in SPSS files (*.sav).

Table 6. Database for the instruments applied in the pilot study

<i>Instruments</i>		<i>Name of Instrument</i>	<i>Name of database</i>
SSME	1	Classroom observation	Instrumento_1_Observación_de_aula.sav
	2a	Student questionnaire	Instrumento_2_Cuestionario_del_alumno_y_EGRA.sav
	3	Matrix of class activities	Instrumento_3_Matriz_de_actividades_de_clase.sav
	4	Parent questionnaire	Instrumento_4_Cuestionario_del_padre.sav
	5	Director's questionnaire	Instrumento_5_Cuestionario_del_Director.sav
	6	Observation at the education institution	Instrumento_6_Observación_de_la_IE.sav
	7	Teacher questionnaire	Instrumento_7_Cuestionario_del_Docente.sav
EGRA	2b	Pre-reading	Instrumento_2_Cuestionario_del_alumno_y_EGRA.sav

Table 7. Database structure of the instruments applied in the pilot study

<i>Instruments</i>		<i>Name of the Instrument</i>	<i>Name of the database</i>
SSME	1	Classroom observation	Estructura_Instrumento_1_Observación_del_aula.doc
	2a	Student questionnaire	Estructura_Instrumento_2_Cuestionario_del_alumno_y_EGRA.doc

<i>Instruments</i>	<i>Name of the Instrument</i>	<i>Name of the database</i>	
	3	Matrix of class activities	Estructura_Instrumento_3_Matriz_de_actividades_de_clase.doc
	4	Parent questionnaire	Estructura_Instrumento_4_Cuestionario_del_padre.doc
	5	Director's questionnaire	Estructura_Instrumento_5_Cuestionario_del_Director.doc
	6	Observation of the education institution	Estructura_Instrumento_6_Observación_de_la_IE.doc
	7	Teacher questionnaire	Estructura_Instrumento_7_Cuestionario_del_docente.doc
EGRA	2b	Pre-reading	Estructura_Instrumento_2_Cuestionario_del_alumno_y_EGRA.doc

III. RESULTS

The following sections contain results for

- a) The pilot study fieldwork
- b) Application time
- c) SSME
- d) EGRA
- e) Connections between SSME and EGRA

3.1 Results of the field study

During the process to collect data, a set of detailed observations was formulated by the team of interviewers and supervisors involved in field work. Below we identify the achievements and difficulties experienced during fieldwork in the departments included in the sample.

Department of La Libertad

ACHIEVEMENTS:

- a) All of the instruments were applied without difficulties according to the established schedule.
- b) The directors of the IEs were always cooperative; this also applies to the teachers and parents in general; all the children wanted to be interviewed.
- c) The test administrators were able to work well together to achieve the goal set.
- d) An adequate environment was set aside for student interviews; students were motivated to participate in the questionnaire because they would receive a pencil case as a gift.

DIFFICULTIES:

- a) The authorizations from the UGELs corresponding to instrument application had not arrived at each IE address; as such, directors were often unfamiliar with this project. [However, some went directly to the UGEL to verify the veracity of the application of these instruments and to ensure that the interviewers had the corresponding authorization to enter the IEs and apply said instruments. In IE José Carlos Mariátegui, personnel denied access to continue instrument application on the second day because this authorization had not been received. To resolve this problem, we contacted the Regional Office of Trujillo, which sent a primary school specialist from the pedagogical unit.]
- b) Some IEs had characteristics that were not foreseen in the planning stage. [The IE José Faustino Sanchez Carrión was a Large School Unit for boys. For this reason, 4 interviews were conducted with female students at the IE República de Panama. The IE 80061 was under construction and was operating temporarily at another IE, which made it difficult to locate. As such, we found only one 2nd-grade class and interviewed a 3rd-grade section. The IE Daniel Alcides Sanchez Carrión was difficult to access and we found that many teachers were absent, particularly in the 2nd grade. As such, we applied the instrument to a section of the 3rd grade.]
- c) In the student questionnaire, children were confused by some of the questions; for example, for the question “si en su casa tienen lavadora” (“if there is a washer in your house”), the students answered “yes,” but curiously enough these students had no electricity and we had to clarify that we meant an actual washing machine as opposed to the sink at which their mothers washed clothes. This will eventually require improvement.
- d) In some of the IEs, the pedagogical day transpires over two shifts, primary in the morning and secondary in the afternoon. Both of these shifts use the same classroom. As such, secondary school students frequently destroy the work done by children in the morning (e.g., posters put up on the classroom walls), which means that the classroom does not fulfill the conditions for “appropriate education material” during the observation. This was the case of IE Santo Domingo.
- e) The low level of student learning was an impediment to applying instruments such as the pre-reading. Students were unable to complete the instrument in a satisfactory manner in the time given (IE 80021).

Department of Junín

ACHIEVEMENTS:

- a) All of the instruments⁸ were applied as planned. At the UGEL Chanchamayo, instruments were applied over a one-day period in the 7 IEs chosen due to the announcement of a regional strike for Thursday, November 8 (second day of application). In the morning, the instruments were applied to students and teachers and in the afternoon parents and directors were interviewed. The interviewer Liliana

⁸ Except for one parent questionnaire; because it was Monday (market day), the parents were not in the vicinity.

Avellaneda and the supervisor supported the other members of the group to meet the goal of finishing work by 6 in the evening.⁹

- b) The majority of the directors received the authorization sent by MINEDU.

DIFFICULTIES:

- a) Five of the eight IEs in the UGEL Huancayo had not received the authorization regarding the application of the diagnostic. Some of the authorizations that did arrive did not include the director's name.
- b) Some IEs had presented characteristics that were not foreseen in the work plan. [In IE 30202, the diagnostic was held in two different locations. In IE 30138 the application began after 10 in the morning, because the director was not present and the teachers opposed instrument application. In IE 30006, Ana Mayer, the 2nd-grade teacher was opposed to instrument application and we could only work with two children. The teacher was absent despite the fact that the director had given authorization. As such, the work was not concluded and coordinator Verónica Zuñiga was consulted regarding this point. She gave permission to apply the instruments in IE Ricardo Menendez. At this school we applied: two student questionnaires, a teacher questionnaire, classroom observation and a matrix of class activities. In IE 31695, the teacher was on sick leave and the director was in charge of the class. The director answered the questionnaire but chose not to answer as a teacher and provide more information. The students were playing and no teacher was present.]
- c) The access route was poor in some cases. A lack of transport to return to the base from the most isolated areas created difficulties.
- d) The announcement of a regional strike that paralyzed transportation created an impediment. Rumors were that this strike was supported by the Union of Peruvian Education Workers (Sindicato Unitario de Trabajadores en la Educación del Perú, SUTEP). SUTEP delayed application in some cases because they were concerned that the instrument would be used against teachers.

Department of Ancash

ACHIEVEMENTS:

- a) All of the instruments were applied successfully on the days scheduled.¹⁰
- b) The teachers, parents, students and directors of the IE were helpful and fulfilled all aspects of instrument application despite the fact that they had not received authorization from UGEL.
- c) A good relationship was established with interviewers.

⁹ Work began at 8 in the morning and concluded at 6 in the evening with the help of more than one test administrator. This exceeded 9 hours of work and is the equivalent of two workdays between 8 am and 1 pm.

¹⁰ Except for one parent questionnaire for IE 86032; the parents were two hours away in another community at the time.

DIFFICULTIES:

- a) In the majority of cases, authorization from the UGEL never arrived.
- b) The director, who was supposedly doing paperwork at the UGEL, was absent.
- c) Some teachers and directors indicated their discontent, saying they always received visits "but nothing changes."
- d) We had to visit parents in their homes because it was a market day.
- e) We had transport for only two hours. In both cases, we had to walk for two hours due to lack of transport. We were initially concerned about the trip to the IE because of criminal activity in the area. This was a more of a problem in areas located far from urban centers. In some cases, we met up with "ronderos" along the way who protect the community from terrorist activities.
- f) Some teachers were uncomfortable with giving us their first names and surnames.
- g) Interviewers were dressed too formally and in some case women were wearing heels; this was inappropriate for the area.

Department of Lima

ACHIEVEMENTS:

- a) We applied all of the instruments.
- b) All of the teachers participated, particularly those in the 2nd grade, as well as parents, the director of the IE, and students.
- c) There was a good work atmosphere and harmonious relations were established with all interviewers.

DIFFICULTIES:

- a) Despite the fact that the schools were close to the corresponding UGEL, they had not received the authorization regarding the diagnostic.
- b) In some IEs, the teachers were at an external training course, which cut the time available to apply the instruments and limited the process's privacy. [In UGEL Rimas, the teachers were being monitored by the Enrique Guzmán y Valle National University, La Cantuta, as part of teacher training. A university professor was present in many of the classrooms. In IE Señor de los Milagros de Los Olivos, we had to apply the instrument in one day during the morning shift and the afternoon shift because (Friday, 16th) the entire school was to undergo training by the UGEL on the following day.]
- c) Students were absent from the classrooms; teachers indicated that students often miss school because parents are constantly traveling.
- d) The teachers were uncomfortable giving their first names and surnames.
- e) Some IEs showed characteristics that were not foreseen in the work plan. The IE Tomaza Méndez de Bringas is now renamed "Virgen del Carmen," which made it difficult to locate. The director at IE 7241 indicated his discontent because he had yet to receive official textbooks for students. In IE San Francisco de Asís de Manchay, the UGEL had scheduled a reading plan for November 16; as such, the student questionnaires were applied on the 15th.

f) The area assigned for the application of the pre-reading instruments was not private.

