The Gambia

Early Grade Reading Assessment (EGRA)

Results from 1,200 Gambian Primary Students Learning to Read in English—Report for the World Bank

January 2008

This publication was produced for review by the World Bank. Support for the development of EGRA was provided by the United States Agency for International Development (USAID) to EdData II, managed by RTI International. The publication was prepared by Liliane Sprenger-Charolles, Directeur de Recherche (DR1, Senior Researcher/Scientist), Centre National de la Recherche Scientifique (CNRS), and Descartes-University, Paris (France).

The Gambia
Early Grade Reading Assessment (EGRA)

Results from 1,200 Gambian Primary Students Learning to Read in English—Report for the World Bank

Prepared by Liliane Sprenger-Charolles, Directeur de Recherche (DR1, Senior Researcher/Scientist), CNRS, and Descartes-University, Paris (France).

Acknowledgements:

We would like to thank the Ministry of Education of The Gambia for undertaking this assessment, in particular Mr. Babucarr Bouy, Permanent Secretary of the Ministry of Education, and the following experts who participated in the local adaptation of the EGRA instrument, carried out the assessments in schools, and processed the data: Ms. Ida Njie, Mr. Ebrima N.H. Jarjou, Mr. Burama L.J. Jammeh, Mr. Omar B. Baldeh, Mr. Sang F.W. Gomez, Ms. Fatou Bittaye, Mr. Mohamed O. Kebbeh, Ms. Faatu Touray, Mr. Amadou M. Touray, Mr. Sulayman Barry, Mr. Lamin S. Sonko, Mr. Fabakary N.B. Jarju, Ms. Fatou Janneh, Ms. Fatou D. Bittaye, Mr. Ousman Senghore, Mr. Salifu Jobe, Mr. Karamba Sonko, Mr. Edward P. Mansal, Mr. Omar Jatta, and Mr. Kimintang Sowe.

In addition, we would like to thank the school directors and teachers, as well as the Gambian students who participated in this study.

Special thanks are extended to Mr. James Stevens from the World Bank for his advice and comments on the document and Ms. Amber Gove from RTI International, who assisted the Ministry of Education in Gambia with the local adaptation of the instrument, the training of enumerators, the assessments in schools, and review of this publication.

Prepared for
World Bank

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

The author's views expressed in this publication are based on the analysis of the results of the EGRA pilot studies conducted in Senegal and in Gambia, as well as on the outcomes of the current literature. They do not necessarily reflect the views of the World Bank, the United States Agency for International Development, or the United States Government.

Table of Contents

Page

ABBREVIATIONS iv

ABSTRACT 1

1. INTRODUCTION 2

2. METHOD 3
   2.1. Protocol used for the Gambian Study 3
   2.2. Participants 5

3. RESULTS 6
   3.1. Group differences 8
   3.2. Correlations and Regression 17

4. SUMMARY AND IMPLICATIONS 21
   4.1. Summary of the results of the experimental tasks and implications for future EGRA applications 21
   4.2. Control variables: SES, cultural environments, and gender 27

5. IMPLICATIONS FOR EDUCATION AND RESEARCH 28
   5.1 Summary of Implications for future EGRA applications 29
   5.2 Implications for educational policy 29

6. REFERENCES 30

Tables

Table 1. Gambian Children Learning to Read in English (N=1,200) 5

Table 2. Comparison of Results by SES Group (Means and SDs of EGRA Tasks) 9

Table 3. Comparison of Results by Gender (Means and SDs of EGRA Tasks) 10

Table 4. Comparison of Results by Nursery School Attendance (Means and SDs of EGRA Tasks) 11

Table 5. Comparison of Results by Literacy Status of the Parent (Means and SDs of EGRA Tasks) 12

Table 6. Differences Due to Home Language (Means and SDs) With and Without SES as a Covariable 16

Table 7. Correlations (N=1,200 students, 86 for the reading comprehension task) 18

Table 8. Prediction of Isolated-Word and Word-in-Context Reading

(1,184 Children) 20

Table 9. Prediction of Reading Comprehension (85 children) 21

Figures

Figure 1. Results for EGRA Tests for Gambian Children in Grades 1, 2, and 3 (Means and SDs) 13

Figure 2. Means and SDs for the Gambian (English) and Senegalese (French) Children 17

ABBREVIATIONS

ANOVA	analysis of variance
CNRS	National Center for Scientific Research (*Centre National de la Recherche Scientifique*, France)
CV	consonant-vowel
CVC	consonant-vowel-consonant
EGRA	Early Grade Reading Assessment
GPC	grapheme-phoneme correspondences
IEA	International Evaluation of Educational Achievement
OECD	Organisation for Economic Co-operation and Development
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PPVT	Peabody Picture Vocabulary Test
SD	Standard Deviation
SES	socioeconomic status
USAID	United States Agency for International Development

ABSTRACT

In most major international assessments (e.g., Programme for International Student Assessment–Organisation for Economic Co-operation and Development [PISA-OECD], and Progress in International Reading Literacy Study–International Evaluation of Educational Achievement [PIRLS-IEA]) children's reading skills are not assessed before fourth grade. For students who are poor readers, it is often too late to carry out efficient and effective remedial instruction. Indeed, to be efficient, remedial instruction should be conducted as early as possible. In addition, most major assessments are only composed of reading comprehension tasks, and do not take into account the level of word reading fluency (including accuracy and speed) and listening comprehension. However, research suggests that reading comprehension is associated with capacity in these complementary tasks.

To complement existing international assessments, a new protocol, Early Grade Reading Assessment (EGRA),[1] has been developed to assess the main skills that are known to predict reading success within the early grades of primary school (first to third grade). During a workshop held in Dakar, Senegal, the EGRA protocol was adapted to the local context in English (and in French) and then pretested in several schools in Gambia (and in Senegal). As a result of the adaptation and pretesting process, the original EGRA testing instrument was changed significantly. The protocol used in Gambia (and in Senegal) include two tasks assessing text comprehension (in spoken and written language), three tasks assessing accuracy and speed in the reading of high frequency words (isolated-words and word-in-context reading) and invented words (pseudowords), two spelling tasks, two phonemic awareness tasks, and a task assessing letter knowledge. EGRA also includes a student survey with questions about the student's cultural and linguistic environment (including home language and parents' literacy status) and socioeconomic status (SES).

A large-scale reading assessment was conducted in Gambia with 1,200 first, second, and third graders (randomly selected from 40 schools) who were learning to read in English. Three analyses were carried out. The first involved a comparison within the group, in which the effect of control variables including gender, home language, and SES was examined in relation to the children's results. In the second analysis, the pattern of correlations between the different tasks (and between these tasks and some control variables) was examined. Finally, regression analyses were carried out in order to determine the predictors of isolated-word and word-in-context reading, and reading comprehension.

The last two analyses indicated that the correlations between SES and the experimental tasks were not significant, and SES was not found to explain variance in word reading. This was also the case for phonemic awareness and the knowledge of

[1]. *EGRA: Early Grade Reading Assessment (RTI International for U.S. Agency for International Development [USAID]). For instruments and reports, see: www.eddataglobal.org.*

letter names, although the correlations between these tests and reading tests were high. The correlations between pseudoword and word reading were very high. Pseudoword reading was the only skill that systematically explained variance in word reading. Finally, the correlations between listening and reading comprehension were high, and listening comprehension was the only skill that explained variance in reading comprehension.

According to the previous analysis, SES was never found to explain word reading. However, comparisons within the group indicated that children from the lower SES category obtained lower scores on all the EGRA tasks than children from the higher SES category. Nevertheless, even after taking into account the effect of SES, the effect of some linguistic factors remained significant. This was especially the case for the effect of home language on the phoneme identification task (e.g. the scores of the Jola and of the Pular speakers were lower that those of the Mandinka speakers).

The role of the linguistic environment also emerged from a comparison between the results of the Gambian children who were learning to read in English and those from a similar assessment conducted with Senegalese children who were learning to read in French. Even after controlling for SES, the reading scores of the Gambian children were always lower than those of the Senegalese children. In addition, while in two cases (pseudoword reading and word spelling), the scores of the Gambian second graders were not found to be higher than those of the first graders, all the differences between first and second graders were significant for the Senegalese children. These results replicate those observed with monolingual children and suggest that it is more difficult to learn to read in English than in French, which is explained by the fact that the relations between graphemes and phonemes are far less transparent in English than in French.

1. INTRODUCTION

In most major international assessments,[2] children's reading skills are not assessed before fourth grade. For students whose reading level is low, it is often too late to carry out efficient and effective remedial instruction. Indeed, to be efficient, remedial instruction should be conducted as early as possible (see the report of the U.S. National Reading Panel, Ehri et al., 2001a and 2001b). In addition, most major assessments are only composed of reading comprehension tasks, and do not take into account the level of word reading fluency (including accuracy and speed) and listening comprehension. However, research suggests that reading comprehension is associated with capacity in these complementary tasks (for reviews, see Perfetti, 1985; Sprenger-Charolles et al., 2006; and Stanovich, 2000).

