

The Kom Experimental Mother Tongue Education Project Report for 2011

Stephen L. Walter and Kain Godfrey Chuo

Executive Summary

For four years, an experimental program in multilingual education has been functioning in Boyo Division in the North West Region of Cameroon. In this program, 12 experimental schools use the local language—Kom—as a medium of instruction while another 12 schools continue to use English as the language of instruction. At the end of year 4, standardized testing was once again completed for children in both comparison and experimental schools in Classes 2 to 4. The findings are presented in the various sections of this report.

Results for Classes 2 and 3 closely paralleled those from previous years of testing. In Class 2, children in the experimental showed performance advantages of approximately 150 percent on all subjects other than oral English.

In Class 3, when tested in Kom, children continued to show an advantage of approximately 130 percent over those in the standard program. However, when tested in English, this advantage slipped to about 80 percent—still quite significant.

Of special interest are the findings from Class 4 in that all instruction in all schools for both programs reverted to the standard English instructional model. Three key findings characterize the data from Class 4. First, the gap between the comparison and experimental children decreased further to about 45 percent. Second, all children performed surprisingly poorly on the assessment given including on content drawn from the curricula of Classes 2 and 3. Third, children coming from the standard English-medium schools are STILL performing, as a group, at a chance level (about 25%) indicating continued inability to read AND general low performance on tasks (such as math operations) which do not require reading ability.

The discussion of the Class 4 results touches both on the predictions which come from multilingual education theory as well as some of the probable reasons for the reduced performance of those coming from the experimental program. These include minimal ability in English, inadequate time on task, instructional issues, etc.

Introduction

The year 2011 marks the fourth year of the operation of the experiment in mother tongue based multilingual education in the Boyo Division of North West Province, Cameroon. In the design, 12 experimental schools have been selected as sites for the implementation of a curriculum using the mother tongue (Kom) as the primary language of instruction. Twelve (12) matching schools were selected as control schools. These 12 schools implement the standard model of instruction in which English is used as the medium of instruction for all subjects and classes¹. The schools participating in the experimental program were matched by type (private or public), location (remote village, small town), and size (smaller or larger) with schools in the standard program. There was a certain amount of selection bias in the identification of schools to be included in the experimental program in that poorer performing schools were somewhat more likely to be included to reduce or eliminate any concerns about hand-picking better schools to create a more favorable outcome.

The intervention

The primary element of the intervention is that of language of instruction. We consider this variable foundational in light of (a) the research findings of Thomas and Collier (1997, 2002), and (b) the fact that all developed countries consider it a given that the language of instruction in their own schools should be the primary language of the country whether English, Spanish, German, French, Danish, Russian, or whatever. The issue of language of instruction becomes controversial in such countries only when dealing with minority populations—be they indigenous or the product of immigration.

The primary intervention, language of instruction, was accompanied by several supporting elements worthy of investigation in their own right, though we have lacked the resources (and local permission) to carefully manipulate and measure the impact of these variables. Nevertheless, they are reported here for the sake of completeness. (1) Teachers received three weeks of training in how to teach reading and writing in their own language since they had neither previous experience nor training in using their own language for literacy. This variable is probably offset by the fact that teachers in the standard program receive regular in-services as well. (2) The teachers and students were provided with textbooks. Textbooks exist as well for children in the standard program but buying them is at the discretion of parents and many fail to do so. (3) Since teaching literacy in Kom was an additional subject in the curriculum for those in the experimental program, some adjustments had to be made in the weekly timetable to make time for this additional subject. This was accomplished primarily by bundling some of the ‘more peripheral subjects’ into a Kom-language reader to be used both for reading practice and to teach the content of the more peripheral subjects. (4) Because teachers were implementing instructional methods new to them (and the community) at least two additional supervisory visits were arranged to provide support and to do on-the-spot problem solving.

At this point, we are treating the intervention as being unitary in nature though we know it is very possible that some of the sub-elements may have had a measurable impact on outcomes. In fact, some education officials in the area have raised the question of whether having textbooks in the English-medium schools might impact outcomes in those schools as well.

¹ In this part of Cameroon, the term “class” is preferred to the term “grade” in reference to the yearly levels of students in basic primary education.

Past findings

The results during the first three years of the Kom Education Project (KEP) showed a marked—even strong advantage in educational achievement for the children in the experimental program. Measured improvements in all learning outcomes except oral English were dramatic—as much as 600 percent in some cases with 100 percent gains being typical. In the case of oral English, the children in the experimental program also outperformed the children in the standard English-medium program but by much smaller amounts ranging from 5 to 35 percent depending on the class. Even this small advantage was contrary to prevailing public expectation that the best way for children to learn and improve their English was to be immersed in an all-English instructional environment.

Previous reports for the years 2008, 2009, and 2010 contain the details of the findings for those years. This year—year 4 of the experimental project—brings us to a new phase of research. Specifically, in Year 4, children who had been in classrooms receiving mother tongue instruction moved also into an all-English instructional environment. The consequences of this transition will be a major theme of this report. We are aware of no research which predicts or documents with any kind of specificity the likely educational consequences of this kind of transition in a developing country after having spent three years in a mother tongue model. Thus, we expect the details of this year's testing to be of great interest especially among those committed to the MLE model and, more generally, to the larger international education community.

The best prevailing model of what to expect in this circumstance is that of Thomas and Collier (1997, 2002) which makes a number of relevant statements. First, their model states quite categorically that so-called early-exit mother tongue instructional programs—of which the KEP program is an instance—will produce weaker educational outcomes than late-exit programs, those providing instructional support in the mother tongue for up to six years. Secondly, the model indicates that early-exit programs will still produce a certain amount of enduring improvement in educational outcomes and is thus still an improvement over immersion programs (such as the standard model in Cameroon of educating children in English or French regardless of whether they speak either of these languages). Thirdly, the model suggests that after children leave early-exit programs and move into a standard language program, they will tend to stagnate and lose ground to mainstream children (children being educated entirely in their first language).²

Since the model put forth by Thomas and Collier is based entirely on research done in the US, they made no claims about its aptness for describing educational experience in developing countries. It is for this reason that we believe readers will be very interested in the findings of this report especially those pertaining to Class 4.