3.2 Results of the application time

Table 8 contains the basic statistics regarding the time used in applying the instruments in the IEs in each department.

Table 8. Time of application: Averages and standard deviations (SDs) for fieldwork activities in the IEs, by department

<i>Activities</i>	<i>La Libertad</i>		<i>Ancash</i>		<i>Junín</i>		<i>Lima</i>	
	<i>Average</i>	<i>SD</i>	<i>Average</i>	<i>SD</i>	<i>Average</i>	<i>SD</i>	<i>Average</i>	<i>SD</i>
Introduction of the IE director	21	10	17	8	21	20	18	10
Interview with the director. Instrument 6 (director's questionnaire)	21	6	25	11	37	26	28	11
Interview with teacher in classroom 1. Instrument 7 (teacher questionnaire)	22	4	26	11	24	16	24	11
Interview with students in classroom 1. Instrument 2b (pre-reading application)								
Student 1	21	5	26	6	25	6	23	7
Student 2	21	5	23	6	25	6	22	6
Student 3	22	6	25	7	24	7	21	5
Student 4	21	4	25	8	22	7	22	6
Interview with teacher in classroom 2. Instrument 7 (teacher questionnaire)	19	4	23	11	26	15	22	13
Interview with students in classroom 2. Instrument 2b (pre-reading application)								
Student 1	20	3	24	6	24	7	20	6
Student 2	21	4	22	7	23	7	22	7
Student 3	21	5	23	6	23	9	22	6
Student 4	20	4	23	7	22	9	23	8
Interview with parents. Instrument 4 (parent questionnaire)	16	6	21	12	26	11	19	12

<i>Activities</i>	<i>La Libertad</i>		<i>Ancash</i>		<i>Junín</i>		<i>Lima</i>	
	<i>Average</i>	<i>SD</i>	<i>Average</i>	<i>SD</i>	<i>Average</i>	<i>SD</i>	<i>Average</i>	<i>SD</i>
Approximate time for coordination	32	15	31	15	41	30	32	11
Verification of how information was filled out in the instruments	44	7	50	20	60	32	46	11
Total times*	340	30	385	88	422	100	362	77

* Classroom and IE observations were conducted during the day without specific time limits; as such, they are considered within the total time used.

The application time was greater in the department of Junín. This result can be explained by the fact that more time was used in the interviews (director, teachers, and parents) and because supervisors and teachers had to coordinate beforehand because teachers in this department refused to participate in government activities.

With regard to the application time for instruments involving students, times were higher in the department of Ancash and Junín. These two departments had a higher number of rural IEs (44%). Departments that were less rural—La Libertad (25%) and Lima (13%)—required less time for application.

3.3 SSME results

The results of the responses obtained in the sample for the different questions on the SSME questionnaire can be found in the Excel file, *EstadisticasbasicasSSME.xls sttsvjrd*. Annex 8 contains a list of questions that have no assigned values and several response categories that achieved less than 5% of total responses. In general, these categories can be regrouped or eliminated.

Table 9 shows the magnitude of the changes made to the instruments after the final pilot test.

The instruments with the highest number of observations were the director's questionnaire, teacher questionnaire, student questionnaire, and class activity questionnaire. The discussion at the end of this report comments on modifications introduced into the instruments as a result of these analyses.

Table 9. Modifications made to the instruments following the pilot study

<i>Description</i>	<i>1</i>	<i>2a</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>Total</i>	<i>%</i>
Total number of original questions without modification	9	20	22	29	68	13	46	207	85
Total number of questions with modifications to words	0	9	0	1	0	2	3	15	6
Total number of questions with additional response alternative	0	6	0	0	7	0	2	15	6

<i>Description</i>	<i>1</i>	<i>2a</i>	<i>3</i>	<i>4</i>	<i>5</i>	<i>6</i>	<i>7</i>	<i>Total</i>	<i>%</i>
Total number of questions with modifications to words and additional response alternative	2	0	0	0	0	4	0	6	2
Total number of additional questions	0	0	0	0	0	0	0	0	0
Total number of questions eliminated	0	0	0	0	0	0	0	0	0
Total number of original questions (other than general data)	11	35	22	30	75	19	51	243	100

3.4 EGRA results

Systems for EGRA coding

The EGRA instrument has 7 sections. Table 10 is a summary of the EGRA coding. Please note that there are two basic types of coding. One is done on a correct-incorrect basis and the other is based on a word or letter count.

Table 10. System to code the EGRA by section and format type

<i>Sections</i>	<i>Subsection description</i>		<i>Coding system</i>	
I Orientation to print	1.1	Where to begin to read	0=incorrect	1=correct
	1.2	Where to read next	0=incorrect	1=correct
	1.3	At the end of the line, where to read next	0=incorrect	1=correct
II Letter name knowledge	2.4	Total number of correct letters in the time allotted	Minimum=0	Maximum=100
III Familiar word identification	3.4	Total number of correct words in the time allotted	Minimum=0	Maximum=50
IV Unfamiliar nonword decoding	4.4	Total number of nonsense words correct in the time allotted	Minimum=0	Maximum=50
V Reading of a passage	5.4	Total number of correct words in the passage for the time allotted	Minimum=0	Maximum=69
Vc Comprehension of the passage	5.6	Do they have a cat?	0=incorrect	1=correct
	5.7	What does it like to do?	0=incorrect	1=correct
	5.8	Is it skinny or fat?	0=incorrect	1=correct
	5.9	Where did the cat hide?	0=incorrect	1=correct

<i>Sections</i>	<i>Subsection description</i>		<i>Coding system</i>	
	5.10	What is Maria's mom going to have?	0=incorrect	1=correct
VI Listening comprehension	6.1	Where does Maria stay on Saturday?	0=incorrect	1=correct
	6.2	What does Maria's dad do?	0=incorrect	1=correct
	6.3	What is the name of Maria's friend?	0=incorrect	1=correct
VII Dictation	7.1	Spelled "mercado" (market) correctly	0=incorrect	1=correct
	7.2	Spelled "Va" (He goes) correctly	0=incorrect	1=correct
	7.3	Spelled "té" (tea) correctly	0=incorrect	1=correct
	7.4	Used appropriate spaces and the correct text direction (left to right)	0=incorrect	1=correct
	7.5	Used capital letters correctly	0=incorrect	1=correct
	7.6	Used punctuation correctly	0=incorrect	1=correct

Results of the EGRA sections

Table 11 describes some of the characteristics of the responses to the section items, as well as giving a reliability assessment and the results of the normality test.

In general, the sections with greatest confidence are Sections I, II, and VII. Sections III, IV, V, Vc, and VI did not reach acceptable reliability levels.¹¹ Additionally, none of the sections shows a normal distribution of scores.

The results also indicate that a lower incidence of valid responses was reached in section Vc and in section VII, which correspond to comprehension of the passage and dictation.

Table 11. Descriptive statistics, reliability, and the EGRA normality test

<i>EGRA sections</i>	<i>Characteristics</i>		<i>Descriptive statistics of responses</i>				<i>Reliability</i>	<i>Normality Test</i>	
	<i>Items</i>	<i>Valid responses</i>	<i>Mean</i>	<i>SD</i>	<i>Variability coefficient</i>	<i>Range</i>	<i>Cronbach's alpha</i>	<i>Z-test</i>	<i>p value</i>
I	3	512	2.2	1.1	52.1	0-3	0.7947	8.202	0
II	100	512	31.2	16.7	53.6	0-99	0.6104	2.611	0
III	50	509	39.5	13.7	34.7	0-50	0.2505	4.982	0

¹¹ An extensive analysis of the reliability of these sections can be found in the annexes.

EGRA sections	Characteristics		Descriptive statistics of responses				Reliability	Normality Test	
	Items	Valid responses	Mean	SD	Variability coefficient	Range	Cronbach's alpha	Z-test	p value
IV	50	507	31.0	12.7	40.9	0-50	0.3186	1.507	0
V	69	502	57.0	18.0	31.6	0-69	0.4123	6.085	0
Vc	5	464	4.4	1.0	22.6	0-5	0.5612	7.432	0
VI	3	509	2.0	0.9	44.3	0-3	0.3295	5.235	0
VII	6	493	3.9	1.6	40.2	0-6	0.6386	3.97	0

The matrix of correlations in Table 12 shows that Sections II to VII demonstrate significant correlations and that section I is significantly correlated with section VI¹² but not with the other sections.

In general, this confirms that the different EGRA sections measure a common aspect with the exception of the orientation-to-print section.