To overcome the flaws of previous international assessments, a new protocol, Early Grade Reading Assessment (EGRA), was developed by RTI International with support from the World Bank and the United States Agency for International Development (USAID). EGRA is designed to assess, within the early grades of

[2] *OECD (e.g., PISA, 2000); IEA (e.g., Elley, 1992, and PIRLS, 2003). See Abadzi (2006) for a review.*

primary school (first to third grade), the main skills that are known to predict reading success. EGRA includes a task assessing reading comprehension, a task assessing listening comprehension, and three tasks assessing accuracy and speed in word reading: high frequency, isolated-word reading; word-in-context reading; and pseudowords (i.e., invented words) reading. Pseudowords can only be read by using grapheme-phoneme correspondences (GPC), graphemes being the basic units of an alphabetic written system, which transcribe each phoneme of the spoken language. To measure student proficiency in the recognition of graphemes and phonemes, EGRA includes a task assessing the knowledge of simple graphemes (letters) and two tasks aimed at assessing the level of phonemic awareness. In addition, EGRA includes two spelling tasks and a student interview aimed at measuring linguistic, cultural, and socioeconomic factors that could influence reading acquisition.

EGRA is designed to be used in different linguistic contexts. We know that, in alphabetical systems, the degree of transparency of GPC produces differences in the *rate* at which children achieve the first few steps toward literacy (Seymour et al., 2003). Indeed, in languages with transparent GPC (shallow orthographies, such as Spanish, Italian, and even French[3]) the progression is very rapid; in languages where the correspondences between graphemes and phonemes are more complex (deep orthographies, e.g., English), reading acquisition can take several years. There are also *qualitative* differences between children who learn to read in languages with a deep orthography as compared to those who learn to read in languages with a shallow orthography. For example, English-speaking children use top-down lexical representations more than German-speaking children to supplement error-prone, bottom-up processes based on inconsistent GPC, especially for vowels; and in English (but less so in French or in German) reading errors involve more often vowels than consonants (for reviews, see Sprenger-Charolles, 2003, and Ziegler and Goswami, 2005). It is necessary to take into account these linguistic differences in the EGRA instrument, and to adjust the level of difficulties of the different tasks across languages as much as possible. We have tried to do this for the Gambian (and Senegalese) assessments.

2. METHOD

2.1. Protocol used for the Gambian Study

A workshop was held in Dakar, Senegal, in April 2007 to introduce Gambian (and Senegalese) officials to the EGRA pilot concept. English and French language versions were presented to the participants and each of these versions was adapted to the local context and pretested in several schools. As a result of the adaptation and pretesting process, the original EGRA testing instrument was changed significantly.

[3] *For statistics on English and French GPC consistency, see Peereman and Content, 1998; and Peereman et al., 2007.*

Tasks

Phonemic awareness

Students were asked to pronounce each sound of a spoken word (identification task) and to provide the number of sounds included in these words (counting task). Two scores were calculated: (1) the total number of words for which the number of phonemes was correctly counted, and (2) the total number of phonemes correctly pronounced.

Pre-reading skills

Concept of print: This task assessed the knowledge of some conventions of alphabetic systems, such as the direction of the writing. This task was not examined further because almost all children (85 percent in Grade 1 and 94 percent in Grade 3) obtained the maximum score for correctly identifying text direction: 1 point for left-to-right and 1 point for top-to-bottom.

Letter-name knowledge (1-minute task): The full set of letters was listed in random order. There were 10 letters to a row in a clear, large, and familiar font, with each letter presented multiple times. Randomization was used to prevent students from reciting a memorized alphabet. The complete alphabet (both uppercase and lowercase) was used. The score was the number of letters (out of 100) each student named correctly in 1 minute.

Reading skills (word reading and reading comprehension) and spelling skills

Word and Pseudoword reading (1-minute task)

Pseudoword (invented words) reading: This assessment includes a list of 50 one-syllable (2–3 letters) pseudowords presented in 10 rows, with five items per row. Forms are supposed to be legal for the language, using letters in legitimate positions, and consonant-vowel combinations that are typical of the language. The pseudowords are not supposed to be homophones of real words (for example, “kab” should not be used as it is a homophone of “cab”). The score was calculated by counting the number of pseudowords each student read aloud correctly in 1 minute.

Familiar word reading: A list of familiar words was selected from early grade reading materials. The list included 50 one-syllable (2–3 letters) words presented in 10 rows, with five items per row. The score was calculated by counting the number of words each student read aloud correctly in 1 minute.

Word reading in context: This assessment included one short narrative paragraph (~60 words). The score was calculated by counting the number of words each student read aloud correctly in 1 minute.

Reading comprehension and listening comprehension

This assessment was based on the text used to assess word-in-context reading. After the student read the text aloud, he or she was asked five simple questions about the passage, including two questions with “yes/no” responses. The score was the number of questions the student answered correctly.

A text similar to the one used to assess reading comprehension (~60 words) was used to assess listening comprehension. After the examiner read aloud this text, the student was asked three simple questions about the passage, including two with "yes/no" responses. The score was the number of questions the students answered correctly.

Spelling skills

A short sentence was read to the child, after which he/she was asked to write it down. The scores for spelling and spacing were as follows: 0 points for incorrect responses, 1 point for partially correct responses, or 2 points for correct responses. The remaining tasks were scored with 0 points for incorrect or 1 point for correct responses: mastering of the direction of the writing, capitalization, and punctuation. Two results were taken into account for the analysis: the spelling of the two key words, and the scores obtained in the four remaining tasks (spacing between words, direction of the writing, capitalization, and punctuation).

Students' Survey

This survey includes questions about the students' cultural and linguistic environment (e.g., home language and parents' literacy status,), and SES. We have created a SES variable based on the total number of yes responses for the 14 items taken into account in the survey (e.g., presence of articles in the home such as water taps, electricity, refrigerators, televisions, fixed-line telephones; and ownership of a car, bike, or mobile phone).

2.2. Participants

The participants were students in first, second, and third grades from Gambia who were learning to read in English. A total of 1,200 students were randomly selected from 40 schools. The students' main characteristics are in Table 1.

Table 1. Gambian Children Learning to Read in English (N=1,200)

				No response
	Gender: Male N=560; Female N=638			**N=2**
Grade et Age Age (Mean and Standard Deviation [SD]):	Grade 1 (N=419) 7.5 (1.4)	Grade 2 (N=389 9.1 (1.3)	Grade 3 (N=392) 10.1 (1.4)	
Age 5…….	N=010 (02%)			
Age 6…….	N=055 (13%)			
Age 7…….	N=184 (44%)	N=018 (05%)	N=002 (01%)	N=1 (Grade 2)
Age 8…….	N=099 (24%)	N=133 (34%)	N=029 (07%)	
Age 9…….	N=048 (11%)	N=118 (30%)	N=121 (31%)	
Age 10…….	N=015 (04%)	N=078 (20%)	N=127 (32%)	

				No response
Gender: Male N=560; Female N=638				**N=2**
Age 11……..	N=003 (01%)	N=019 (05%)	N=055 (14%)	
12 and above 12…	N=005 (01%)	N=022 (06%)	N=058 (15%)	
Education before Grade 1: Nursery=467 (39%), including 100 with Islamic Schooling (*Madrassa* or *Daara)*				N=6
Currently attending an Islamic school: N=710 (59%)				N=1
Socioeconomic Status (Mean and Standard Deviation): 4.9 (3.2)				
Parents' literacy status	**No one can read:** N=485 (40%)	**Only one reader:** N=359 (30%)	**Both are readers:** N=356 (30%)	N=0
Home language				N=11
English	N= 14 (including 6 with another language)….. …… 01%			
Mandinka	N=491 (including 3 with another language, besides English) 41%			
Pular	N=275 (including 19 with another language, besides English) 23%			
Wolof	N=195 (including 21 with another language, besides English) 16%			
Jola	N=108 (including 14 with another language, besides English) 09%			
Other	N=106 (including 21 with another language, besides English) 09%			

3. RESULTS

Three analyses were carried out. The first involved a comparison within the group, in which the effect of six control variables on the children's results was examined. The control variables were gender, grade level, attendance at nursery school, home language, parents' literacy status, and SES. For the SES variable, we created a factor based on the total number of yes responses for the 14 items taken into account in the children's survey (e.g., articles in the home such as water taps, electricity, and refrigerators). Because a similar assessment had been carried out with Senegalese children who are learning to read in French, the scores of the Gambian children were also compared to those of the Senegalese children (N=502). In the second analysis, the pattern of correlations between the different tasks (and between these tasks and some control variables) was examined. The third analysis, regression, was carried out

to determine the predictors of three crucial reading skills: isolated-word reading, word-in-context reading, and reading comprehension. Only three control variables were used for these analyses (chronological age, grade level, and SES).

The variables considered in the experimental tasks were the following: phonemic awareness (counting and identification), pre-reading skills (letter-name knowledge), reading skills (pseudoword, isolated-word, and word-in-context reading, plus reading comprehension), listening comprehension, and spelling skills (word spelling and other formal aspects of writing). To facilitate the result reporting, the scores were converted into percentages, except the scores for the 1-minute tasks.