The analysis of Class 4 data will present summary results of the testing of learning outcomes for the year and will also include some preliminary thinking and analysis as to the reasons for these results. This reflection will include a review of results in the light of the theoretical predictions coming from the best models on multilingual education. In this process, attention will be given

² It is important to note at this point that in countries like Cameroon, few or relatively few children receive their entire education in their first language. In Cameroon's case, that is because few children speak standard English or standard French as their first language.

to distinctive features of basic education in the project area as these contrast significantly from the educational practice upon which the Thomas and Collier model is based.

Some notes on the educational assessments which were carried out

To assist first time readers, some background information is provided here on the nature of the assessments and assessment process used in the project area. As indicated above, the experimental project consists of 12 experimental schools in which the local language (Kom) is used as the language of instruction in Classes 1-3 for all subjects but English (as a subject). These 12 experimental schools have been paired with 12 control schools in which all instruction is in English, the standard educational model of the area (and the province).

Each year, standard tests (developed by the research team) are administered to children in both control and experimental schools **in the language of instruction**. This means, obviously, that children in the control or comparison schools are tested with instruments rendered in English and those in the experimental schools are tested with instruments rendered in Kom. Testing of proficiency in oral English was done in English.

Content for all assessment instruments came directly from the curriculum materials used for instructional purposes. In the case of math, the instruments were identical with the exception of the language of presentation. In the case of reading and language arts, the instruments can best be described as 'highly comparable.' That is, if a sample text for reading comprehension was taken from Lesson 3 in the standard program, a sample text for reading comprehension in Kom was also taken from Lesson 3 of the instructional textbook developed for that program. If the English instrument contained 5 words to test word recognition, the Kom-medium instrument also contained 5 words drawn from parallel lessons and of similar length and perceived difficulty. The oral English assessment was identical for all students at a given level.

The assessment designs for Classes 1 and 2 were identical apart from the content of the instruments which were adjusted according to the content of the respective curricula. The assessment design for Class 3 included an innovation in that ALL children in both programs took the same reading test in English. Those in the experimental program also took an additional reading/language arts test in Kom since that is their language of instruction for reading. This modification in the assessment design was motivated by a desire to better measure the nature of the reading skills being developed in the experimental schools with respect both to English AND to Kom. Consequently, no assessment of oral English was given to Class 3 students.

The assessment for Class 4 was quite straightforward; identical tests in English given to all students coming from both experimental and control schools. No test of oral English was given in Class 4 though the results for this year suggest there may be a need for a more thorough assessment of proficiency in English especially at the basic level of vocabulary development.

The following table summarizes the assessment design for all classes in the experimental program.

Table 1. Description of assessment design being used in the KEP project.

	Standard Program		Experimental Program	
	Language of instruction	Language of testing	Language of instruction	Language of testing
Class 1 Language arts/reading Math English	English English English	English English English	Kom Kom English	Kom Kom English
Class 2 Language arts/reading Math English	English English English	English English English	Kom Kom English	Kom Kom English
Class 3 Language arts/reading Math	English English	English English	Kom Kom	Kom and English Kom
Class 4 Language arts/reading Math	English English	English English	English English	English English

All testing was done by non-school personnel to increase the level of reliability of assessment results.

Some comments on presentation of results

In general the presentation of results will proceed from Class 1 to Class 4. The Class 1 results are repeated from testing done in 2010 since testing of Class 1 was not done in 2011.

Within each Class, the presentation of results will include the following:

Comparative results across years. Listing of results for all previous years for which data is available. These comparisons across time serve primarily to monitor changes between years as well as the relative consistency or reliability of the test instruments.

A more detailed presentation of results for the current year. For each class, the assessment instrument will be broken down into the identifiable sub-skills which were included on the instrument of assessment. This presentation is useful to determine more specifically where students were strong or weak. Such information could be useful to inspectors and headmasters to help them determine specific instructional topics not receiving adequate attention (or too much attention).

Presentation of results of any independent variables such as gender or age which might have had an impact on results. There is interest in many circles on the impact of a range of variables which may influence educational delivery in developing countries. Not all such variables have been tracked in the current research program. Those which have been and which seem to affect outcomes will be reported.

One or more performance profiles. Technically, a performance profile is a distribution. For the purposes of this paper, we will refer to these distributions as performance profiles since that

term seems to better communicate the intent of these graphs. For the most part, the performance profiles will be presented in terms of performance ranges or categories on a scale from 0 to 100 since all of the assessments have been scaled accordingly. Most of these distributions or profiles will be divided into ten ranges (called bins by most researchers) for uniformity of presentation and interpretation. The performance profiles provide a visually effective way of comparing the performance of two (or more) groups or subpopulations. Normally, the two groups compared will be those coming from English-medium versus Kom-medium schools.

Interpretive comments about the findings for each class. At the end of each section presenting the results for a given class will be a section which explores both unusual findings as well as the conditions which may account for such findings. Occasionally, the interpretive sections will venture into the realm of making recommendations for steps which could be taken to improve educational outcomes.

Summary of test results from 2011

Table 2 summarizes the major findings for all three classes included in the 2011 assessment.

Table 2. Performance on standardized tests administered to both Standard and Experimental (KEP) schools in 2011.

Test Component	Class 2			Class 3			Class 4		
	Standard	KEP	Gain ³ (%)	Standard	KEP	Gain (%)	Standard	KEP	Gain (%)
Language Arts	22.0	61.0	177.3	24.2	41.1	69.8	28.1	38.7	37.7
Math	21.5	54.3	152.6	21.0	41.7	98.6	27.3	40.0	46.5
Oral English	52.9	61.3	15.0	*	*		*	*	*
Overall Test	34.4	59.2	72.1	23.0	41.3	79.6	27.6	39.5	43.1

Several preliminary interpretive notes are in order. First, apart from the very basic oral English assessment, children in the standard program consistently score in the 20-30 percent range. Since the assessment instruments used primarily a multiple choice format, this range can be considered equivalent to random guessing. While an examination of individual test scores indicates that some children in the standard program are clearly learning, the group mean scores persist in the range one associates with random guessing.