Table 12. Matrix of correlations of EGRA sections

		<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>Vc</i>	<i>VI</i>	<i>VII</i>
<i>I</i>	r	1	-0.083	-0.004	-0.032	0.002	0.066	0.14	-0.072
	Sig.	—	0.06	0.93	0.478	0.96	0.159	0.001**	0.109
	N	512	512	509	507	502	464	509	493
<i>II</i>	r	-0.083	1	0.458	0.542	0.399	0.27	0.172	0.373
	Sig.	0.06	—	0**	0**	0**	0**	0**	0**
	N	512	512	509	507	502	464	509	493
<i>III</i>	r	-0.004	0.458	1	0.818	0.891	0.503	0.232	0.598
	Sig.	0.93	0**	—	0**	0**	0**	0**	0**
	N	509	509	509	507	502	463	507	491
<i>IV</i>	r	-0.032	0.542	0.818	1	0.773	0.424	0.185	0.528
	Sig.	0.478	0**	0**	—	0**	0**	0**	0**
	N	507	507	507	507	502	463	505	489
<i>V</i>	r	0.002	0.399	0.891	0.773	1	0.491	0.204	0.553
	Sig.	0.96	0**	0**	0**	—	0**	0**	0**
	N	502	502	502	502	502	460	500	484

¹² This correlation between orientation to print and listening comprehension can be considered spurious.

		<i>I</i>	<i>II</i>	<i>III</i>	<i>IV</i>	<i>V</i>	<i>Vc</i>	<i>VI</i>	<i>VII</i>
Vc	r	0.066	0.27	0.503	0.424	0.491	1	0.219	0.345
	Sig.	0.159	0**	0**	0**	0**	—	0**	0**
	N	464	464	463	463	460	464	462	447
VI	r	0.14	0.172	0.232	0.185	0.204	0.219	1	0.113
	Sig.	0.001**	0**	0**	0**	0**	0**	—	0.012*
	N	509	509	507	505	500	462	509	491
VII	r	-0.072	0.373	0.598	0.528	0.553	0.345	0.113	1
	Sig.	0.109	0**	0**	0**	0**	0**	0.012	—
	N	493	493	491	489	484	447	491	493

* : Correlation significant at 1%

** : Correlation significant at 5%

Exploring the global EGRA score

To explore the possibility of reporting a single score for the EGRA by considering all the sections of the EGRA, we conducted an analysis of items that is described in Table 13.

The analysis was conducted by separating by qualification type into correct-incorrect and a count of letters or words. In this report, the items in the first group are called EGRA dichotomic and those in the second group are called EGRA count.

EGRA DICHOTOMIC AND EGRA COUNT

The results indicate that even though an acceptable confidence level was obtained with EGRA, it is more convenient to separate EGRA into EGRA count and EGRA dichotomic.

Please note that items from section I and Section 6.2 demonstrate better psychometric behavior (greater validity based on the correlation item vs. the remaining items) in the EGRA dichotomic than in the complete EGRA.

Table 13. Analysis of the EGRA items

Section	<i>All the EGRA items</i>		<i>EGRA count</i>		<i>EGRA dichotomic</i>	
	<i>Correlation item vs. remaining items</i>	<i>Alpha if the item is eliminated</i>	<i>Correlation item vs. remaining items</i>	<i>Alpha if the item is eliminated</i>	<i>Correlation item vs. remaining items</i>	<i>Alpha if the item is eliminated</i>
1.1	-0.04	0.68			0.17	0.62
1.2	-0.02	0.68			0.23	0.61
1.3	0.03	0.68			0.24	0.61
2.4	0.43	0.66	0.48	0.92		

Section	All the EGRA items		EGRA count		EGRA dichotomic	
	Correlation item vs. remaining items	Alpha if the item is eliminated	Correlation item vs. remaining items	Alpha if the item is eliminated	Correlation item vs. remaining items	Alpha if the item is eliminated
3.4	0.81	0.55	0.85	0.77		
4.4	0.78	0.56	0.83	0.79		
5.4	0.74	0.56	0.78	0.79		
5.6	0.18	0.68			0.27	0.61
5.7	0.26	0.67			0.23	0.61
5.8	0.16	0.68			0.26	0.61
5.9	0.34	0.67			0.22	0.61
5.10	0.41	0.67			0.38	0.59
6.1	0.13	0.68			0.12	0.62
6.2	0.08	0.68			0.16	0.62
6.3	0.16	0.68			0.23	0.61
7.1	0.43	0.67			0.40	0.58
7.2	0.21	0.67			0.18	0.62
7.3	0.38	0.67			0.35	0.59
7.4	0.42	0.67			0.40	0.59
7.5	0.31	0.67			0.18	0.62
7.6	0.19	0.67			0.12	0.63
Cronbach's alpha		0.67		0.86		0.62

In the dichotomic EGRA, we can see that despite the fact that the items in Section 6.1 and Section 7.6 show less correlation with the rest of the items in this scale, their elimination does not translate into a substantive gain in the scale's reliability.

Formulating an EGRA for count and speed

To calculate the EGRA score for the count, we must conduct a more exhaustive analysis of section II-letters, section III-words, section IV-nonsense words, and section V-reading of a passage, regarding the count times for the elements, the total elements read, the number of elements read incorrectly, and the number of elements read correctly.

READING TIMES

Reading time is the total time elapsed when the stopwatch has stopped. Table 14a describes some statistics and correlations between sections.

Table 14a. Descriptive statistics and correlation matrix for reading times

<i>Sec</i>	<i>Mean</i>	<i>Median</i>	<i>Mode</i>	<i>SD</i>	<i>Variability coeff.</i>	<i>Min</i>	<i>Max</i>	<i>III</i>	<i>IV</i>	<i>V</i>
II letters	59.9	60	60	1.0	1.7	40	60	-0.035	-0.007	0.045
III words	55.4	60	60	7.6	13.7	6	60	1	0.421**	0.682**
IV nonsense	59.3	60	60	2.8	4.7	35	60	0.421**	1	0.376**
V passage	53.0	60	60	9.2	17.4	25	60	0.682**	0.376**	1

* Correlation significant at 1%

** Correlation significant at 5%

In relation to the passage reading times, we observe a lower average in section V involving passage reading, and the variability is significant.

The time used to read letters (section II) is not related to the time used in reading words (section III), but the times for reading known words (section III), nonsense words (section IV) and words within a passage (section V) are correlated.

TOTAL ELEMENTS READ

Table 14b presents statistics regarding the total elements read in the allotted time.

Table 14b. Descriptive statistics and correlation matrix for elements read

<i>Sec</i>	<i>Mean</i>	<i>Median</i>	<i>Mode</i>	<i>SD</i>	<i>Variability coeff.</i>	<i>Min</i>	<i>Max</i>	<i>III</i>	<i>IV</i>	<i>V</i>
II letters	35.9	33	34	16.3	45.4	1	100	0.454**	0.552**	0.388**
III words	40.1	47	50	13.0	32.4	0	50	1	0.823**	0.881**
IV nonsense	32.4	32	50	12.2	37.7	0	50	0.823**	1	0.755**
V passage	56.6	69	69	18.8	33.2	0	69	0.881**	0.755**	1

* Correlation significant at 1%

** Correlation significant at 5%

Regarding total elements read in the allotted time, we observe that the lowest average occurred in the case of nonsense words (IV).

The number of letters read (II) is associated with the number of words read (III), but known words (III), nonsense words (IV), and words in a passage (V) show stronger relationships to each other.

ELEMENTS READ INCORRECTLY

Table 14c presents statistics on the total elements read *incorrectly* in the allotted time.

Table 14c. Descriptive statistics and correlation matrix for elements read incorrectly

<i>Sec</i>	<i>Median</i>	<i>Mode</i>	<i>SD</i>	<i>Variability coeff.</i>	<i>Min</i>	<i>Max</i>	<i>III</i>	<i>IV</i>	<i>V</i>
II letters	4.7	4	3	3.0	63.8	20	0.348**	0.382**	0.277**
III words	0.9	0	0	1.7	188.9	19	1	0.714**	0.637**
IV nonsense	1.7	1	0	2.4	141.2	34	0.714**	1	0.563**
V passage	0.7	0	0	1.4	200.0	13	0.637**	0.563**	1

* Correlation significant at 1%

** Correlation significant at 5%

In the correlations for total elements read incorrectly in the allotted time, the average for reading letters (II) was the highest. It is interesting that reading of nonsense words (IV) resulted in reporting of more errors than did reading of familiar simple words (III).

As mentioned previously, the number of letters read incorrectly (II) is related to the number of words read incorrectly (III), but the reading times of known words (III), nonsense words (IV), and words within a passage (V) show stronger relationships to one another.

ELEMENTS READ CORRECTLY

Statistics for total elements read correctly in the allotted time are shown in Table 14d.

Table 14d. Descriptive statistics and correlation matrix for total elements read correctly

<i>Sec</i>	<i>Median</i>	<i>Mode</i>	<i>SD</i>	<i>Variability coeff.</i>	<i>Min</i>	<i>Max</i>	<i>III</i>	<i>IV</i>	<i>V</i>
II letters	31.2	29	31	16.7	53.5	99	0.458**	0.542**	0.399**
III words	39.5	46	50	13.7	34.7	50	1	0.818**	0.891**
IV nonsense	31.0	31	50	12.7	41.0	50	0.818**	1	0.773**
V passage	57.0	68	69	18.0	31.6	69	0.891**	0.773**	1

* Correlation significant at 1%

** Correlation significant at 5%

Total elements read correctly in the time allotted shows a higher average in the section for reading words in a passage (V).