There was a significant problem with the construction and scoring of the reading comprehension questions. Indeed, 82 percent of the Gambian children were unable to correctly read more than five words per minute for the 60-word text used to assess reading comprehension. Furthermore, 27 percent of these children (319 of all the children tested) were found to have a reading comprehension score between 1 and 5 (correct answers to the comprehension questions) whereas they were able to read only between 1 and 5 words. Note that it was necessary to have read at least 15 words to obtain a score of 1 correct response in the reading comprehension task (to obtain a score of 2, 3, 4, or 5 correct responses required reading, respectively, 25, 33, 50, and 54 words). These striking results may be attributable to several facts. First, students were allowed to read the text to its end, even if it took them longer than 1 minute. However, because only the words correctly read in 1 minute were scored and coded, it is not possible to know the total number of words each child had correctly read. Second, when the child was not able to read a word, after 3 seconds the examiner was instructed to provide the missing word, mark that word as incorrect, and encourage the child to continue. Therefore, in those cases where the examiner provided a large number of words, the task became a mixture of reading and listening comprehension. Third, it is probable that some students guessed on the responses to some questions, especially in the case of the yes/no questions. Finally, the examiner was instructed to ask the comprehension questions regardless of how many correct words the child had read. Therefore, for the correlation and the regression analyses, we have only examined the reading comprehension scores of the 86 children (7 percent of the population) able to read at least 20 words in 1 minute. This choice was made based on the fact that reading 20 words in 1 minute corresponds to reading 40 words in 2 minutes, a very low benchmark compared with U.S. results. For example, at the end of first grade in the U.S., 28 words per minute corresponds to the average reading level of students at the 25th percentile, and is considered at-risk for poor reading performance. For the comparison within the group, we have taken into account the original data.

Large floor effects (where students received 0 points because they were unable to perform the task or scored incorrectly on all of the items attempted) were also observed in most of the other reading tasks. For instance, the scores of 80 percent of the first graders were at the floor level for the word-reading task, and 91 percent of these children were totally unable to read a pseudoword. For third grade children, 0-point scores were still the case for approximately 50 percent and 76 percent of the

children for word and pseudoword reading, respectively. Floor effects were also prominent when assessing whether the children knew the correct spelling of two high frequency words (95 percent and 80 percent for first and third graders, respectively). Alternatively for first and third graders, floor effects were less important for the second spelling task (37 percent and 16 percent, respectively), for the letter-name task (35 percent and 4 percent, respectively), for the listening comprehension task (58 percent and 20 percent, respectively), and for the two phonemic awareness tasks (around 50 percent and 30 percent for each task for students in both grades). There were no large ceiling effects, except for the first pre-reading task.

3.1. Group differences

These analyses involved a comparison within the group, in which the effect of six control variables on the children's results was examined. The control variables were gender, grade level, attendance at nursery school, home language, parents' literacy status, and SES. Because a similar assessment had been carried out with Senegalese children who were learning to read in French, the scores of the Gambian children were also compared to those of the Senegalese children (N=502).

According to the results of international assessments (e.g., PISA, 2000), it is expected that the children from the higher SES category outperformed those from the lower SES category and that, at least in reading tasks, girls outperformed boys. The other factors considered likely to have a positive effect on results are schooling in a higher grade (compared to a lower grade); attending a nursery school; having at least one literate parent; learning to read in the mother tongue; and learning to read in a language with a more transparent orthography than the orthography for the English language.

For SES, the participants were split into two groups based on the total number of yes responses for the 14 items taken into account in the children's survey: those from the lower SES category (SES scores lower than 8) and those from the higher SES category (SES scores equal to or more than 8). For the parents' literacy status, the participants were also split into two groups: a group with children who reported at least one literate parent, and a group of children without literate parents.

The significance of the differences has been checked by t-tests. In the cases where there were unequal variances between the groups, the threshold of significance taken into account was for "equal variance not assumed" (threshold $p < .05$, at least).

Children's Socioeconomic Status (SES)

The population was divided into two groups: the children with a SES score equal to or more than 8 (the higher SES category, approximately equivalent to the first or top quartile), and those with a SES score below 8 (the lower SES category, equivalent to the second, third, or fourth quartiles).

The results are presented in Table 2. The scores of the children with a low SES level were systematically inferior to those of the children from the other group.

Table 2. Comparison of Results by SES Group (Means and SDs of EGRA Tasks)

	Higher SES (>8) N=271	Lower SES (<8) N=929	Significant Differences
Phoneme Counting: %	37.7 (27.7)	24.7 (26.6)	p < .01
Phoneme Identification: %	38.1 (28.5)	24.1 (27.1)	p < .01
Letters/Minute	33.8 (26.7)	21.8 (21.4)	p < .01 unequal variance
Pseudowords/Minute	3.34 (9.78)	0.67 (2.78)	p < .01 unequal variance
Isolated Words/Minute	6.71 (14.4)	1.33 (4.0)	p < .01 unequal variance
Words in Context/Minute	12.5 (25.6)	2.88 (8.35)	p < .01 unequal variance
Reading Comprehension: %	33.8 (35.1)	14.7 (24.3)	p < .01 unequal variance
Listening Comprehension: %	61.2 (37.2)	37.3 (37.5)	p < .01
Spelling 1: %	15.2 (30.7)	3.69 (14.6)	p < .01 unequal variance
Spelling 2: %	28.6 (25.0)	21.3 (15.7)	p < .01 unequal variance

Gender, nursery school attendance, and literacy status of the parent

Because of the effect of SES on the children's results, two analyses have been carried to examine the other control variables. First, the significance of the differences has been checked by t-tests (same procedure as above). Differences that appeared to be significant in these analyses have then been examined in the analyses of variance (ANOVAs), with SES as a covariable.

Table 3 highlights the relationship between gender and student results. Boys obtained higher scores than girls on 4 of the 10 tasks: 2 out of the 3 tasks involved spoken language processing, and 2 out of the 7 tasks involved written language processing. There was no change in the results when SES was taken into account.

Table 3. Comparison of Results by Gender (Means and SDs of EGRA Tasks)

	Boys (N=560)	Girls (N=638)	Significant Differences	With SES
Phoneme Counting: %	29.2 (27.8)	26.3 (26.9)		
Phoneme Identification: %	30.0 (28.3)	24.9 (27.6)	$p < .01$	$p < .01$
Letters/Minute	26.2 (24.3)	22.9 (22.1)	$p < .05$ unequal variance	$p < .01$
Pseudowords/Minute	1.45 (5.56)	1.11 (5.19)		
Isolated Words/Minute	2.79 (7.58)	2.32 (8.33)		
Words in Context/Minute	5.30 (13.8)	4.84 (15.6)		
Reading Comprehension: %	21.4 (30.0)	16.8 (26.4)	$p < .01$ unequal variance	$p < .01$
Listening Comprehension: %	48.3 (39.8)	37.8 (37.0)	$p < .01$ unequal variance	$p < .01$
Spelling 1: %	6.87 (20.3)	5.72 (19.8)		
Spelling 2: %	23.6 (19.0)	22.3 (17.9)		

The effect of nursery school attendance is presented in Table 4. Student enrollment in a nursery school was found to have a systematic positive effect. There was only one change when SES was taken into account: the difference for the second spelling task became nonsignificant.

Table 4. Comparison of Results by Nursery School Attendance (Means and SDs of EGRA Tasks)

	Yes (N=468)	No (N=726)	Significant Differences	With SES
Phoneme Counting: %	32.3 (28.6)	24.6 (26.1)	p < .01 unequal variance	p < .05
Phoneme Identification: %	32.6 (29.1)	23.9 (26.9)	p < .01 unequal variance	p < .01
Letters/Minute	29.3 (26.0)	21.4 (20.8)	p < .01 unequal variance	p < .01
Pseudowords/Minute	2.15 (7.79)	0.71 (2.78)	p < .01 unequal variance	p < .05
Isolated Words/Minute	4.25 (11.4)	1.46 (4.25)	p < .01 unequal variance	p < .01
Words in Context/Minute	8.52 (20.8)	2.86 (8.23)	p < .01 unequal variance	p < .01
Reading Comprehension: %	24.7 (32.6)	15.4 (24.4)	p < .01 unequal variance	p < .01
Listening Comprehension: %	50.2 (38.9)	37.8 (37.7)	p < .01 unequal variance	p < .01
Spelling 1: %	10.1 (25.7)	3.89 (14.9)	p < .01 unequal variance	p < .01
Spelling 2: %	25.0 (21.5)	21.6 (15.9)	p < .01 unequal variance	

The results of the comparison between the children who had said that at least one of their parents is able to read and write and the children who reported that neither parent was literate are presented in Table 5. Only three differences remained significant when SES was taken into account: one in a task involving written language processing (letter knowledge), and two in tasks involving spoken language processing (listening comprehension and phoneme identification).