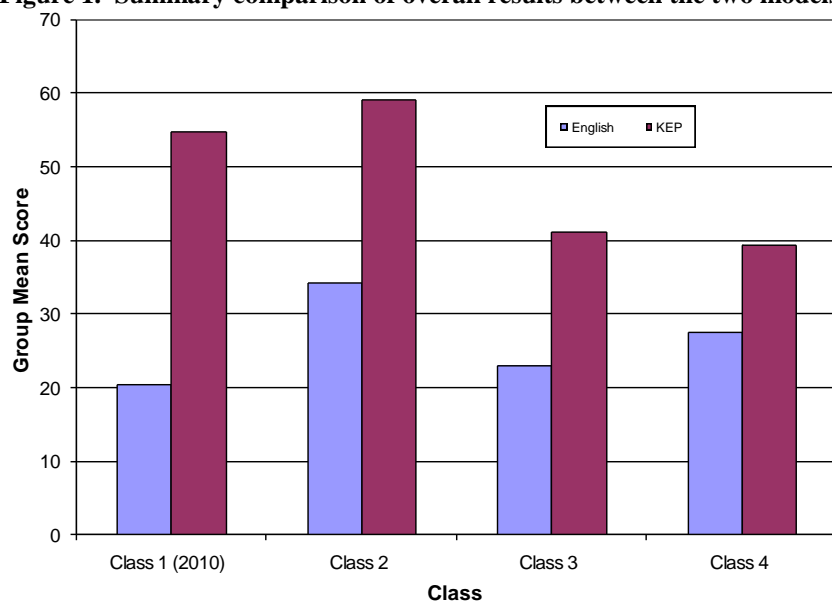
Second, we note what appears to be a significant drop in performance in the KEP program between Class 2 and the next two classes. The most likely reason for this *apparent* drop is the fact that the group mean scores reported in Table 2 for Classes 3 and 4 reflect performance on an English-medium assessment instrument rather than a Kom-medium instrument which is the language of instruction in Class 3. As will be demonstrated in a subsequent section of the paper, when children in the experimental program in Class 3 were tested using Kom-content reading instruments, they scored in the 65 percent range.

³ Many of the tables included in this report will contain a column called 'Gain (5).' The numbers in this column express the extent of the improvement or gain in efficiency for those in the KEP program relative to those in the standard program which, in this case, is treated as the 'normal' or default program.

Probably the most intriguing feature of Table 2 is the general similarity between mean scores in Class 3 and Class 4 (when Class 3 results are based on English-language assessments). Children in/from the KEP program are consistently scoring around 40 percent. Children in the standard program averaged 23 percent in Class 3 with an up-tick to 27.6 percent in Class 4. Subsequent sections will take a closer look at the possible significance of this finding.

Finally, note that no test of proficiency in oral English was given to children in Classes 3 and 4. The reason for this change was that the entire test for children in Classes 3 and 4 was in English so a separate oral test was not deemed necessary. This decision may well be reconsidered in future assessments.

Figure 1. Summary comparison of overall results between the two models in all four classes for 2011.



Interpretation of the results shown in Figure 1 must be done with caution. First, we note what appears to be strong performance of the KEP children in Classes 1 and 2. We believe this to reflect what we will term “instructional adequacy.” Both children and teachers⁴ speak Kom as a first language so there is a high level of effective communication in the classroom. The classroom atmosphere is dynamic and electric as children compete to answer questions and demonstrate their abilities. Conversely, the performance of the children in the English-medium schools is effectively lower than pictured having been elevated by scores on the oral English assessment substantially above those of reading and math. Children in these classrooms are quiet, hesitant, and unwilling to participate actively. The fact that many of the teachers speak English poorly further detracts from instructional adequacy.

In Classes 3 and 4 we observe a narrowing of the differential between the two groups. The children in or from the experimental program lose—from a testing point of view—the advantage of being tested in their language of instruction. Children in the English program reflect performance levels similar to those observed in Classes 1 and 2 though obviously the tests for

⁴ A few of the teachers are not actually native speakers of Kom but are highly fluent having lived in the area for many years.

Classes 3 and 4 are more difficult than the earlier tests. The significance of these comparisons will be investigated more fully when we take a closer look at the results from Classes 3 and 4.

Results for Class 1

(This section is largely retained from the 2010 report for the sake of completeness. Testing in 2011 did not include Class 1 since we already had three years of results and these had been quite consistent across the years.)

Results of testing in Class 1 across multiple years

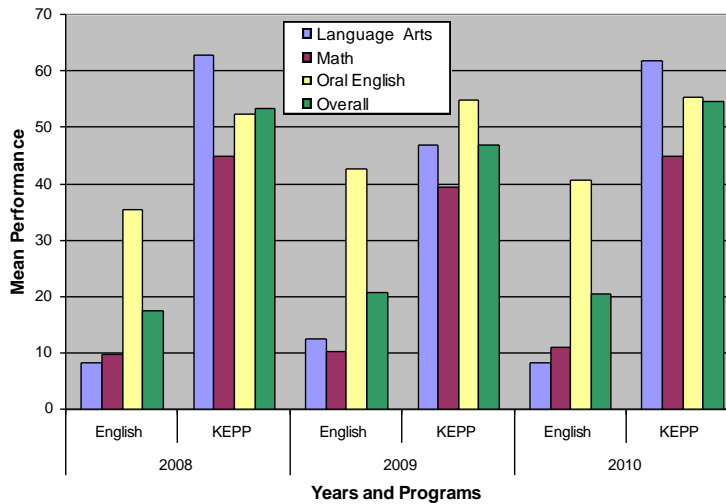
Now that the program has operated for four years, we are able to compare results for a single class across this span of time. This sort of comparison serves several possible purposes. First, we can gather evidence as to whether the testing process has been stable through time. Second, we can look for evidence as to whether the instructional process has changed—for better or for worse—during the time of the experimental program. Third, the increased mass of data serves to further strengthen the effect of the key variable of interest in the program—the use of the first language as a language of instruction.

Table 3. Class 1 results for three years (2008-2010).