Likewise, the number of letters read correctly (II) is associated with the number of words read correctly (III), but known words (III), nonsense words (IV), and words within a passage (V) are more closely correlated with each other.

The number of words read correctly in the section on simple words (III) was higher than the number of nonsense words read (IV).

ELEMENTS READ PER MINUTE

Examining Tables 14b and 14c regarding letters and words read, and taking into account the time allowed of 60 seconds (1 minute), we can describe these results per minute. Thus, we can say that:

According to Tables 14b and 14c, if we look at the average of letters read per minute, it is approximately 36 letters, of which the average number of letters read incorrectly per minute is 5 letters.

We can see from the same tables that the average of simple words read per minute is around 40 words with an average of 1 word read incorrectly.

It is also observable that the average of simple words read per minute is around 32 with an average of 2 words read incorrectly.

NUMBER OF ELEMENTS READ AND PERCENTAGE OF SUCCESS

In this subsection, we study the effect of the number of elements to be read on the success rate.

To evaluate the longitudinal effect of the elements presented in the percentage of success in the reading of elements, we prepared a dispersion diagram (Figure 5) of these two variables for the sections on letters, familiar words, nonsense words, and words within a passage of text.

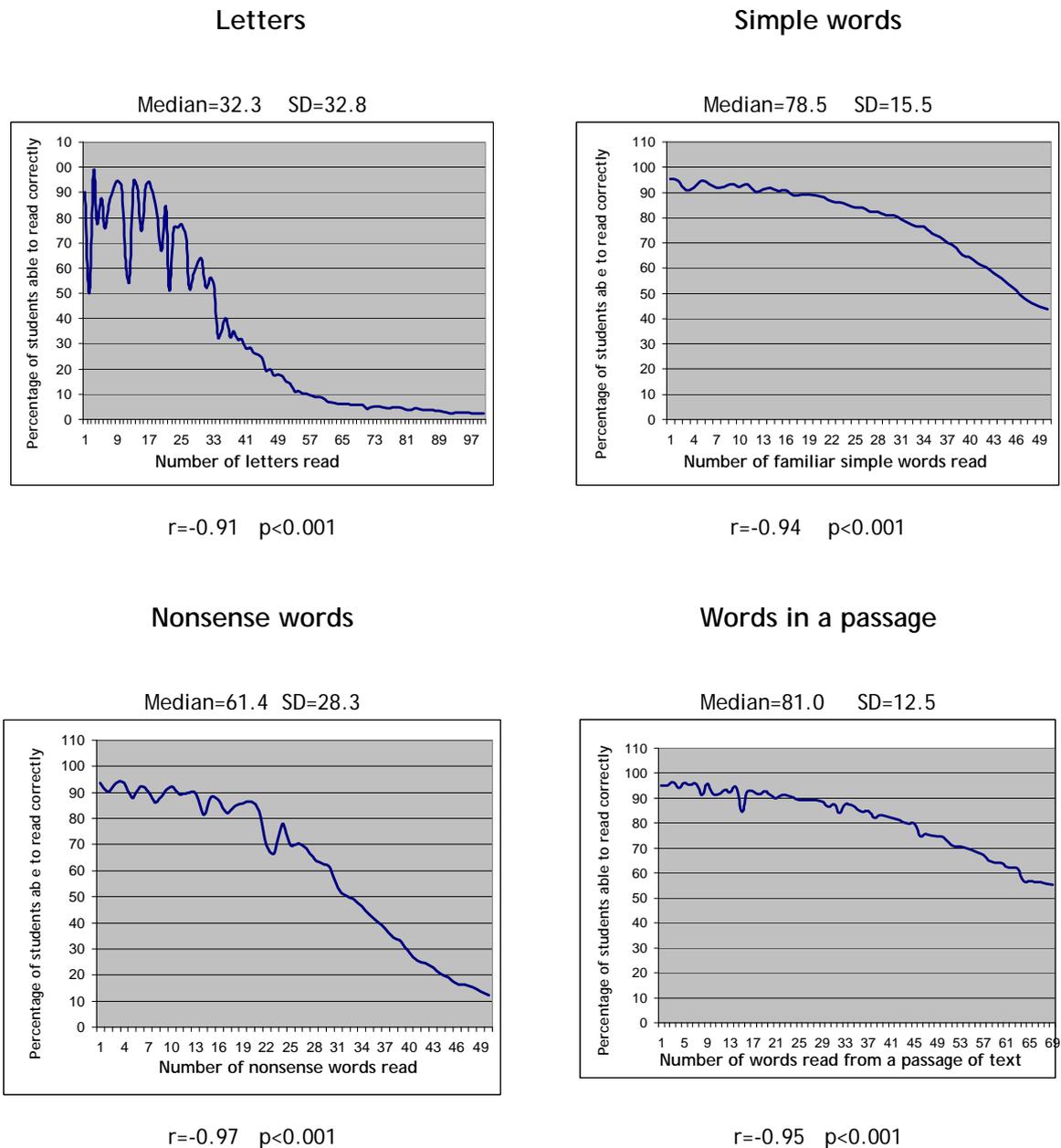
Regarding letter name knowledge, we observe that up to letter 33, more than 50% of the students know the name of the letters. In general, we observe a decreasing rate of knowledge of the letters.

For knowledge of simple words, we see that up to word 46, more than 50% of the students can read this type of word. In general, we observe a decreasing rate of knowledge of simple words.

For decoding of simple nonsense words, we can see that up to word 31, more than 50% of the students can read this type of word. In general, we observe a decreasing rate of reading for nonsense words.

In relation to reading a passage, we observe that all of the words were read by at least 50% of the students. We also observe a decreasing rate of words read.

Figure 5. Effect of the number of elements read in the EGRA



If we analyze the correlation coefficients in Figure 5, we can conclude that the number of letters read or words read increases and the reading rate decreases. Additionally, we can see that the average of the percentage of success for common words is higher than is the case for nonsense words but lower than for words in a passage. The average percentage of success for letters is very low due the number of letters read and the low percentage of students able to read said letters.

EXPLORING THE POSSIBILITY OF REDUCING THE COUNT SECTIONS

In the previous subsection, we identified the effect of the number of elements to be read on the percentage of success. To establish if it is possible to reduce the elements in the EGRA section based on counts, we conducted a reliability test for a number of elements less than that incorporated into EGRA. The objective was to determine if a version with a lower number of elements, still in the order in which they appear in their respective sections, would have a reliability level that was either the same as or higher than that obtained previously. This is shown in Table 15.

Table 15. Cronbach's alpha reliability for a different number of elements read, according to the order in which they appear

<i>II: Knowledge of letters</i>		<i>III: Common words</i>		<i>IV: Nonsense words</i>		<i>V: Words from a passage</i>	
<i>Letters</i>	<i>Alpha</i>	<i>Words</i>	<i>Alpha</i>	<i>Words</i>	<i>Alpha</i>	<i>Words</i>	<i>Alpha</i>
33	0.63	25	0.41	25	0.73	25	0.54
40	0.71	30	0.44	30	0.78	30	0.53
45	0.68	35	0.42	35	0.87	35	0.38
50	0.72	40	0.40	40	0.55	40	0.39
55	0.70	45	0.44	45	0.46	45	0.41
60	0.66	50	0.25	50	0.32	50	0.42
70	0.64					55	0.42
80	0.65					60	0.41
90	0.67					65	0.40
100	0.61					69	0.41

The results of the reliability test for knowledge of letters indicate that it is possible to have versions including a lower number of letters that maintain and perhaps even improve the reliability of this section of the EGRA. An acceptable number could be 55 letters, given that reliability decreases for numbers lower or greater than 55.

In terms of common words, higher reliability values are found for a lower number of words in this section. As such, we suggest that the section contain around 40 words.

With regard to nonsense words, we obtained reliability values that were higher for a lower number of words. As such, we suggest that the section contain around 35 words given that after this number reliability for this section decreases.

In terms of words in a text passage, we obtained the same reliability values for a lower number of words. As such, we suggest using around 55 words, given that with this number, we can obtain a slight improvement in reliability.

Despite the fact that reliability also depends on the type of elements presented, the analysis conducted suggested that, in general, it is possible to cut down the count section of the EGRA to a lower number of elements to be read without causing a reduction in reliability.

Given that there is not a high percentage of “did not answer” or “did not read” responses, it *is* possible to establish a global score for EGRA dichotomic by totaling the corresponding scores from the dichotomic items in EGRA.

ANALYZING THE COUNT SECTIONS ACCORDING TO SPEED

The values obtained in the EGRA count can be transformed into speed values through the following expression:

$$\text{Speed value} = \text{Elements} / \text{Time} ,$$

where

Elements = Number of elements read in the allotted time, from 0 to the number of elements considered

Time = time in seconds used to read the number of allotted elements that is greater than 0 if the student read an element in a 60-second time period.