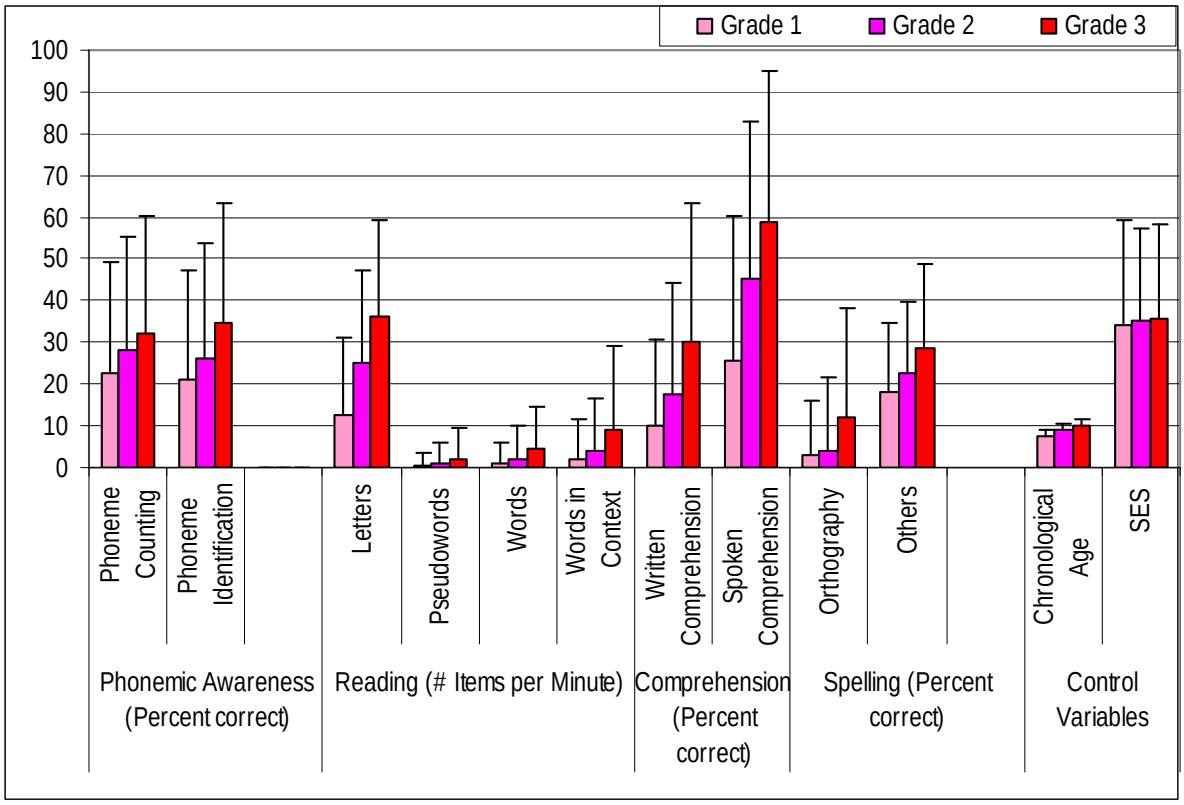
Table 5. Comparison of Results by Literacy Status of the Parent (Means and SDs of EGRA Tasks)

	At least one literate parent (N=715)	No literate parent (N=485)	Significant Differences	With SES
Phoneme Counting: %	30.3 (27.9)	23.7 (26.1)	p < .01 unequal variance	
Phoneme Identification: %	30.4 (29.0)	22.6 (25.9)	p < .01 unequal variance	p < .05
Letters/Minute	27.2 (24.1)	20.5 (21.2)	p < .01 unequal variance	p < .01
Pseudowords/Minute	1.63 (6.38)	0.75 (3.27)	p < .01 unequal variance	
Isolated Words/Minute	3.32 (9.58)	1.40 (4.48)	p < .01 unequal variance	
Words in Context/Minute	6.44 (17.4)	3.01 (9.27)	p < .01 unequal variance	
Reading Comprehension: %	21.9 (30.1)	14.8 (24.6)	p < .01 unequal variance	
Listening Comprehension: %	47.8 (39.0)	35.3 (37.0)	p < .01 unequal variance	p < .01
Spelling 1: %	8.22 (22.8)	3.45 (14.8)	p < .01 unequal variance	
Spelling 2: %	24.2 (19.5)	21.0 (16.6)	p < .01	

Grade-level effect

Figure 1 shows the results of EGRA tasks for the first, second, and third graders (N = 419, 389, and 392, respectively). There were no SES differences between first and second graders or between second and third graders based on the control variables, and the chronological age of children in higher grades was systematically higher than the chronological age of the children in lower grades. On the experimental tasks, the scores of second graders were higher than those of first graders, except in two cases (pseudoword reading and first spelling task). The differences between second and third graders were all significant ($p < .05$, at least), except for the phoneme counting task ($p < .06$).

Figure 1. Results for EGRA Tests for Gambian Children in Grades 1, 2, and 3 (Means and SDs)



Effect of the language spoken at home

The goal of this analysis is to determine the potential effect of the child's home language on the variables from the experimental tasks. The Gambian children were speaking at least six different languages at home, with a sufficient number (statistically speaking) of children within each linguistic group (excluding the English-speaking group there were at least 14 children in each group, including bilingual children, counted once for each language spoken). The other languages were Wolof (174 children), Pular (254, plus 2 children with missing achievement data), Jola (94), Mandinka (486, plus 2 children with missing achievement data), plus 106 children for whom the home language was not specified, and 56 for whom there were two home languages, English not included. Because there were too many different multilingual children (reporting multiple linguistic groups, e.g., Wolof-Pular, Pular-Mandinka), these children were not included in the analysis. The analysis was carried out with the variables from the experimental tasks for which the results of all the children were available (the results of the reading comprehension task were thus not taken into account). We also examined the SES of the different linguistic groups. ANOVAs were carried out with the six linguistic groups.

There were SES differences between the linguistic groups. Indeed, the SES of the English speakers was higher than the SES of the children from all other linguistic groups, the SES of the Wolof-speaking group was also higher than the SES of all the remaining linguistic groups (except for children speaking an unknown language, where there was no SES difference between these two groups), and the SES of the

Pular and Jola speakers was lower than the SES of the children from all the other linguistic groups, with no difference between Pular and Jola speakers. Because of these SES differences, two analyses were carried out. First, the significance of the differences was checked by t-tests (same procedure as above). Differences that appeared to be significant in these analyses were then examined in an ANOVA, with SES as a covariable. A summary of the results is provided in Table 6. Differences that were to the advantage or to the disadvantage of a specific linguistic group are indicated respectively by '>' or '<' (threshold: $p < .05$).

The English speakers (highest SES) were very rarely found to outperform the other linguistic groups. Only one difference (out of seven significant differences according to t-tests) was still significant after having controlled for SES: the English speakers surpassed the Pular speakers in word-in-context reading.

Second, the SES of the Wolof speakers was higher than the SES of all but one of the other linguistic groups: the children who spoke an unknown language (no significant difference). The Wolof speakers outperformed those children in two tasks (letter-name and word-in-context reading). In the comparisons with the other groups, only two differences remained significant after having controlled for SES (out of the 12 significant differences according to t-tests): the Wolof speakers outperformed the Mandinka speakers in the word-in-context reading task and the first spelling task.

Third, there were no SES differences between the groups that had the lowest SES scores: the Pular and the Jola speakers. However, the Pular outperformed the Jola speakers in the phoneme-counting task, the reverse being observed for the listening comprehension task. The comparison with the other linguistic groups indicated that, after having controlled for SES, only six differences (3 out of 13 significant differences according to t-tests for the Pular speakers, and 3 out of 6 for the Jola speakers) remained significant. In the phoneme identification task, the scores of these two groups were lower than those of both the Mandinka speakers and the speakers of an unknown language; in the phoneme-counting task, the Jola speakers scored lower than the Mandinka speakers; and in the word-in-context task, the Pular speakers scored lower than the English speakers.

Fourth, the SES of the Mandinka speakers was lower than that of the English and Wolof speakers but higher than that of the Jola and Pular speakers. After having taken SES into account, there were still some significant differences between these children and the other linguistic groups (5 out of the 11 significant differences according to t-tests). The Mandinka speakers obtained lower scores than the Wolof speakers in the word-in-context reading task and in the first spelling task whereas they obtained higher scores than the Jola and Pular speakers in the two phonemic tasks, except in one case (where the difference in the phoneme identification task was not significant for the Pular speakers).

Finally, the scores of the speakers of an unknown language were higher than those of the children from a lower SES group only twice (out of the five significant differences according to t-tests): when compared to both the Pular and the Jola speakers in the phoneme identification task. In addition, although there was no SES difference

between these children and the Wolof speakers, their scores were lower in two tasks (letter-name knowledge and word-in-context reading).

In the previous analysis, the results were discussed in both directions (e.g. Jola vs. Pular, and Pular vs. Jola). The total number of significant differences must thus be divided by two. The main finding was that most of the differences in the two phonemic awareness tasks (five out of the eight that were significant according to t-tests) remained significant after having taken SES into account: one in the phoneme counting task (the scores of the Jola were lower than those of the Mandinka speakers) and four in the phoneme identification task (the scores of the Jola and of the Pular speakers were lower that those of both the Mandinka speakers and the speakers of an unknown language). Alternatively, only three other differences remained significant after having taken SES into account (out of five): in word-in-context reading, the English speakers surpassed the Pular speakers, and the Wolof speakers surpassed the Mandinka speakers; and in the first spelling task, the Wolof speakers surpassed the Mandinka speakers. These findings call for a more in-depth scrutiny of the effect of the language spoken at home on the children’s academic results.