	2008			2009			2010		
	English	KEP	Gain (%)	English	KEP	Gain (%)	English	KEP	Gain (%)
Language Arts	8.4	62.8	648	12.7	47.0	270	8.5	61.8	627
Math	9.9	44.9	354	10.4	39.5	280	11.1	44.9	305
Written test	9.1	53.8	491	11.7	43.7	274	9.2	54.3	490
Oral English	35.4	52.5	48	42.8	55.0	29	40.8	55.5	36
Overall	17.6	53.4	203	20.9	47.1	125	20.6	54.8	166

The test results for both programs have been quite stable from one year to the next. There was something of a drop-off in the performance of the KEP program in 2009 which was traceable to two of twelve schools having major staffing problems (loss of teachers without replacement by a similarly qualified teacher for sustained periods of time). In comparing the three years, it would appear that 2009 represents the most distinctive of the years with the English-medium program scoring slightly higher than in the other years and the Kom-medium program scoring somewhat lower than in the other two years. The pattern of results over a three year period of time strengthen the claim that the observed differences between the two models are, in fact, due primarily to model-specific effects, i.e., the use (or not) of the mother tongue as a medium of instruction. Figure 2 provides a graphical presentation of the same data on the major measures of assessment.

Figure 2. Comparison of Class 1 students by year and by program.



Several features of differential performance are obvious from the graph. First, the English-medium program is performing at a level well below half of that of the Kom-medium program overall. Second, the two programs are not far apart on the measure of oral English though the Kom-medium program is consistently the higher of the two. Third, the specific measure on which there is the greatest difference of performance is that of language arts. The Kom-medium program in Class 1 has generally outperformed the English-medium program on this measure by a factor of 7 to 1. In as much as reading is the most fundamental objective of Class 1, a huge advantage accrues to children in the Kom-medium program.

Breakdown of results of the testing from 2010

The following table contains the results of the testing done in Class 1 broken out by all major content areas of the testing.

Table 4. Detailed listing of test results from Class 1 by program.

	English-medium	Kom-medium (KEP)	Statistics
Language Arts	8.5	61.8	P = 0.000
Word recognition	5.4	66.3	
Grammar	12.0	57.7	
Reading comprehension	8.8	60.3	
Math	11.1	44.9	
Counting	18.6	54.9	
Place value	9.5	40.5	
Addition	11.6	48.4	
Subtraction	9.5	42.6	
Written Test	9.2	54.3	
Oral English	40.8	55.5	
Social communication	72.3	82.7	
Following instructions	55.2	76.3	
Free response	15.8	27.9	
Overall	20.6	54.8	

The data indicate a very large testing/learning advantage (54.8 percent vs. 20.6 percent) for those in the Kom-medium program. The differential advantage is greatest for the written part of the test and much less for the Oral English section of the test (55.5 percent vs. 40.8 percent). The latter finding is somewhat surprising given the common assumption that children in an English-

medium program 'should' learn more English than those in a program using the mother tongue as a language of instruction with instruction in English limited to second language instruction (English as a subject). The evidence, however, does not support this 'common' assumption.

Not too surprisingly, the greatest differences are in the area of language arts. One of the major objectives of Class 1 is to teach children to read. Since the children in the Kom-medium program are learning to read in a language they know, they showed good progress in doing so. On the contrary, the children in the English-medium program are blocked from learning to read because they neither understand the instruction of the teacher nor know the language in which they are supposed to be learning to read. Therefore, they make little progress as reflected by the data.

Gender

Girls in Class 1 show a small but consistent advantage over males in most measures tested. This is true both in the English and Kom-medium programs with the advantage being greater in the English-medium program. In some cases, the difference approaches what would be characterized as statistical significance though we have little basis for explaining why this might be the case. Interestingly, some preliminary and very experimental research on native academic capability showed a slight advantage for females in general in the Kom area. At this point, we have no principled explanation for this finding though it is consistent with the educational testing being done in the experimental project.

Fulfulde speakers⁵

Since the data recorded whether students were Fulfulde speakers, we examined that as an independent variable. The obvious assumption—working hypothesis—is that these children would fare poorly in the Kom-medium classes and would, therefore, be best served by English-medium instruction. The findings for Class 1 are reported in Table 5.

Table 5. Data on the performance of Fulfulde speakers by program.

	English-medium	Kom-medium
Fulfulde speakers	17.0	55.9
Non-Fulfulde speakers	20.7	54.8

These data, taken at face value, certainly do not support the view that Fulfulde-speaking children are benefited by attending an English-medium program OR harmed by attending a Kom-medium program. Fulfulde-speaking children in the Kom-medium program outscored their peers in the English-medium program by a factor of three. At this point, however, we can't put too much weight on these findings as there were only 4 Fulfulde-speaking children in each program.

⁵ The presence of another linguistic minority living in the midst of a first linguistic minority raises a variety of questions about the efficacy of an instructional model using the language of the more numerous group as a language of instruction in schools. This issue applies in the current research setting so will be reported as a potential variable affecting educational outcomes.

Some interpretive comments

Looking at performance on the language arts task across the classes, we note a very large advantage in Class 1 (627 percent) for the experimental schools in language arts (the beginning skills of reading). The extremely low mean score for Class 1 students from English-medium schools in language arts reflects both non-learning (because of the language barrier) AND inability to deal with the assessment process (even though instructions were given in the local language AND practice items were included on the instrument).

While the extreme difference in performance on the reading assessment would be easily attributable to some of the difficulties in learning to read English (which is not easy even for native speakers), the large difference in performance in math cannot be so easily accounted for via this explanation. Few if any language skills are needed to add $2 + 2$ or to recognize that 3 is bigger than 1. However, effective communication or instruction is needed to learn these concepts. The Class 1 math data provide strong evidence that language of instruction is likely a very significant issue in learning math as well as learning how to read.

Why or how did the children in the experimental program do better on the oral English assessment than did the children in the all-English classroom? This finding is very counter-intuitive for Cameroonian parents as well as educators who have always assumed that English (or French) is best learned by maximizing the amount of time being exposed to English in the classroom. We posit several possible explanations. First, researchers (Bialystok, 1988; Galambos and Hakuta, 1988; Cummins, 1976) have suggested that meta knowledge developed about a second language via one's first language facilitates conscious assimilation of a second language. Secondly, oral English in the KEP program was taught using the methodology of second language learning. Contrastively, the learning of English in the English-medium program is a more indirect artifact of general exposure. Children are drilled on social niceties such as "How are you?" and "What is your name?" being required to memorize such expressions without internalizing them in a natural or conscious way.