Table 16 shows descriptive statistics, an analysis of items, and a matrix of correlations between these speeds.

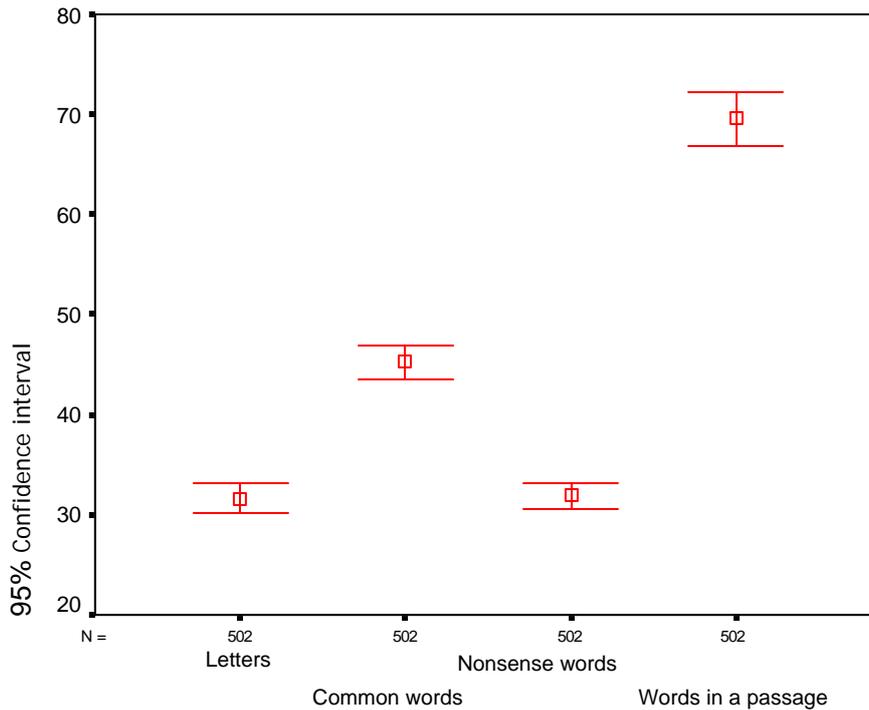
Table 16. Descriptive statistics, item analysis, and matrix of correlations for reading speed

Sec	Descriptive statistics					Item analyzed		Matrix of correlations		
	N	Mean	SD	Variab. coefficient	Range	Correlation item vs. rest of items	Alpha if item is eliminated	III	IV	V
II	512	31.3	16.8	53.7	0-104.2	0.47	0.87	0.455**	0.546**	0.401**
III	509	44.7	19.9	44.5	0-107.1	0.85	0.72	1	0.82**	0.841**
IV	507	31.7	14.0	44.2	0-85.7	0.86	0.77	0.82**	1	0.796**
V	502	69.6	31.4	45.1	0-165.0	0.80	0.80	0.841**	0.796**	1
							alpha = 0.84			

* Correlation significant at 1%

** Correlation significant at 5%

As can be discerned from Table 16 and the diagram below, the results indicate that the highest speed obtained is in the section to read words in a text passage (70 words per minute), followed by reading common words (45 words per minute). The lowest reading speeds correspond to reading of letters (31 letters per minute) and the reading of nonsense words (31 words per minute). We can see that the variability in speed is great, particularly with regard to the speed for reading letters.



CONFIDENCE INTERVALS AT 95%

We also find that it is possible to combine the different reading speeds into a single scale, considering that the reading speed for these different elements is very high, particularly the speeds for reading common words, nonsense words, and words within a text passage.

We propose three indexes for a combined EGRA speed value:

- a) Average of speed values
- b) Average of standardized speed values
- c) Factor obtained from an analysis of main components

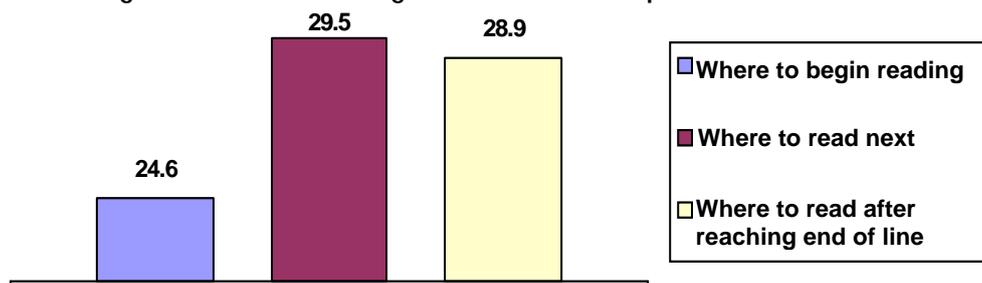
Putting together an EGRA dichotomic

Prior to conducting this type of analysis, it is important to review the number of incorrect, no-response, and not-read results obtained.

SECTION I. ORIENTATION TO PRINT

In this section, no “no response” answers were obtained. In Figure 6, we can see that 25% of the students had problems indicating where reading should begin, 30% had difficulties indicating where reading should continue, and 29% had trouble pointing to where to read after the line had been read.

Figure 6. Percentage of incorrect responses in section I



Alternatively, regarding orientation to print, we can regroup the sample into those that had no problems with this part (group 4) and three groups that had problems (groups 1, 2, and 3). The following table defines the four cases that constitute these groups.

Types of groups represented:

1. Problems beginning, continuing, and changing lines
2. Problems beginning but no problems with continuing or changing lines
3. No problems beginning but problems in continuing or changing lines
4. None of these problems

<i>Group</i>	<i>Case</i>	<i>Frequency</i>	<i>Percentage</i>
1	000	73	14.3
2	001	22	4.3
	010	9	1.8
	011	22	4.3
3	100	40	7.8
	101	16	3.1
	110	26	5.1
4	111	304	59.4
	Total	512	100.0

According to this classification, around 60% had no problems in dealing with whole words, given that they could successfully begin reading, continue, and change lines. Around 15% could not manage to meet any of the requirements for this section.

On the other hand, 10% had problems beginning but were able to continue with some success, and 15% had no problems beginning but then encountered difficulties moving forward.

SECTION V. PASSAGE READING AND COMPREHENSION

In Figure 7, we can see the percentages of students that responded incorrectly to the reading. We see that the highest percentage is for the last question for the last paragraph read. With regard to the questions not read, we find a higher percentage for the last question in this section (see Figure 8).

Figure 7. Percentage of incorrect responses in section V

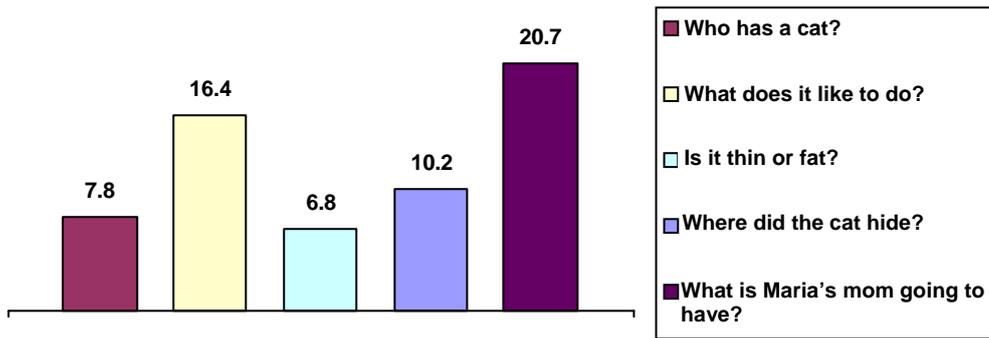
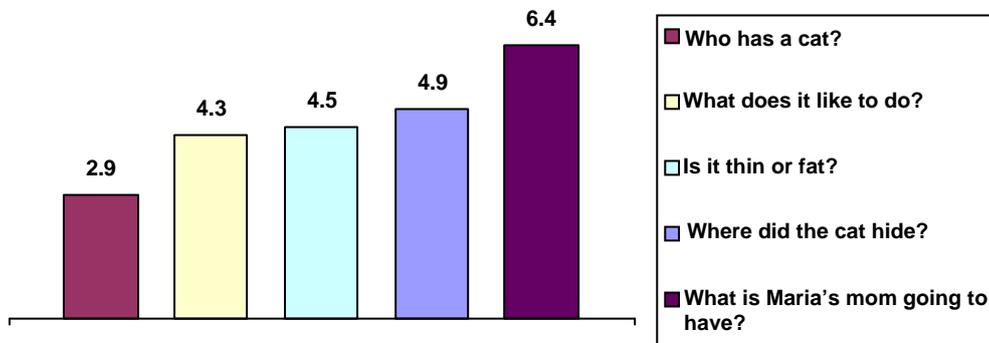


Figure 8. Percentage of nonresponses (unable to read) in section V



SECTION VI. LISTENING COMPREHENSION

In Figure 9, we can see the percentages of students that responded incorrectly to the listening comprehension component. We can see that the highest percentage is found in the last question, which corresponds to the last paragraph read. In Figure 10, we find that the percentages for "no response" and "not read" are low.