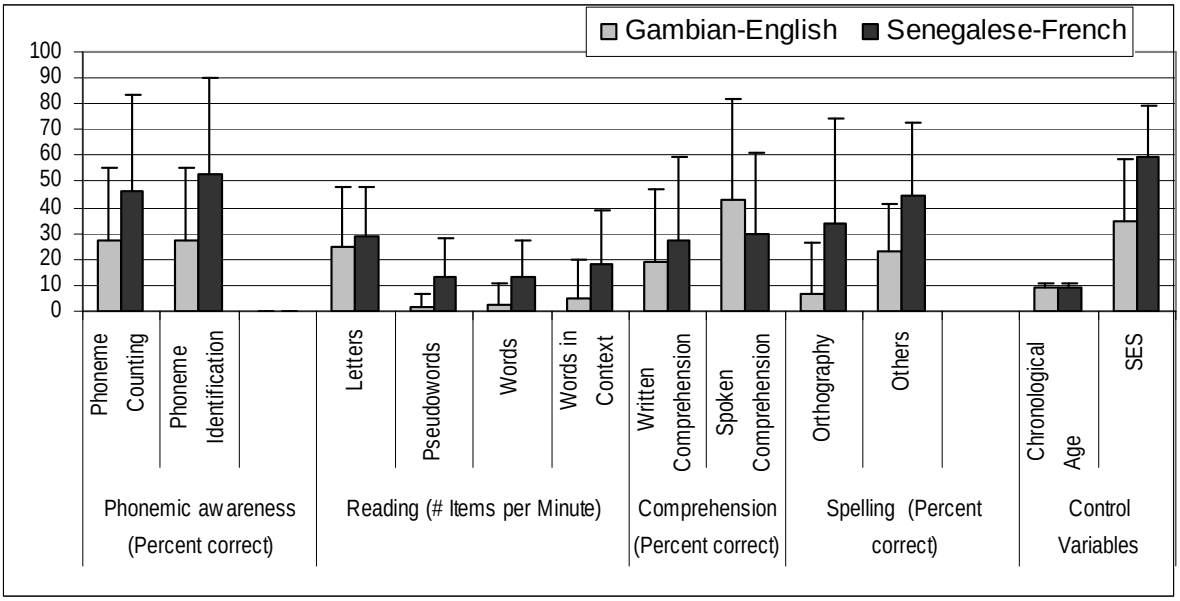
Table 6. Differences Due to Home Language (Means and SDs) With and Without SES as a Covariable

	1	2	3	4	5	6	1. English					2. Wolof					3. Pular					4. Jola					5. Mandinka					6. Other				
	English N=14	***Wolof N=174***	***Pular N=254***	***Jola N=094***	***Mand. N=486***	***Other N=106***	2	3	4	5	6	1	3	4	5	6	1	2	4	5	6	1	2	3	5	6	1	2	3	4	6	1	2	3	4	5
Socio-economic Status	8.29 (3.41)	6.09 (3.69)	3.59 (2.55)	3.99 (2.80)	5.10 (3.12)	5.60 (3.12)	> *	> *	> *	> *	> *	< *	> *	> *	> *		< *	< *		< *	< *	< *	< *		< *	< *	< *	< *	> *	> *		< *		> *	> *	
Phoneme Counting (%)	1.50 (1.56)	1.45 (1.33)	1.31 (1.41)	1.00 (1.26)	1.49 (1.38)	1.47 (1.35)								> *					> *				< *	< *	< *	< *				> *					> *	
	ANOVAs with SES as a covariable (N=No SES Diff.)																		N					N	*					*							
Phoneme Identificat (%)	3.93 (4.65)	4.14 (3.99)	3.02 (3.82)	2.94 (3.54)	4.26 (3.95)	4.49 (3.94)								> *						< *	< *		< *		< *	< *			> *	> *				> *	> *	
	ANOVAs with SES as a covariable (N=No SES Diff.)																			*	*				*	*			*	*					*	*	
Letters /Minute	37.0 (26.4)	28.5 (24.4)	22.9 (22.4)	23.5 (24.1)	23.8 (23.1)	21.9 (20.2)		> *		> *	> *		> *		> *	> *	< *	< *									< *	< *				< *	< *			
	ANOVAs with SES as a covariable (N=No SES Diff.)															N																		N			
Pseudo-words /Minute	3.86 (8.35)	1.48 (4.98)	0.77 (3.23)	1.61 (5.84)	1.19 (5.21)	1.46 (7.39)																														
	ANOVAs with SES as a covariable (N=No SES Diff.)																																			
Isolated Words /Min	7.71 (14.0)	4.18 (9.71)	1.40 (5.05)	2.63 (8.32)	2.33 (7.26)	2.33 (8.53)							> *		> *			< *		< *								< *	> *							
	ANOVAs with SES as a covariable (N=No SES Diff.)																																			
Words (Context) /Min	16.1 (22.9)	8.74 (18.6)	2.99 (9.13)	6.28 (18.1)	4.17 (13.1)	4.37 (14.3)		> *					> *		> *	> *	< *	< *										< *					< *			
	ANOVAs with SES as a covariable (N=No SES Diff.)							*							*	N	*											*					N			
Listening Comp (%)	2.00 (1.18)	1.41 (1.27)	1.14 (1.10)	1.40 (1.10)	1.28 (1.17)	1.31 (1.14)		> *		> *	> *		> *				< *	< *	< *					> *			< *					< *				
	ANOVAs with SES as a covariable (N=No SES Diff.)																		N					N													
Spelling 1 (%)	0.79 (1.37)	0.43 (1.02)	0.16 (0.61)	0.28 (0.92)	0.21 (0.73)	0.25 (0.75)							> *		> *			< *										< *								
	ANOVAs with SES as a covariable (N=No SES Diff.)														*													*								
Spelling 2 (%)	1.14 (0.66)	1.03 (0.84)	0.81 (0.64)	0.93 (0.75)	0.93 (0.76)	0.92 (0.75)							> *					< *		< *									> *							
	ANOVAs with SES as a covariable (N=No SES Diff.)																																			

Comparison of results of children learning to read in English and French

Because a similar assessment has been carried out with Senegalese children who are learning to read in French, the scores of the first, second, and third grade Senegalese children (N=502) were compared to those of the first, second, and third grade Gambian children (N=1200). The results are presented in Figure 2.

Figure 2. Means and SDs for the Gambian (English) and Senegalese (French) Children



There was no difference between the Gambian children and the Senegalese children in chronological age. However, the SES of the Gambian children was lower than that of the Senegalese children. Therefore, the analysis of the results of the experimental tasks was carried out both with and without the covariable SES. Even after controlling for SES, the scores of the Senegalese children were always higher than those of the Gambian children, except for the listening comprehension task, where the Gambian children's scores were higher. These results replicate those observed with monolingual children and suggest that it is more difficult to learn to read in English than in French, because the relationships between graphemes and phonemes are far less transparent in English than in French.

3.2. Correlations and Regression

Correlations

A correlation indicates the strength of a relationship between two measures. Thus, we expected to observe high correlations between tasks supposed to assess similar skills: for instance, between word-in-context and isolated-word reading, and between written and spoken language comprehension.

The correlations have been calculated between all the EGRA experimental tasks, except the pre-reading test, which assessed whether the children were aware of the direction of the writing because there were strong ceiling effects (most children score 100 percent). For the reading comprehension task, only the children who were able to read at least 20 words in 1 minute of the 60-word text have been included in these analyses. The rationale for this choice is explained above in Section 3. We have also examined the correlations between the experimental tasks and three control variables (chronological age, grade level, and SES). The results are presented in Table 7. The correlations superior or equal to .41, .33, and .25 are indicated by ***, **, and *, the thresholds of significance being .001, .01, and .05, respectively, for the correlations involving the smallest number of children, those between the reading comprehension task and the other tasks (N = 86, i.e. the 86 children with a minimum of 20 correct words per minute). The choice of a standard threshold of significance, whatever the size of the population, allows for easier comparisons. In addition, the highest correlations (more than .60) are highlighted in grey.

Table 7. Correlations (N=1,200 students, 86 for the reading comprehension task)

	1 (PC)	2 (PI)	3 (L/M)	4 (P/M)	5 (MI/M)	6 (MC/M)	7 (RC)	8 (SC)	9 (S1)	10 (S2)
1. Phoneme Counting: %										
2. Phoneme Identification: %	***.51									
3. Letters/Minute	**.33	*.32								
4. Pseudowords/Minute	.18	.22	***.41							
5. Isolated Words/Minute	.21	.24	***.53	***.86						
6. Words in Context/Minute	.22	.23	***.54	***.80	***.93					
7. Reading Comprehension: %	*.25	.19	***.43	**.39	***.47	***.47				
8. Listening Comprehension: %	**.33	*.32	***.47	.22	*.31	**.34	***.42			
9. Spelling 1: %	.19	.24	***.52	***.66	***.77	***.77	***.49	*.31		
10. Spelling 2: %	.20	*.25	***.46	***.50	***.58	***.59	***.54	*.29	***.62	
11. Chronological Age	.11	*.25	*.31	.08	.10	.11	-.20	*.29	.14	.19
12. Grade Level	.14	.20	***.41	.12	.16	.20	-.09	**.35	.19	.23
13. Socioeconomic Status	*.25	*.25	*.25	.23	*.32	*.31	*.30	*.30	*.28	.19

For the correlations between the control variables and the experimental tasks, eight correlations with SES, although weak, were significant (.25 to .32); nonsignificant

correlations were observed between SES and pseudoword reading or the second spelling task. The correlations between the chronological age of the children and the results they had obtained in the experimental tasks were low (mean = .14). A similar pattern of results was observed for the correlations between the grade level in which the children were enrolled and the experimental tasks (mean = .19).