Finally, we note that scoring was not exceptional even in the KEP program, especially in math. We believe this to be due to (a) less capable instruction and (b) less time on task than is needed to master the material specified in the curriculum. Both are issues widely observed in developing countries and both will be further developed in later sections of this report.

Results for Class 2

As of 2011 we now have results for three years of testing in Class 2. We were especially interested in the results for Class 2 this year in that there had been considerable disparity between the results for this class between 2009 and 2010.

Comparison of results for Class 2 across years

We now have three years of data for Class 2. While the data for each year vary somewhat, a comparison of results for each of the years allows us to better establish what is 'normal' for the two models in the research setting.

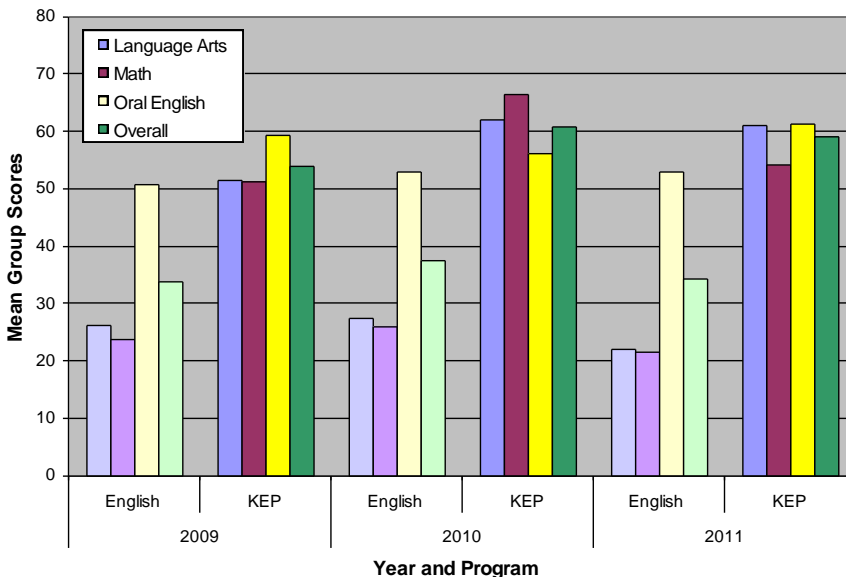
Table 6. Class 2 results for the years 2009-2011.

	2009			2010			2011		
	English	KEP	Gain (%)	English	KEP	Gain (%)	English	KEP	Gain (%)
Language Arts	26.2	51.5	98.6	27.4	62.0	126	22.0	61.0	177
Math	23.7	51.2	116	26.0	66.5	156	21.5	54.3	153
Written test	25.0	51.4	105	26.8	64.1	139	21.8	57.8	165
Oral English	50.8	59.5	17.1	53.1	56.1	5.6	52.9	61.3	15.9
Overall	33.8	54.1	60.0	37.5	60.9	62.4	34.4	59.2	72.1

Three observations seem worth making. First, Class 2 students in the English-medium program, like their peers in Class 1, have consistently scored at the level of random guessing in all areas except oral English. Second, the differential between English-medium and Kom-medium students is generally in the range of a 100 to 150 percent advantage for those in the Kom-medium program. Third, the children in the experimental program have consistently outperformed their peers in the standard program on the assessment of oral English proficiency though the advantage is rather modest. Fourth, the performance of children in the experimental program has definitely trended up since 2009. We suspect this represents “growth” in teacher’s ability to teach effectively in the experimental model after gaining some experience. It is also possible that the explanation is more pragmatic as 2009 was a problematic year for several of the schools in the experimental program for reasons having nothing to do with the program (internal ideological conflicts within the schools’ sponsor). In this case, the data from 2010 and 2011 would be considered ‘typical’ with the results from 2009 being somewhat aberrant.

Figure 3 provides a graphical presentation of the same information in Table 6 for Class 2 for the years 2009 to 2011.

Figure 3. Graphical comparison of performance of Class 2 students across 3 years (2009-2011).



In the graphic the bolder colored bars mark the performance of children in the experimental program while the sets of pastel or more lightly-colored bars denote the performance of children

in the standard English-medium program. The main observation to be made from the graph is that children in the Kom-medium program manage to do as well as the children in the English-medium program in learning English even while doing markedly better in mastering the curricular content of Class 2. The additional data from 2011 helps to establish that the results from 2010 appear to be more typical of program effect in Class 2 than those from 2009.

A more detailed examination of test results for Class 2 for 2011

Table 7 provides a more detailed examination of the results from Class 2 in 2011. In the table, sub-skills have been identified with levels of performance indicated for each sub-skill for each program.

Table 7. Comparison of results for Class 2 by program (2011).

	English-medium	Kom-medium (KEP)	Gain (%)	Statistics
Language Arts	22.0	61.0	177	P = 0.000
Grammar	22.9	56.6	147	
Reading comprehension	20.6	67.7	229	
Math	21.5	54.3	153	
Addition	25.2	58.4	132	
Subtraction	16.1	46.0	186	
Equational form	23.3	58.5	151	
Written Test	21.8	57.8	165	
Oral English	52.9	61.3	16	P = 0.350
Social communication	85.8	84.1	-.02	
Following instructions	59.2	68.9	16	P = 0.000
Free response	32.2	44.9	39	
Overall Test Result	34.4	59.2	72	

With the exception of oral English and the three components of oral English, all tested measures exhibit large differences between the two programs. As noted above, the performance of children in the standard program on the various sub-skills tested (apart from oral English) is at or below the level of random guessing.

On the other hand, performance on the measure of oral English reflects a greater degree of similarity of performance between the two programs. Not surprisingly, performance is lowest on the *Free response* item in which children are shown a stimulus such as a picture and asked to make any comments they wish in English about the picture. By 'common logic' performance in the area of oral English should show a distinct advantage for those in the English-medium program. The fact that this is not the case calls into question one of the major arguments in favor of an educational model using only a language of wider communication..