Figure 9. Percentages of incorrect responses in section VI, listening comprehension

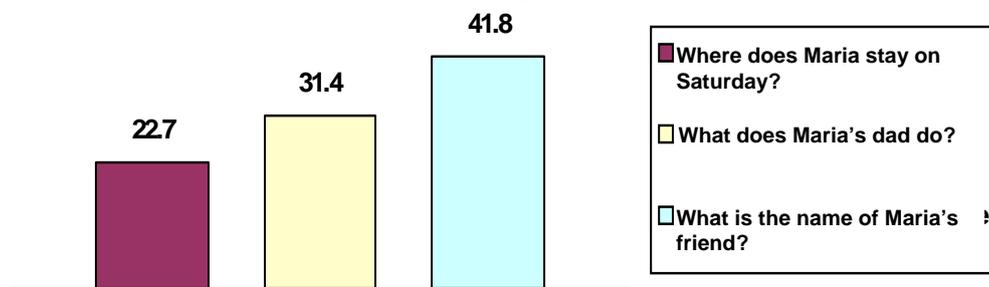
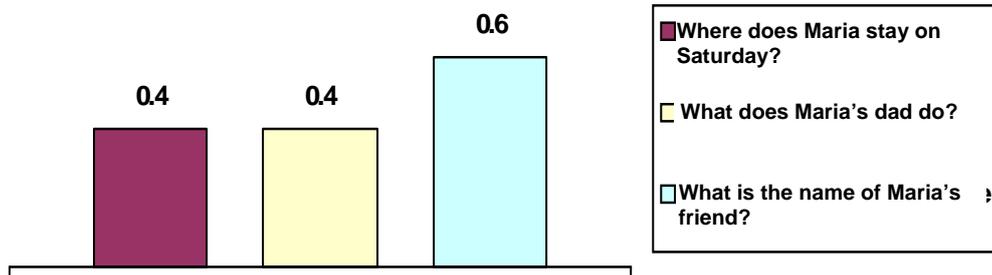


Figure 10. Percentages of non-response (not read) in section VI



SECTION VII. DICTATION

Figure 11 shows the percentages of students that fall into different categories with regard to the dictation section. We observe high percentages in the category for “none.” In Figure 12, we find that the percentages for “no response” and “not read” are low.

Figure 11. Percentages of incorrect responses for questions in section VII

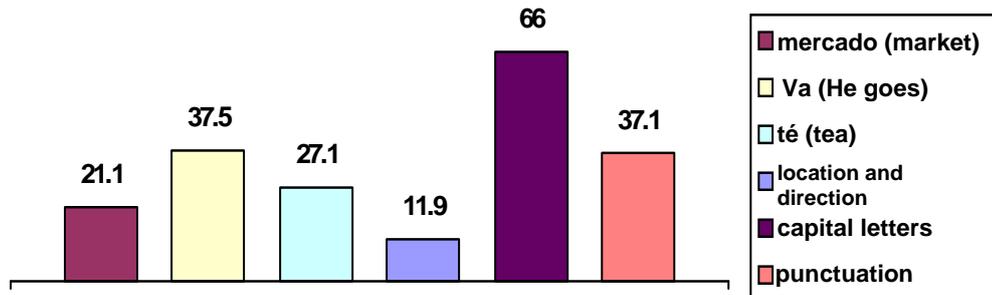
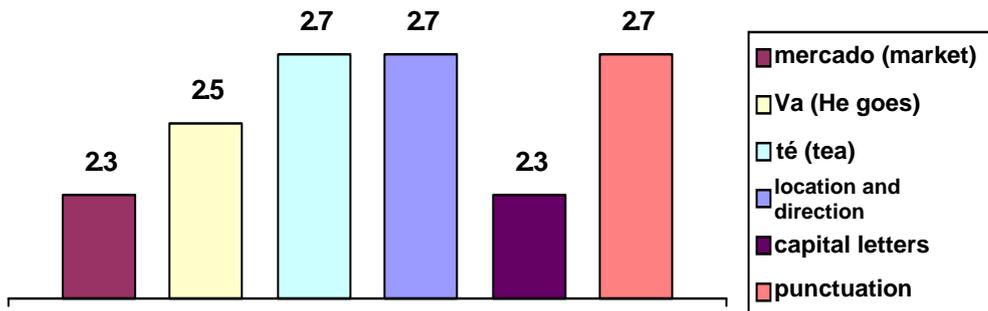


Figure 12. Percentage of nonresponse (not read) in section VII



Given that the percentages for “no response” and “not read” are not high, it is possible to calculate a global score for the EGRA dichotomic by totaling the scores corresponding to the dichotomic items in the EGRA.

NOTE ON MODIFICATIONS INTRODUCED IN SECTIONS V AND VII, INSTRUMENT 2B, PRE-READING

It is important to clarify the text changes and scoring criteria that were incorporated into the reading and dictation sections.

Section V, Passage reading and comprehension

Change in the reading text was from “debajo de la casa” (under the house) to “debajo de la cama” (under the bed).

Section VII, Dictation

The questions were re-edited to:

- 7.1 Spelled “mercado” (market) correctly.
- 7.2 Spelled “Va” (He goes) correctly.
- 7.3 Spelled “té” (tea) correctly.
- 7.4 Used correct spaces and text direction (left to right).
- 7.5 Used capital letters correctly.
- 7.6 Used punctuation correctly.

The options to score for these questions include:

- 0 None
- 1 Some correct
- 2 All correct

Given that the extremes represent clear scores and in practice option 1 was subject to doubt, given that it was confused with 0, the recommendation for this section is that response options be pared down to two elements: correct (options 1 and 2) and incorrect (option 0). In practice, this can be done through the database post-pilot.

3.5 SSME and EGRA relationships

In order to demonstrate the explanatory nature of EGRA and SSME, we show comparisons of reading speed for words in a text passage according to some of the students’ characteristics such as level of instruction, sex, and geographic point of origin, and whether the student attended preschool or kindergarten.

In Table 17, the mean comparison results using an F-test indicate that there are significant differences for boys in third grade, girls, and students who had training prior to primary school. By contrast, zone of residence—urban or rural—made no such difference in speed. In general, it can be said that the explanatory effect of these variables coincides with those reported in the literature.

Table 17. Comparison of performance in reading speed for a text passage according to some student characteristics

<i>Characteristic</i>	<i>Category</i>	<i>N</i>	<i>Median</i>	<i>SD</i>	<i>F</i>	<i>p value</i>
Grade	2nd grade	354	64.6425	31.0913	31.233	0**
	3rd grade	148	81.3156	28.9535		
Sex	Male	250	65.4287	28.0585	8.755	0**
	Female	252	73.6547	33.9294		
Area	Urban	345	69.4686	31.4144	0.009	0.92
	Rural	157	69.7547	31.4183		
Attended prekindergarten, nursery school, kindergarten	No	102	64.0309	26.7579	4.117	0.04*
	Yes	396	71.0786	32.3372		
	Total	498	69.6351	31.3798		

*: Differences significant at 1%

** : Differences significant at 5%

4. DISCUSSION

4.1 Fieldwork results

According to the fieldwork results, the following aspects are positive and others require modification for future applications.

Positive aspects

- All of the instruments were applied without difficulty according to the established schedule. We obtained information from:
 - 128 classroom observations
 - 512 students
 - 128 matrices of classroom activities
 - 126 parents¹³
 - 64 directors
 - 64 observations of IEs
 - 127 teachers¹⁴
- The applications were conducted over two days as foreseen, but in 8 IEs, it was applied in one day by using the morning and afternoon shifts.

¹³ In 2 IEs, the parents were not available because the assessment date coincided with market day.

¹⁴ In one IE, the director was substituting for a teacher who was on sick leave, and the director would not answer the teacher questionnaire on her behalf.

- With few exceptions, IE directors were cooperative; this also applied to teachers and parents as well as students in general; all of the children wanted to be interviewed.
- The group of interviewers worked together well to achieve the goal indicated.
- The environment for the interview was appropriate and students were motivated by the gift of a pencil case.

Aspects that need to be improved in future applications

- The authorizations corresponding to instrument applications had not arrived at each of the IE addresses; as such, the directors were unfamiliar with the Project. This translated into a delay in the application process. This situation was exacerbated if the director was absent with the excuse that he was doing paperwork at the UGEL.
- Some IEs showed characteristics that were not foreseen in the work plan such as IEs that were not co-ed, IEs located at other addresses due to construction, IEs with names different from those on the official list, IEs working out of two different locales.
- Some IEs did not complete the application due to special situations in which the teacher may have opposed the process or was out on leave. In all these cases, the absences were covered by completing questionnaires at other IEs.
- Accessing the IEs is difficult, particularly for rural and peripheral IEs. The access routes, lack of transport to get back from the most distant areas, the need to walk up to two hours, crime, and community watches due to terrorism (“ronderos”) are some of the difficulties we found.
- We identified some difficulties in the responses to some of the student questionnaires and in the pre-reading tests: There were difficulties in class observation because the classrooms were shared with secondary students, there were no education materials available, or there was a lack of privacy in the area assigned for the application of the pre-reading instrument.
- The announcement of a regional teacher strike generated some initial difficulties at the beginning of the application process.
- Difficulties arose from the school dynamics such as the discontent of some teachers and directors because they are always visited but “everything remains the same,” individuals uncomfortable with giving their first name and last names, and student absenteeism.