For the correlations between the EGRA experimental tasks, first, the phonemic counting task was significantly correlated with the phonemic identification task (.51). The mean of the correlations between these two tasks and the three reading tasks (pseudowords, isolated word, and words in context) were .20 for the phoneme counting task, and .23 for the phoneme identification task. Second, the mean of the correlations between the letter per minute task and the three reading tasks was .49. Third, the correlations between pseudowords and isolated word or word in context, as well as the correlations between the two tasks involving words, were very high (respectively .86, .80, and .93). Fourth, in spite of the very strong floor effects observed in the first spelling task, the correlations between that task and the word and pseudoword reading tasks were also very high (mean = .73), the correlations between the other spelling task and the word and pseudoword reading task being lower (mean = .56). Finally, between the two tasks involving reading and listening comprehension, the correlations were lower than expected (.42), as were the correlations between word and pseudoword reading and reading comprehension (.39 to .47). This is likely due to the fact that, by dropping those cases where children were able to read less than 20 correct words per minute, we have reduced the variation in the results.

Regressions

The goal of the regression analyses is to identify the predictors of word reading levels (isolated word and word in context), and reading comprehension. For this analysis, we have taken into account as predictors the variables from the experimental tasks, plus some control variables (chronological age, grade level, and SES). We have calculated the total part of the variance explained by all the predictors and the *unique part of the variance* explained by each of these predictors. Except for the prediction of reading comprehension, the regression analyses were conducted using the scores of all students.

The predictors for the reading of isolated-word and words-in-context reading were phonemic awareness skills (phoneme counting and identification), letters per minute, pseudowords per minute, listening comprehension, and spelling skills (spelling tasks 1 and 2). Control variables (chronological age, grade level, and SES) were also taken into account. The results are presented in Table 8.

A very large amount of the variance in isolated-word reading (more than 83 percent) was explained by the variables entered in the model. The unique contribution of the three control variables was very low (less than 1 percent) and nonsignificant. Among the variables from the experimental tasks, the unique contribution of pseudoword reading skills was significant (19 percent). The sole other variable that explained a unique and

significant part of the variance in word reading came from the first spelling task (around 3 percent). Less than 1 percent of the variance in word reading was explained by the unique contribution of the other variables from the experimental tasks.

As was the case in the previous analysis, a very large amount of the variance in word-in-context reading was explained by the variables entered in the model (more than 77 percent). The unique contribution of the three control variables was very low (less than 1 percent) and nonsignificant. Among the variables from the experimental tasks, only the unique contributions of pseudoword reading and of the first spelling task were significant: 13 percent and 4 percent respectively. Less than 1 percent of the variance in word reading was explained by the unique contribution of the other experimental variables.

Table 8. Prediction of Isolated-Word and Word-in-Context Reading (1,184 Children)

	Isolated Word	Word in Context
	Total explained variance	
	.835	.773
	Added (unique) and significance	
Socioeconomic Status	.003	.001
Chronological Age	.001	.002
Grade Level	.000	.000
Phoneme Counting	.000	.000
Phoneme Identification	.000	.001
Letters per Minute	.008	.007
Pseudoword per Minute	.194*	.130*
Spoken Comprehension	.000	.003
Dictation-1 (orthography)	.029*	.041*
Dictation-2	.003	.004

For reading comprehension, the number of children involved in this analysis is low. Therefore, only seven predictors have been taken into account: phonemic awareness (counting and identification), pseudoword reading, isolated-word reading, word-in-context reading, spoken comprehension, and word spelling. We have not taken into account the results of the letter per minute task, and those of the second spelling task because they were never found to significantly contribute to the explanation of the variance in the previous analyses. For the same reason, control variables such as the age

of the children, their grade levels, and their SES were not included in the analyses. The results (see Table 9) indicate that 46 percent of the variance in reading comprehension was explained by the variables entered in the model. Among the variables from the experimental tasks, only listening comprehension and spelling skills uniquely and significantly explained the variance in reading comprehension (14 percent and 6.5 percent, respectively). Less than 1 percent of the variance in reading comprehension was explained by the unique contribution of the other variables from the experimental tasks, except 1.7 percent of the variance explained by the outcomes of the phoneme counting task.

Table 9. Prediction of Reading Comprehension (85 children)

	Total explained variance
	.458
	Added (unique) and significance
Phoneme Counting	.017
Phoneme Identification	.003
Pseudoword per Minute	.007
Isolated Word per Minute	.006
Word in Context per Minute	.000
Spoken Comprehension	.141*
Dictation-1 (orthography)	.065*

4. SUMMARY AND IMPLICATIONS

4.1. Summary of the results of the experimental tasks and implications for future EGRA applications

Pre-reading skills

The scores obtained in the reading orientation test can only vary from 0 to 2 and reached the ceiling level for almost all the children from first grade. Consequently, it is difficult to use the results of this test, especially in the analysis of the correlations. Therefore, it is unnecessary to include this task in future EGRA applications, unless local educators suspect that children have not acquired these basic skills.

The correlations between the letter-per-minute task and the other 1-minute tasks (pseudoword, isolated word, and word in context) were significant. However, the scores obtained in the letter-per-minute task were never found to contribute to the explanation of

the variance in word reading or in reading comprehension. This result is in line with those generally reported in the literature. According to Wagner et al. (1997), for instance, when reading skills are taken into account, the rapid naming of letters added no significant part of the variance in reading level.

In addition, the letter name sometimes differs from letter sound. Furthermore, the use of GPC requires that the student masters the sound of the letters. A result supporting the claim that it is better to use letter sound than letter name is that scores obtained in this task were less strongly correlated with the pseudoword and word (isolated or in context) reading for the Gambian children than for the Senegalese children (see Sprenger-Charolles, 2007), whatever the language they were learning to read (French or Wolof). This result might be due to the fact that, for the tasks in French and in Wolof, both the letter name and the letter sound were accepted as correct responses and not only the name of the letters as was the case for the English task. A letter-sound task is thus more appropriate than the letter-name task.

The letter knowledge task might thus be dropped from the EGRA protocol. However, because it is crucial to assess the knowledge of the basic visual units of an alphabetic system, this task might be replaced by a task aimed to assess the ability to discriminate true letters (A, a, P, p, B, b) from nonalphabetic symbols (such as ☺, ♣, ✶, ♪) and reversed letters (such as Ə, Λ, Ш, Я, ʎ). The inclusion of this task would allow the assessment of visual skills, not examined in the current EGRA protocol. This new task would be a 1-minute test, and will include 60 letters and 20 signs (10 visual symbols and 10 reversed letters, with 1 to 3 signs, symbols, or letters per line) spread over 10 lines. Two points could be awarded for the correct designation of a reversed letter and 1 point for the correct designation of a symbol.

Word reading skills

The correlations between pseudowords and isolated words or words in context were highly significant. Therefore, even for the children who were learning to read in English, there were no strong dissociations between the phonological reading procedure that may be used to read new words (pseudowords) and the orthographic reading procedure that may be used to read high frequency words. Furthermore, only pseudoword reading uniquely explained the results of the word reading tasks (isolated words and words in context). These results are not congruent with what can be expected according to Seymour's double foundation model of reading acquisition, which is mainly based on the results obtained by English-speaking children (2003). They are more in line with some models of reading acquisition based on data obtained by English-, French-, German-, and Spanish-speaking children, for example (e.g., Sprenger-Charolles et al., 2006; Ziegler and Goswami, 2005), that state that reading acquisition strongly depends on mastering GPC (see also Share, 1995).

Another important finding was that floor effects were very large in these three tasks. For instance, 80 percent of the Gambian first graders were totally unable to read a word and 91 percent a pseudoword. In third grade, this was still the case for approximately 50 percent and 76 percent of the children for word and pseudoword reading, respectively.

For future EGRA applications, it will thus be necessary to simplify these tasks more than current tasks. It may be necessary to rely only on very short, high frequency words with regular GPC primarily in the two first lines of the test in order to allow poor readers to read at least some items. Likewise, pseudowords have to be as short and simple as possible and to include only simple GPC, still primarily in the two first lines of the test. It would be preferable to build up pseudowords from the words used in the isolated word reading task by changing only the consonant(s), not the vowel(s) because, in English, it is mainly vowel pronunciation that poses problems.

Reading comprehension

According to the literature, the level of reading comprehension is supposed to be explained by the listening comprehension level and by the level of word reading. In the regression analyses, we have thus taken into account the results obtained in the three word reading tasks and in the listening comprehension task, plus those obtained in the first spelling task and in the two phonemic awareness tasks. A large part of the variance in reading comprehension was explained by the variables entered in the model. However, only listening comprehension and word spelling skills (not word reading skills) uniquely and significantly explained the variance in reading comprehension.

The fact that word reading skills were not found to predict the level of reading comprehension might be due to the inclusion in the regression analysis of only the children able to read at least 20 words per minute from the text. This choice was made because 82 percent of the Gambian children were unable to correctly read more than 5 words per minute of the text used to assess reading comprehension. In addition, 27 percent of these “nonreading” children were found to have a reading comprehension score between 1 and 5 (correct answers to the comprehension questions), while it was necessary to have read at least 15 words to obtain a score of 1 correct response in the reading comprehension task.

These striking results may be attributable to several facts. First, as mentioned previously, students were allowed to read the text to its end, even if it took them longer than 1 minute. However, because only the words correctly read in 1 minute were scored and coded, it is not possible to know the total number of words each child had correctly read. Second, when the child was not able to read a word after 3 seconds, the examiner was instructed to provide the missing word, and encourage the child to continue. Therefore, in those cases where the examiner provided a large number of words, the task became a mixture of reading and listening comprehension. Third, it is probable that some students guessed on the responses to some questions, especially in the case of the yes/no questions.