Gender

On all major measures, female students scored slightly higher than male students but the difference was far from being statistically significant.

Age

The variable of age plays an interesting and possible significant role in the Class 2 data. The age range for the entire population tested was 5 to 13. There were 40 children who would typically be considered under-age (5 and 6) in the KEP program and 58 in the standard program. Even though national guidelines state that children should not start Class 1 until the age of 6, the issue of under-age children persists in the area.

Interestingly, the variable of age behaves differently in the two programs. In the standard program 5 and 6 year olds, as a group, perform well below the level of the 7 and 8 year olds suggesting they don't belong in Class 2. On the opposite end of the age scale, over-age children (those 9 and above) don't perform any better than the 7-8 year olds.

In the KEP program, however, there is much more of a linear relationship between age and performance ($r = .32$). The 5 year-olds still do not belong in Class 2 scoring less than half the mean for normally-aged children in the KEP program. The 6 year-olds, however, scored just as well as the 7 year-olds. Furthermore, as children get older, mean scores increase substantially with the 10 year olds scoring 25 percent higher than the 7 year-olds and the 11 and 12 year olds scoring about 60 percent higher than the 7 year-olds.

A possible interpretation of these findings is that age and maturity matter little (beyond a basic age) in the standard program because the primary barrier to learning is knowledge of English which is equally absent in older as well as younger children. Conversely, in the case of the KEP program, older more mature children benefit more from instruction since language is not an impediment to their learning.

Fulfulde-speaking children

As in the case of Class 1, we examined the performance of Fulfulde-speaking children in an effort to better understand the impact of the experimental Kom-medium program on these children. The results for Class 2 are presented in Table 8.

Table 8. Performance of Fulfulde-speaking students by program.

	Kom-medium Program	English-medium program
Fulfulde-speakers	73.8	42.8
Non-Fulfulde speakers	60.7	37.3

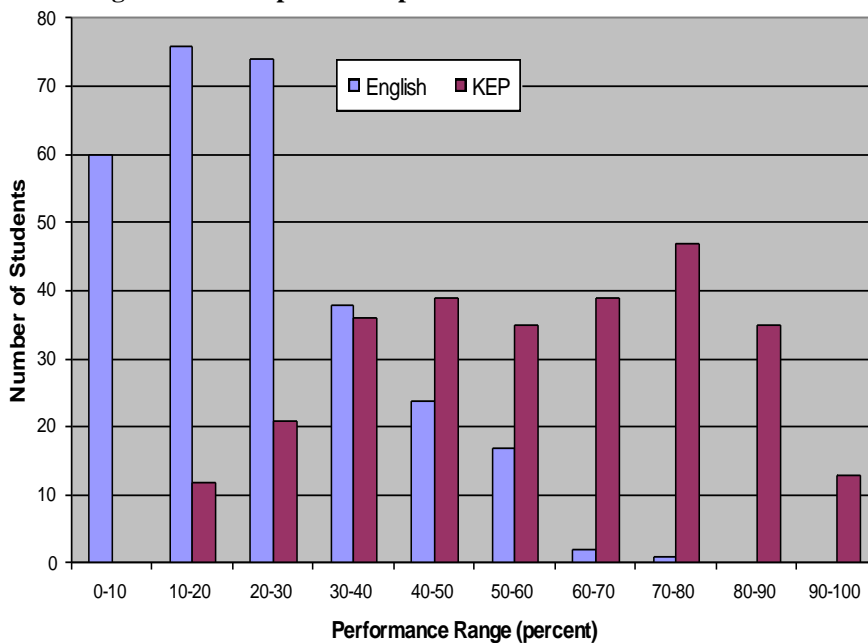
As we have consistently observed, the Fulfulde-speaking children tend to score higher than other children regardless of the program they are in. What most stands out, however, is the fact that they are not at all compromised by attending Kom-medium classes. Contrary to what some have expected, it would appear that the Fulfulde-speaking children actually benefit more than Kom-

speaking children by attending the Kom-medium schools. (Data from later classes will suggest that this result for Class 2 may be somewhat accidental.)

Performance profile

Figure 4 provides a graphical comparison of the performance of Class 2 children on the written portion of the 2011 assessment showing the sharp contrast between the experimental and control groups. The profiles are highly distinct. The profile for children in the English-medium program is bunched to the far left (technically, a skewed right distribution). Approximately 80 percent of children scored below 30 percent. Practically speaking, both the profile and an actual scan of test papers indicates that the vast majority of children are merely guessing or turning in empty test papers.

Figure 4. A comparison of performance on the written assessment for Class 2 (2011).



Meanwhile, the performance profile of children in the KEP program is approximately normally distributed with a mean of around 60 percent. This distribution or performance profile indicates good progress in mastering curricular content. The fact that there is still a sizable group of children scoring below 50 percent indicates that achievement is still undesirably low for too many children in the KEP program.

Interpretive comments on the Class 2 results

The dominant feature of the Class 2 results is the low level of performance in the standard program. With the exception of oral English, performance on every aspect of the assessment is at or below the level of random guessing. In fact, in most cases, students would have scored higher if they had simply closed their eyes and randomly selected an answer. The overwhelming weight of evidence is that, outside of a little progress in learning basic English, these students have, as a group, learned nothing during two years of school though clearly there are a few who have learned something when individual test papers are examined.

The Class 2 data provide clear evidence of the impact of the language of instruction variable. Under the educational conditions which prevail in the Kom region, children being instructed in a second language they do not speak by teachers who are also speaking that language as a second language, are learning very little. At the same time, children being instructed in a familiar language spoken by teachers who speak the language well themselves have demonstrated substantial mastery of the subject matter. Most of these students can read with a high level of comprehension—mean score of 67.7 percent in Table 7—and have demonstrated as well their ability to learn other content when instructed in a familiar language.

In Figure 4, the number of students who performed at each level between 0 and 100 on the written part of the assessment has been broken out by program—experimental or standard. So, for example, we observe that 60 out of 292 students in the English program scored below 10 percent on the assessment. Nearly 50 percent scored below 20 percent and 72 percent scored below 30 percent. Only 28 percent of students in the standard program scored above this level. Conversely, 88 percent of children in the experimental program scored above the 30 percent level, 61 percent scored above the 50 percent level and 17 percent scored above 80 percent.