- We had to visit parents at home because they did not always come to the IEs; this may be a problem if interviewers do not have the appropriate attire for walks of this kind.
- In some IEs, the teachers were in an external training process that limited the application time and the privacy of the process because other people were present in the classrooms.

On the whole, we believe that the fieldwork was completed successfully, meeting the goals foreseen. It is important to point out that the experience of the network of interviewers and supervisors helped to resolve some of the initial difficulties during the fieldwork: a) lack of information from the directors regarding the evaluations, b) maps and information on the IEs that did not coincide with the observations, and c) external difficulties, particularly due to the threat of a teacher strike and teacher training. All these aspects can be improved for future instrument application.

4.2 Regarding the application time for instruments

The study times in the pilot constitute an important source of information to make future decisions regarding instruments if they must be applied in only one day. The effective working day for evaluation at IEs is around 4½ hours, but according to the results found, the instruments require an average application time of 6 hours, which may be even higher if the coordination tasks identified in the fieldwork have not been done.

There are two possibilities for work in future applications: apply the instruments over two workdays or apply instruments in only one day by cutting down the number of instruments or components to be applied in the time available. Up to 25% of the total number of questions can be cut down. Further analysis must be conducted to identify at which point the questions can be cut down in specific questionnaires. A suggestion is to remove several of the control questions used in the questionnaires, reduce the coordination time, and conduct verification after the application.

The time used by interviewers to evaluate 8 students in an IE is 180 minutes on average given that the individual time was 23 minutes. To reduce these times, it is important to consider simultaneous applications or to use reduced versions of the instruments.¹⁵

4.3 The SSME results

This instrument constitutes an efficient tool to understand school dynamics. After the pilot application, the FDA technical teams thought that it would be a good idea to make two kinds of modifications to the instruments: general and specific.

¹⁵ See the section “Formulating an EGRA for count and speed” above on possible changes or reductions in the EGRA.

General modifications

After evaluating the results of the pre-pilot questionnaire conducted in October 2007, we redesigned the instrument format to facilitate the interviewers' job of gathering information for the pilot questionnaire. The objective was to provide more fluidity to the coding work and information-gathering process.

We made modifications to the general identification questions in the instruments given that the first versions of the initial questions did not allow for an adequate identification of aspects such as the questionnaire, application location, the education institution, grade of applications, etc. We have added specific identification items such as: data for the questionnaire (questionnaire code, name and interviewer code), general data (department code, province, district and area), data on the education institution (UGEL code, name, and school identification code for the education institution), classroom data (grade code, classroom code and name of teacher), all of which may vary depending on the nature of the instrument.

Specific modifications

Table 9, duplicated here from an earlier section, shows the magnitude of the changes made to the instruments after the final pilot test.

Table 9. Modifications made to the instruments following the pilot study

<i>Description</i>	1	2a	3	4	5	6	7	Total	%
Total number of original questions without modification	9	20	22	29	68	13	46	207	85
Total number of questions with modifications to words	0	9	0	1	0	2	3	15	6
Total number of questions with additional response alternative	0	6	0	0	7	0	2	15	6
Total number of questions with modifications to words and additional response alternative	2	0	0	0	0	4	0	6	2
Total number of additional questions	0	0	0	0	0	0	0	0	0
Total number of questions eliminated	0	0	0	0	0	0	0	0	0
Total number of original questions (other than general data)	11	35	22	30	75	19	51	243	100

Of the total number of questions present in the instruments, 85% maintained their original form and 15% were modified. The instruments that had the largest number of modifications were: the student (2a) and the education institution (7); the instruments that underwent the lowest number of changes were the matrices for class activities (3) and pre-reading (2b). The types of modifications made included changes in words (6%),

additional response alternatives (6%), and both (2%). Questions were not added or eliminated as a result of the pilot study.

Next, we provide some details on the changes introduced following the analysis of the pilot study. The changes suggested in questions (and their codes, also given here), refer to the pilot instruments (see Annex 3).

INSTRUMENT 1: CLASSROOM OBSERVATION

Name of teacher (Quest. CO8)

Due to the political atmosphere in the education sector in Peru, we recommend that the requirement to include the teacher name be eliminated so that the interviews will be anonymous.

Name of the subject studied (Quest. CO12 and CO13).

In CO12, we changed the term "Comunicaciones" (Communications) to "Comunicación Integral" (Integrated Communications), given that this is the official name of the course.

In CO12 and CO13, we added an option for comments about textbook availability. After reading the responses, we decided to take the comment section out of the questionnaire (12.01 in the pre-pilot) given that it was not clear if the presence or the absence of a book was because the student did not have the book, and if he or she had the class, or other combinations.

INSTRUMENT 2A: STUDENT QUESTIONNAIRE

Status of the questionnaire (S2)

This question, which is found in the general data section, should be excluded because an incomplete questionnaire is replaced by another.

Language spoken at home (S14)

In some IEs, we found a high incidence of "other" languages. The explanation is that in some cases, this option was used to indicate Spanish when the student meant to say that they speak only Spanish at home. As such, it would be best to write "Only Spanish."

Exercise book (S16)

"Exercise book" was translated as "Cuadernos de ejercicios."

The option "más de uno" (more than one) could be confused with "more than one page" and as such may conflict with other options. In the Options section, we propose changing the option "más de uno" to "more than one exercise book." We also recommend adding the option "Menos de la cuarta parte," or "less than a fourth."

Teacher comments and corrections (S17)

This question must be distinguished from the following question, S18. We propose changing to "Anote cuántas páginas tienen señales no textuales del docente en forma de

anotaciones, correcciones, señales de visto bueno" (Write down how many pages have nontextual signs from the teacher in the form of comments, corrections, indications of approval).

Textual comments (S18)

We propose changing to "Anote cuántas páginas tienen comentarios textuales del docente" (Indicate how many pages have textual comments from the teacher).

Extra classes (S30)

We propose changing to "En este año, alguna vez tu profesor te dió clases extras" (This year, your teacher taught extra classes [Exclude makeup classes for one student or all students]).

Reasons for absence (S35)

We propose changing the option "día del mercado" (market day) to "Ayudar en trabajo en casa" (helping at home).

Initial level of instruction (S38)

We propose "inicial" (preschool) instead of "kinder" (kindergarten; same with the director's questionnaire).

Existence of a bathroom (S40.04 y S40.05)

The option "baño" (bathroom) was ambiguous in some cases. We propose changing "baño" to "retrete" (latrine).

Question S40.10

We propose changing "furgoneta" (delivery van) to "moto" (motorcycle).

INSTRUMENT 3: MATRIX OF CLASS ACTIVITIES

In this instrument, we added three additional classroom data points (grade, teacher name, and classroom code) in the general section. We recommend emphasizing the time registered for the three observations to ensure valid comparisons with other schools.

INSTRUMENT 4: PARENT QUESTIONNAIRE

Grade (P9)

The responses to this question can be arbitrary when the general instructions indicate that parents from other grades can be included. In future evaluations, the parents must correspond with the students evaluated, connecting as such the grade and the teacher.

Parent meeting (P18)

In some cases, parent meetings for the whole IE were considered. For this reason, it is important to specify if the meetings refer to the classroom investigated. We propose changing this question to "¿Aproximadamente cuántos padres asistieron a la última reunión relacionada al aula? [Si no se sabe, ponga 88 en las celdas]" (Approximately how

many parents attended the last meeting for this classroom? [If this is not known, put 88 in the cells]).

INSTRUMENT 5: DIRECTOR'S QUESTIONNAIRE

Initially, we eliminated this question on the type of IE (public or private) because the target population included only public IEs. We propose including this question in studies that do not have this limitation.

Duration of the school day (HT26)

It should be emphasized that the software converts the school day into minutes. This is done for the final calculation and the data in the database.

Who takes roll call (HT40.01)

In the pilot, we kept the option "senior teacher" to see if this special designation existed for this type of teacher. In the end, we proposed changing "senior teacher" to "docente."

Problems with the Spanish (HT58.1)

We propose reformulating this question to make it more disaggregated given that the responses do not indicate which member of the group has the problem.

INSTRUMENT 6: OBSERVATION OF THE EDUCATION INSTITUTION

Existence of a bathroom or latrine (SO14 a SO17)

We propose changing "baño" (bathroom) to "inodoro/retrete" (toilet, latrine). In the pre-pilot, the word "baño" was ambiguous and very heterogeneous.

Classrooms in use (SAM1)

According to the instructions, this question is separate from those that follow (SAM2 to SAM4) and does not need the three options (observation times) given that it is a unique data item for a particular IE. We recommend lowering slightly the options for time of observation.

Observation of classroom in activity (SAM2 to SAM4)

We recommend emphasizing that it is important to register the time observed for the comparative study with other IE.