These problems make it necessary to revise the test used to assess both word-in-context reading and reading comprehension. Instead of a 60-word text, it would be preferable to use a shorter text, with 4–5 sentences and no more than 10 words in each sentence. The child should first read the text. The time and the number of words read correctly in 1 minute should be recorded as before with two exceptions: first, the number of words read correctly in 2 minutes should also be recorded; second, when the child is not able to read a word, the examiner should not provide it (just encourage the child to continue when he/she spends more than 3 seconds on a word). Afterward, the sentences should be presented one after another, with questions being asked immediately after the child reads each sentence. This procedure would reduce floor effects and lessen the memory load. In addition, it would be preferable to avoid verbal responses to the questions, especially when the children are not native speakers of the language being tested. One way to avoid this problem is to rely on a set of three pictures (not two, to minimize the probability of guessing), with only one corresponding to the correct answer. A similar protocol could be used to assess the listening comprehension.

Listening comprehension

One important finding for the listening comprehension task came from the cross-linguistic comparison, where the scores of the Gambian children were found to be higher than those of the Senegalese children. This result is striking for two reasons: first, the SES of the Gambian children was lower than the SES of the Senegalese children and it is very often assumed that the linguistic level depends on SES (e.g., PISA, 2000). Second, there were fewer Gambian children who reported that they spoke English at home (14 children, out of 1,200) than Senegalese children who reported that they spoke French at home (38 out of 502).

One possible explanation is that the English test was easier than the French test. Indeed, there were only three questions in the English test whereas there were five questions in the French test, two with yes/no responses in both cases. To determine whether this interpretation is correct, it is necessary to examine the responses to each of the questions. Unfortunately, only the total number of correct responses was included in the data files. For future EGRA applications, the scores for each response should be entered and yes/no questions should be avoided.

Another explanation might be that the Gambian children's English vocabulary level was higher than that of the Senegalese children's French vocabulary. To determine whether this alternative interpretation is correct, it is necessary to know the vocabulary level of the children in the language they are learning to read. For future EGRA applications, an assessment of the children's vocabulary level in the language they are learning to read could be considered. A short version of the commercially available Peabody Picture Vocabulary Test (PPVT) could be included, for instance. A simplest method would be to check the children's knowledge of the names of more and less known parts of their body (by asking them to point to body parts such as their *nose*, their *mouth*, their *eyes*, their

elbow, their *chin*, and their *hip*) and of more and less known objects from the school environment (by asking them to show objects such as a *table*, a *chair*, and a *bench*), and their understanding of spatial terms (by asking them, for example, to put a pencil *under* and *above* a sheet of paper, and then *in front* of him/her and *behind* him/her).

In addition, as indicated above, a procedure similar to that proposed to assess reading comprehension should be used to assess listening comprehension. Compared to the procedure used in the present Gambian study, this new procedure would lessen the difference between the two tasks. Indeed, most people can easily speak at a rate of at least 200 words per minute and, for the listening comprehension task, enumerators were instructed to read at a rate of approximately 120 words per minute, which is much faster than the rate at which most of the Gambian participants performed (5 words per minute; and 16 or 18 words per minute for the Senegalese participants who were learning to read in Wolof or in French). Even for readers able to read less than 60 words in 3 minutes, it is easier to answer questions after the examiner reads aloud the 60-word text than after the participants read aloud the same text by themselves, because it is less difficult to recall a specific piece of information provided in the beginning of a text after half a minute (the approximate time it would take for the test enumerator to read the passage aloud) than after, for instance, 3 minutes (the time it might take for a slow reader to finish the text), due to memory capacity and the time lag between reading the text and answering the comprehension questions.

Spelling skills

Very large floor effects were observed in the spelling task assessing whether the children know the correct orthography of high frequency words presented in a sentence (95 percent and 80 percent for first and third graders, respectively). In spite of these floor effects, the correlations between that task and the word and pseudoword reading tasks were strong (.66 to .77) and higher than those of the task that assessed whether the children were able to correctly use some other specific written conventions (.50 to .59). In addition, according to the regression analysis, only the results obtained in the first task predict both word reading and reading comprehension.

The second task could thus be deleted from the protocol for future EGRA applications, and the first task could be simplified. Only four isolated words would need to be used, including two very short, high frequency words with regular GPC, to allow very beginning or very poor spellers to spell at least some items.

Phonemic awareness

The results of the correlations indicated that the relationship between phonemic awareness skills and reading skills was, in most cases, not significant. Furthermore, the scores obtained in those tasks were not found to contribute to the explanation of the variance in word reading, or in reading comprehension.

In addition, the examination of the effect of the home language on phonemic awareness suggests that these tasks, and especially the phoneme identification task, are very sensitive to the mastering of the phonetic properties of the phonemes of a specific language. Indeed, among the eight differences that were significant according to t-tests, five dealt with the phoneme identification task and three with the phoneme counting task. Even after having taken SES into account, five differences remained significant, four for the phoneme identification task (the scores of the Pular and of the Jola speakers were lower than those of both the Mandinka speakers and the speakers of an unknown language) and one for the phoneme counting task (the scores for the Jola speakers were lower than those of the Mandinka speakers).

These results indicate that the phoneme identification task was more sensitive than the other tasks to the specificity of the spoken language. In addition, that task poses a very serious problem: indeed, it is not possible to sound out a consonant without a vowel. However, to use GPC, it is necessary to be able to discriminate the phonemes. It is therefore important to assess the children's phonemic awareness level. One possibility could be to replace the phoneme identification task with a classical phonemic discrimination task. The aim of this type of task is to assess the ability to discriminate spoken words for which the differences in pronunciation are very slight, such as the difference between "path" and "bath," "tear" and "deer," and "coat" and "goat" (children should be required to show, among three pictures, the picture that corresponds to "path" and the picture that corresponds to "bath," for instance).

A counting task will allow the assessment of the participant's ability to segment the speech stream into phonemes, a skill also required to use GPC, which differs from the ability to discriminate among different phonemes. For future EGRA applications, it would be advisable to use a counting task including only simple phonemes, without complex phonemes such as diphthongs (e.g., the vocalic sounds of the words "late," and "boy"), glides (e.g., the 'w' and the 'y' sounds of the words "we" and "yes"), and affricates (e.g., the "ch" of the word "chair"), and without phonemes that are not clearly sounded out (e.g. the postvocalic 'r' of "birds").

However, theses tasks (especially the phonemic discrimination task) were very challenging for the enumerators. Consequently, these two tasks could be replaced by another task allowing the assessment of both phoneme discrimination and phoneme segmentation skills: a spelling task involving simple pseudowords. Children should be required to spell two consonant-vowel (CV) and two consonant-vowel-consonant (CVC) pseudowords. As for the phonemic counting task, it would be advisable to avoid phonemes that are not clearly sounded out, and complex phonemes such as diphthongs, glides, and affricates. The score would be based on the number of consonants and vowels within each pseudoword that are correctly spelled.

Other points

There were differences in the total number of correct responses between some tasks (phoneme counting and phoneme identification; reading and listening comprehension). In order to facilitate the comparisons within a same language, as well as between different languages, it would be important to use a similar number of items across languages for future EGRA applications.

Finally, reading difficulties are very often assumed to be due to poor phonological short-term memory (memory related to the ability to relate letters to sounds). A short assessment of phonological memory should thus be included in the future EGRA protocol. The task could involve the repetition of two- to six-syllable pseudowords (two for each length). As for the phonemic awareness tasks, it would be advisable to avoid phonemes that are not clearly sounded out, and complex phonemes such as diphthongs, glides, and affricates.

4.2. Control variables: SES, cultural environments, and gender

The SES was weakly correlated with the results of the experimental tasks: only some of these correlations were significant, but low. Furthermore, the regression analysis indicated that SES did not explain word reading skills. These results are striking given that SES was found to have an effect on all the results according to the comparison between the children with low versus high SES scores.

Alternatively, the effect of the linguistic factors (home language and language in which the children were learning to read) appears stronger than that of SES. For instance, the language spoken at home by the Gambian children was found to have an effect on their results, even when SES differences between the groups were taken into account. In particular, most of the differences observed between the linguistic groups in the two phonemic awareness tasks remained significant, especially those for the phoneme identification task (the scores of the Jola and of the Pular speakers were lower that those of both the Mandinka speakers and the speakers of an unknown language). Only three other differences remained significant after having taken SES into account: in word-in-context reading, the English speakers surpassed the Pular speakers, and the Wolof speakers surpassed the Mandinka speakers; and in the first spelling task, the Wolof speakers surpassed the Mandinka speakers. In all the other tasks, none of the differences that were significant according to t-tests remained significant after having taken into account SES. The analysis of the effect of the home language calls for a very careful consideration of the data concerning this issue in the student survey. In particular, the home language should be coded and reported in the final database. The use of the category “other” is not suitable without coding and entering the language name.