As mentioned in the introduction, the KEP intervention consists of (a) use of the L1 as a language of instruction, (b) the provision of basic textbooks (a reading primer and a math book), and (c) some adjustment in the timetable to make space for teaching reading in the first language. The combined weight of these has produced the results described above. The research design employed has not permitted sorting out the impact of each variable. In the baseline study done before the KEP intervention was launched, an effort was made to examine the impact of whether children had textbooks. At that time, we found no difference in performance based on whether children had textbooks, though there were so few who did that the data carry minimal weight.

We have not manipulated the variable of time on task (the timetable) so cannot dismiss this variable nor assign any specific impact to it.

Classroom observations and visits reveal a sharp contrast between the standard and KEP classrooms. The KEP classrooms are active, electric, and filled with a high level of interaction between the teachers and the students. In the standard classrooms, one observes a heavy reliance on choral response at the direction of the teacher chanting, for example, such statements as “Two plus two = 4”, or “This is a cat.” Children do little on their own and often contribute to the choral responses without paying any real attention to what the teacher is saying, writing, or demonstrating. This pattern of classroom interaction, so widely attested and so characteristic of L2 instruction, produces little or no learning as measured by standardized assessments.

Eventually, of course, some learning does take place since some children will ultimately take and pass the First School Leavers Examination and go on to secondary and beyond. Analysis of data from subsequent years of schooling will help us better understand this process. The Class 2 data and, as we shall see, the Class 3 data remind us that there appears to be a prolonged delay before children in the L2 model actually begin to truly learn. The data thus provide support for the claims by Cummins (reference to be supplied later) and others that it takes 5 to 7 years for a child to learn a language well enough to be effectively educated in that language.

Results for Class 3

Overall comparisons

The testing design employed for Class 3 allows us to address some interesting questions about the relative effectiveness of the mother-tongue component of the experimental program. This will be explained below.

In the research design, Class 3 children in all 24 schools took the same test of knowledge and skills in language and math in English. This test measured knowledge of English grammar, word recognition skills, and reading comprehension. The test in math covered a range of topics reflecting the curricular content of the Grade 3 math textbook. Children in the experimental KEP program also took a reading test in Kom having a structure similar to the English reading test. The texts used for reading comprehension (in both languages) used local names and events to avoid potential threats to comprehension posed by using unfamiliar story content.

Comparison of results for Class 3 across years

The same instrument was used between 2010 and 2011 for carrying out assessments at the level of Class 3. Also, those in the KEP program underwent reading assessments in both English and Kom. The data reported below in Table 9 allows us to generate a number of insights about the relationship between language and the development of reading skills.

Table 9. Comparison of results between programs for Class 3 for the years 2010 and 2011 when the testing is done in English.

	2010 (English Testing)			2011 (English reading)			2011 (Kom Reading)		
	Standard	KEP	Gain (%)	Standard	KEP	Gain (%)	Standard	KEP	Gain (%)
Language Arts	19.9	40.6	104	24.2	41.1	69.8	24.2	66.4	172
Mathematics	21.8	49.5	127	21.0	41.7	96.7	21.0	41.7	96.7
Overall	20.6	43.8	113	23.0	41.3	79.6	23.0	53.6	133

The results between the two years show some changes though these do not appear to systematically favor either program. We are not aware of any systematic changes in educational practice in the area which would account for the differences. It is true that a practice test was given this year before the final assessment was done.

The two most evident changes were significant improvement in the standard program in reading achievement (though it is still at the level of random guessing) and a larger drop in performance on the math subtest of almost 8 percentage points in the KEP program. Otherwise, we note that although the test was taken in English by those in the experimental programs, the children in the experimental program still performed at a level twice as high as that of those in the standard program. When the KEP children were tested for reading skill development in Kom, their performance was 62 percent higher than when tested in English.

A more detailed examination of performance by Class 3 children in 2011

In Table 10 we break down results of the testing by the sub-skills included in the assessment both for language arts and for math.

Table 10. More detailed comparison of results by program for Class 3 (2011).

	English	KEP	Gain (%)	Reading in Kom	Gain (%)
Language Arts	24.2	41.1	69.2	66.4	174
Grammar	27.4	42.1	53.6	65.8	140
Word recognition	21.0	41.4	97.1	*	*
Reading comprehension	23.2	39.6	70.7	67.9	193
Math	21.0	41.7	96.7	*	*
Number system	27.6	46.8	69.6	*	*
Sets	14.1	23.8	68.8	*	*
Simple calculations	18.8	45.2	140.4	*	*
Story problems	14.8	35.7	141.2	*	*
Overall Test Result	23.0	41.3	79.6	*	*

The primary assessment in Class 3 was administered to all children in English—both for reading and for math. Children in the KEP program also took an assessment of reading ability in Kom, their actual language of instruction for Class 3. The first assessment allows us to gain insight into performance when children from both programs are placed on an equal footing with respect to language of assessment. Any observed difference must, therefore, be largely attributable to effects stemming from the experimental intervention which has motivated this study. At the same time, the inclusion of the assessment of reading skill in Kom gives us a “true” read on the development of reading skills in the experimental program.

A quick scan of scores for the children in the standard (English) program again (or still) reflects group performance at or below the level of random guessing. About the only plausible conclusion to be drawn is that children in the standard program, as a group, are still not reading at all. Of course, this characterization as non-readers has to be recognized as being intertwined with weak or very weak skills in English. However, the fact that the KEP children scored in the range of 40 percent on all sections of the English test pushes us to one of two conclusions: either (a) their level of English is substantially higher or (b) they have learned to read reasonably well and can leverage this ability to perform at a higher level despite weak skills in English.

The evidence from the math section of the test lends additional support to the argument that the primary effect influencing performance is language of instruction rather than proficiency in English. Note, for example, the disparity in the performance when doing basic calculations. This is not a language-based task yet the KEP children scored 140 percent higher on this sub-skill. The primary implications would seem to be that the KEP children have made more progress in mastering math facts and processes BECAUSE of their instruction in a language they understand.