Teachers and students outside the classroom (SAM3 and SAM4)

These questions were difficult to quantify due to movements that take place when the school has more than one patio.

INSTRUMENT 7: TEACHER QUESTIONNAIRE

Interview status (T2)

This question should be eliminated because a teacher who refuses to give the interview is replaced by another teacher.

Grades taught by the teacher (T17M)

In the pre-pilot, this question had multiple options and created confusion as to whether they taught in this classroom, in other shifts in the school and/or in other schools. To specify the responses and in light of the fact that this application is limited to full-grade schools (there are no teachers who teach several grades in a classroom), we recommend that the question contain a single response option. We propose changing to “¿Qué grado or grados enseña en este salón, en este turno, este año académico? [MARQUE LA OPCION QUE SEA APLICABLE—caso de escuela polidocente completa]” (Which grade or grades are taught in this classroom, during this shift, in this academic year? [MARK THE OPTION APPLICABLE—case of the full-grade school]).

Children repeating a grade (T22 y T23)

It is important to take into account that in Peru, there is no first-grade repetition, although some students do repeat the second grade. In this case, we suggest including those students that repeat the year.

Time spent on activities (T31)

For practical purposes and for ease of calculation, it is important for interviewers to show T31.08 prior to filling out the question options.

4.4 EGRA results

In this study, we analyzed different aspects of the EGRA’s psychometric behavior such as reliability, the possibility of instituting global scores, and the feasibility of cutting out some of EGRA’s sections.

Analysis indicates that it is possible to report EGRA in two main ways: considering the variables associated with the count (EGRA count) such as reading time, number of elements read, number of elements read incorrectly, number of elements read correctly; and considering the questions that are dichotomically answered either incorrectly or correctly (EGRA dichotomic). The EGRA count and the EGRA dichotomic show good reliability values.

Several ways of performing the EGRA count and the EGRA dichotomic have been proposed, including using the average of the counts and totaling the correct number of responses.

We have also analyzed the possibility of cutting down the EGRA, particularly in the letter reading section, without sacrificing its psychometric characteristics. These results can be considered if it is necessary to cut down the EGRA so that it can be applied in less time.

4.5 Relationships between SSME and EGRA

In the section corresponding to results, we have conducted an initial analysis that indicates the importance of these two instruments within the school environment. One is a performance indicator (EGRA) and the other provides information on different factors that can be associated with this performance (SSME). Given the study’s emphasis is on adapting

instruments, and the fact that it is a pilot, no further analysis has been done of the relationships between these two instruments. We have limited our analysis to showing the validity of the EGRA by demonstrating its concordance with the results obtained in the literature, such as the differences in the EGRA by school grade, student sex, and kindergarten preparation.

V. RECOMMENDATIONS

In general, the times reported in the pilot application and the strategy to apply the instruments over two days reflect a stopgap approach. It is possible to apply the instruments in only one day if prior coordination with the institutions has taken place.

Part of the time used to apply the pilot was used to confirm school selection with UGEL, obtain the director's authorization, and inform the teachers and parents about the evaluation.

Additionally, the pilot application took place in a context of controversy between the teachers' union and the Ministry of Education regarding the teacher evaluation and the evaluation of schools in general.

Positively speaking, the evaluation was completed in a 6-hour day, which could be reduced even more if we were to consider the new version of the SSME and the possibility of cutting out some of the EGRA sections.

Nevertheless, it is important to consider the rural reality, in which the day can be prolonged because the evaluators need to travel to the schools. In some cases, the time cannot be planned due to weather conditions and the need to update maps of existing routes. Under these circumstances, it would be a good idea to apply the EGRA on one day and the SSME on the other.

5.1 Sampling recommendations

In the pilot study, the target population was defined at public IEs that were full grade and that had a minimum of 3 sections in grades 2, 3 and 4 of primary for basic regular education. This represents 24% of the IEs' populations but only 68% of the student population.

Some recommendations on this point:

- Extend the study to multi-teacher full grade (which corresponds to 44% of the IEs) and single teacher (32% of the IEs), which represent 25% and 7% of the target student population respectively.
- Extend the study to private IEs, which represent 21% of the IEs and 17% of the total target population of students.

- Extend the study to populations that speak other important languages in the country, including Quechua and Aymara.
- Apply the instruments to IEs in different UGELS for the departments selected and not just in two.
- Coordinate with the Ministry of Education to obtain the list of IEs to define the sample framework.

5.2 Recommendations to modify the instruments used in the national application

The instruments were adapted based on the findings of the pilot study in public full grade IEs. In order to extend this for national application, the following recommendations should be taken into account:

- In a national application, it would be a good idea to include questions on the type of public/private IE in all of the instruments. For example, in the director's instrument, this question was eliminated because we assumed (correctly) that all of the IEs were public. We suggest adding the public/private question to the other instruments.
- In a national application, a question should be included on the type of IE as they are commonly referred to in the country ("unidocente," "polidocente," "multigrado," etc.). In the same way, it is necessary to verify this information in the IEs, maintaining the information provided by the Ministry of Education as a reference.
- Regarding the question on "grado" (grade), an exploratory review of the answers to this question indicated that responses can be arbitrary if the general instructions indicate that selected students be replaced by children from other grades. If this happens, the teacher interviewed may not teach the child under study. This comment also applied to the other instruments, particularly the teacher and parent instruments, which request information on grade.
- In the current political scenario (program to assess teachers) in the case of Peru, we recommend that future applications and other questionnaires be anonymous. As such, the teacher's name should be excluded.
- Regarding the presence of textbooks in the classroom, it is important to note that in Peru, texts in the private sector can be taken home. As such, it is possible that texts may or may not be present on the interview day. This question is valid for the public sector, where texts should be in the classroom.

- Register of classroom activities: We recommend emphasizing that the time should be registered for the three observations so that valid comparisons can be made with other schools.
- In terms of observation at the education institution, we recommend not forgetting to register the time at which the observation took place, to facilitate comparisons with other education institutions.
- In section VII of the pre-reading instrument (dictation), the score options for responses can be reduced from three (originally proposed) to two: correct and incorrect (with one or more mistakes).
- The versions of the instruments used for the fieldwork can be found in Annex 3 of this report. The recommended versions (post pilot) with the most significant changes are included in Annex 1, accompanied by a users' guide (Annex 2).

5.3 Recommendation for training for interviewers and materials

- Our main recommendation is to maintain the system for organizing training that was used in this study (defining coordinators, supervisors, and interviewers, who will receive cascaded training). This will encourage increased familiarity with application and supervision activities.
- In terms of materials, in Annex 2, we provide a guide developed for this study that was used for fieldwork and eventually modified with the results obtained.
- Along with this material, Annex 1 provides the instruments that were later modified after the pilot study. The instruments contain the most significant changes made following analysis.

5.4 Recommendations for fieldwork organization

Fieldwork organization depends on, among other aspects, defining the target population, the sampling system used, the IEs selected in the sample, and the application time. In the pilot study, two weeks were used by each application team.

- Given the number of instruments to be applied and considering that the study day in the country is approximately 5 hours long, we suggest maintaining the two-day application period for each IE. We propose that the first day be used for the following instruments: a teacher instrument, instruments for students in a classroom, the director instrument, and the instrument for the education institution. On the second day, the instruments can be applied to another classroom along with the parent instrument
- We suggest maintaining the information regarding a parallel authorization for the IEs—one from the Ministry of Education for the Regional Offices, the Regional

Offices to the UGELs, and the UGELs to the IEs; as well as another system where the interviewers present their credentials and authorizations directly to the IEs.

- We suggest maintaining the criteria of making appointments with the parents from one day to the next to apply the corresponding questionnaire.
- We suggest reconfirming the anonymous nature of the instruments and maintaining that it is not necessary to obtain consent forms, particularly from teachers, due to recent difficulties and the controversy generated by the teacher evaluations in the country.
- Despite the fact that the study discarded applying an instrument to the director of the UGEL and his officials due to overlapping with other operations and applications at a national level, we suggest incorporating this instrument into future evaluations.

5.5 Recommendations for gathering and handling data

Some recommendations include:

- Confirm and verify that all the IEs receive the official letter from the Ministry of Education in a timely manner. It should have the updated name of the authority and provide directors with timely notification of instrument application.
- Update the database entries indicating the accessibility of the IEs. The majority do not coincide with the information obtained in fieldwork—for example, the name of the director, address and telephone number of the IE.
- Update the database of maps with the information provided by the Ministry of Education.
- Deliver small gifts to other children (pencils and other items).
- Try to make prior contact with the IE participants in the pilot study.
- Reinforce quality control over instruments in Phase 1 (joint review by the supervisor and interviewers of the quality of the information obtained by interviewers), creating an additional control mechanism that is managed by the instrument specialist and that allows for corrections with interviewers and their respective supervisor.
- Consider extending the activity to build the database by one week to allow for double entry to minimize data-entry mistakes.

Annexes

[NOTE: Annexes are available only in the second volume of the Spanish version of this report, *Instantánea de efectividad en la administración de centros educativos: Estudio piloto in Perú—Anexos*]