The examination of the relationship between reading skills and the language in which the children were learning to read also suggests that the classical sociological explanation of reading difficulties is not sufficient. Indeed, even when controlling for SES, the scores of

the Gambian children who were learning to read in a language with a deep orthography (English) were lower those of the Senegalese children who were learning in a shallower orthography (French) in all the assessments involving written language processing. These results reproduced those found in the literature showing that in languages with transparent GPC (shallow orthographies, such as Spanish, and even French), the learning progression is more rapid than in languages with a deep orthography (e.g., English) where the correspondences between graphemes and phonemes are more complex (for reviews, see Sprenger-Charolles et al., 2006, and Ziegler and Goswami, 2005).

The effect of the grade level of the children on their results also suggests that the opacity of the orthography exerts an effect on reading acquisition. Indeed, the Gambian second graders who were learning to read in English did not correctly read a higher number of pseudowords in 1 minute than the first graders; and they did not obtain a higher score in the first spelling task. Alternatively, all the differences between the first and the second graders were significant for the children that were learning to read in French or in Wolof (see Sprenger-Charolles, 2007). These data, even if not longitudinal, suggest that the improvement of reading skills with grade level is less significant for the children who were learning to read in a language with a deep orthography (English) than for those who were learning to read in a language with a shallower orthography (French).

One result in particular was found to be inconsistent with what is expected in this type of assessment (e.g., PISA, 2000): the superiority of the boys over girls, especially in the tasks involving reading and listening comprehension. Another result is more in line with the expectations: attendance at a nursery school was found to have a systematic positive effect on the results. Note that attendance at the nursery school was not found to have a positive effect on the results of the Senegalese children, whatever their group (French or Wolof). Finally, the results of the comparison between the children with at least one literate parent and the children with nonliterate parents indicated that, when SES was taken into account, there were only few significant differences between the groups: one difference in the letter-name task and, surprisingly, two differences in tasks involving spoken language processing (listening comprehension and phoneme identification).

5. IMPLICATIONS FOR EDUCATION AND RESEARCH

In light of the current results and their implications for educational policy, the EGRA protocol is largely more relevant than those used in most international assessments (PISA-OECD and PIRLS-IEA, for example). Indeed, these assessments mostly take into account the understanding of different types of written texts, mostly for children in fourth grade and beyond. However, some improvements are still to be made to the EGRA protocol, especially in the context of an assessment taking into account children who are learning to read in languages which, in most cases, are not their mother tongue.

5.1. Summary of Implications for future EGRA applications

In the student survey, the analysis of the relationship between home language and reading skills calls for a very careful consideration of the data concerning this issue. In particular, the home language should be coded and reported in the final database. The use of the category "other" is not suitable without coding and entering the language name. In addition, Arabic, which was never listed as a language spoken at home, should be added to the list, especially in countries where Islamic culture is prevalent. Finally, as was done in the Gambian survey, but not in the Senegalese surveys, it would be preferable to systematically ask the child what language his or her parents read (when they were said to be able to read).

In the different experimental tasks, for the reasons explained in the previous section, five tasks could be deleted (the two pre-reading tasks, the second spelling task, and the two phonemic awareness tasks). Due to large floor effects, other tasks should be drastically simplified (the three 1-minute reading tasks and the first spelling task). It may also be necessary to revise the reading comprehension task in order to lessen floor effects and to render that task as similar as possible to the listening comprehension task, especially as regards the memory load. In addition, the pre-reading task assessing letter-name knowledge could be replaced by a task assessing the ability to differentiate true letters from reversed letters and nonalphabetic symbols, and the two phonemic awareness tasks could be replaced by a pseudoword spelling task. Finally, two tasks could be added. First, it might be necessary to develop a test to assess the vocabulary level of the children in the language in which they were learning to read. An assessment of the children's phonological short-term memory could also be included in future EGRA applications.

Given the effect on the children's responses of both their linguistic environment and the language in which they are learning to read, it is also necessary to very carefully select the items used in the tasks requiring language processing in order to avoid as much as possible some biases due to linguistic differences. On the one hand, it would be better to avoid the use of some language-specific features. On the other hand, it would be important to build up tasks of a similar level of difficulty in each of the languages in which EGRA will be developed. For reading tasks, these types of control, which should be thoroughly realized in future EGRA applications, would be more difficult to design for the English experimental tasks than for the tasks developed in French or in Spanish, for instance, because GPCs are less consistent in English than in Spanish and French. In addition, these controls can be achieved only if there is descriptive data and statistics on the characteristics of the orthographic and phonological system of the languages studied. There is thus a need to provide summaries of these characteristics to local teams.

5.2. Implications for educational policy

The present report once again highlights the fact that reading acquisition depends on the degree of transparency of GPC. Indeed, as suggested by the reading results of the Gambian children, learning to read is very difficult in English, a language where the

correspondences between grapheme and phoneme are not very consistent. In addition, this difficulty is worsened by the fact that the English language is not the children's mother tongue.

It would thus be important to clearly explain the main characteristics of the English orthography to teachers, and to develop reading materials including only very frequent and simple words with, as much as possible, regular GPC. To master GPC, it is also necessary to differentiate the phoneme of the language in which the children are learning to read. Therefore, teachers should be trained in both the English phonological system and the phonological system of the children's home language. This knowledge would help them be aware of the interferences between different phonological systems that could impede not only the acquisition of a new spoken language, but also reading acquisition in that language (see Labov, 1972 and 1995; see also Linan-Thompson and Vaughn, 2007).

In addition, we know that to be efficient, remedial training has to be conducted as early as possible (see Ehri, 2001). It should thus necessary to assess the main basic reading and reading-related skills (those that should be included in the final EGRA protocol), as early as possible and at the latest, by the end of the first year of primary school. These assessments would help teachers identify children with severe reading difficulties, for which remedial programs could be developed.

A last important point to highlight is that nursery school attendance was found to have a very positive effect on the subsequent reading capacity of the children, a result which was not noticed in the study with Senegalese children. This finding provides to the Ministry of Education of The Gambia some justification for further investments at this level.

6. REFERENCES

Abadzi, H. (2006). *Efficient Learning for the Poor.* Washington, DC, The World Bank.

Ehri, L.C., S. R.Nunes, S.A. Stahl, & D.M. Willows (2001a). "Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta- analysis." *Review of Educational Research* 71(3):393–447.

Ehri, L.C., S.R. Nunes, D.M. Willows, B.V. Schuster, Z. Yaghoub-Zadeh, and T. Shanahan (2001b). "Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis." *Reading Research Quarterly* 36(3):250–287.

Elley, W. (1992). *How in the world do students read*? Edited by the International Association for the Evaluation of Educational Achievement (IEA).

Labov, W. (1972). *Language in the Inner City.* Philadelphia: U. of Pennsylvania Press (Traduction française: *Le Parler Ordinaire.* Paris: Editions de Minuit).

Labov, W. (1995). “Can reading failure be reversed: A linguistic approach to the question.” In V. Gadsden and D. Wagner (Eds.), *Literacy Among African-American Youth: Issues in Learning, Teaching and Schooling* (pp. 39–68). Cresskill, NJ: Hampton Press.

Linan-Thompson, S., and S. Vaughn (2007). *Research-Based Methods of Reading Instruction for English Language Learners: Grades K–4*. Alexandria, VA, Association for Supervision and Curriculum Development.

Peereman, R., and A. Content (1998). *Quantitative analysis of orthography to phonology mapping in English and French (on-line)*. http://homepages.vub.ac.be/~acontent/OPMapping.html.

Peereman, R., B. Lété, and L. Sprenger-Charolles (2007). “Manulex-Infra: Distributional Characteristics of Infra-Lexical and Lexical Units in Child-Directed Written Material.” *Behavioral Research Methods, Instruments and Computers* 39:593–603.

Perfetti, C. A. (1985). *Reading ability*. New-York: Oxford University Press.

PIRLS (2003). *International Report: IEA’s Study of Reading Literacy Achievement in Primary Schools.* Chestnut Hill, MA: Boston College.

PISA (2000). *Knowledges and skills for life. First results from PISA 2000*. OECD, Paris, 2000.

RTI International (2007). *Early Grade Reading Assessment: Protocol*. RTI International for USAID. Durham, NC. Web site: www.eddataglobal.org.

Seymour, P. H. K. et al. (2003). “Foundation literacy acquisition in European orthographies.” *British Journal of Psychology* 94:143–174.

Share D.L. (1995). “Phonological recoding and self-teaching: Sine qua non of reading acquisition.” *Cognition* 55:151–218.

Sprenger-Charolles, L. (2003). “Reading acquisition: Cross linguistic data.” In T. Nunes, & P. Bryant (Eds.), *Handbook of children’s literacy* (pp.43-66). Dordrecht: Kluwer.

Sprenger-Charolles, L. (2007). *Evaluation des compétences en lecture (en francais et en wolof) d’enfants sénégalais des trois premiers grades du primaire: Rapport pour la Banque Mondiale*. Web site : www.eddataglobal.org (also in English).

Sprenger-Charolles, L., P. Colé, and W. Serniclaes (2006). *Reading acquisition and Developmental dyslexia*. Psychology Press (Developmental essay series).

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.

Wagner, R.K., J.K. Torgesen, C.A. Rashotte, S.A. Hecht, T.A. Barker, S.R. Burgess, J. Donahue, and T. Garon (1997). “Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: A five year longitudinal study.” *Developmental Psychology* 33:468–479.

Zieger, J., and U. Goswami (2005). “Reading acquisition, developmental dyslexia and skilled reading accross languages: A psycholinguistic grain size theory.” *Psychological Bulletin* 13(1):3–29.