Note also the low performance by children from both programs on problems having to do with sets (union and intersection). Both programs performed at or below the level of random guessing. Even though sets are introduced into the curriculum in Cameroon in Class 1, the concept of sets is clearly not being learned. The consistently poor performance on problems involving sets suggests that the concept of sets is probably too abstract for young children (and maybe for their teachers as well). We would recommend removing this topic from the

curriculum for the early grades and giving the corresponding time to basic math facts which are more useful to children and which clearly need more instructional time to develop mastery.

Comparison of mastery in language arts

As noted above, the children in the experimental Kom-medium schools took an additional test in Kom which tested the growth of language skills—including reading—in Kom. The test in English measured knowledge of English grammar, word recognition skills, and reading comprehension. The test in Kom measured only mastery of Kom grammar and reading comprehension in Kom. All test content for both tests came from the Class 3 curriculum. Neither test was a translation of the other so that there was unlikely to be any kind of direct transfer between the two tests. Otherwise, the tests were analogous in structure with 9-10 test items on grammar and a story followed by questions testing for comprehension of story content.

This additional test permitted an interesting three way comparison as follows.

Mode 1	Mode 2	Mode 3
Skill development in English language arts by children being instructed only via English	Skill development in English language arts by children first taught to read in their mother tongue (Kom) with Kom as the primary language of instruction	Skill development in Kom language arts by children being taught primarily in the first language (Kom).

A comparison of results between Mode 1 and Mode 2 helps answer two questions. First, “To what extent does beginning instruction in L1 raise the level of mastery of the basic skills of reading and writing?” Second, “To what extent does the instructional use of the mother tongue compromise (or facilitate) students’ ability to learn and function in the second language (English)?” A comparison of results between Modes 2 and 3 will test the extent to which language of instruction inhibits or facilitates mastery of desired academic skill in reading mastery. The results of this testing scheme are shown in Table 11.

Table 11. Comparison of scores of Class 3 students in three different test modes (2011)

	English medium instruction (N=344)		Kom-medium instruction (N=231)		Kom-medium instruction (N=231)	
	English-medium test		English-medium test		Kom-medium test	
	Mean	SD	Mean	SD	Mean	SD
Grammar	27.4	15.4	42.1	18.2	65.8	21.4
Comprehension	23.2	19.0	39.6	22.0	67.9	30.2
Language arts	24.2	12.2	41.1	15.4	66.4	21.8

The mean score of those taught only in English on the combined measure of language arts was 24.2 percent (out of 100). The score of the children being instructed in Kom on the exact same test in English was 42.1 percent, an improvement or gain in efficiency of 69 percent. In the third column we note that the mean score of the Kom-medium children on a Kom- language assessment of language arts –structurally similar to the English test—was 65.6 percent, a level of performance more than 20 percentage points higher than the performance of the same children on the English-medium test of the same skills—a relative gain of 56 percent. If we take 65.5 percent to be a “normal” level of reading skill for Kom children being educated in Kom, we can then compute penalties or “costs” for the two variables of knowledge of English and language of

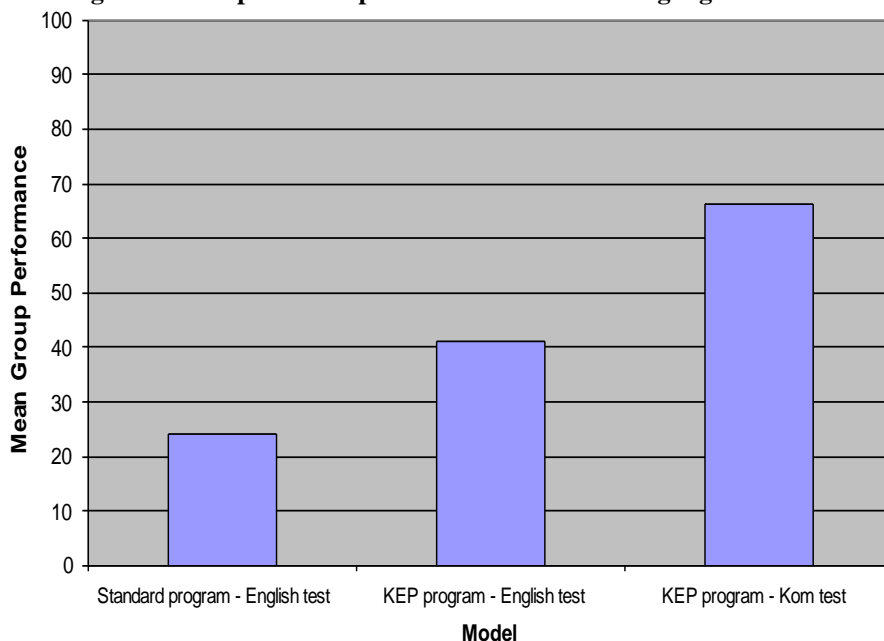
instruction. The decrease in performance from 65.8 to 42.1 percent due merely to the language of the test suggests that the penalty or cost for this variable is 38 percent $((66.4 - 41.1) / 66.4)$. The language of instruction variable alone imposes a further penalty or cost of 41 percent. Collectively, the cost or penalty is 64 percent. If expressed in economic terms, we can say that for every 1000CFA being spent, only 360CFA of educational return or value is being created.⁶

If, on the other hand, we take the performance of the English-medium children on the English-language assessment to be “normal”, we can then compute gains in return or efficiency as a result of adjusting the two variables referenced above. Improved performance based on the variable of language of instruction variable by itself (24.2 percent to 41.1 percent) represents a gain of 70 percent. Shifting the assessment to the language of instruction for the KEP children results in a separate gain of 62 percent. Doing both produces a gain of 174 percent. Cast in economic terms, this means that each 1000CFA spent produces a return of 2740CFA.

The data are very suggestive of a set of conclusions (answers to the questions raised above) though we need more data and experience to be confident in our conclusions. **First**, the Class 3 data clearly support the conclusion that the Kom-medium children are learning more and learning more efficiently than the children in the English-medium or standard classrooms. **Second**, these data clearly support the conclusion that the Kom-medium children are NOT being compromised in their mastery of the curriculum including the English curriculum since they scored almost twice as high as did the children in the English-medium classes. **Third**, the data suggest that reading skill is clearly superior when the language of reading and testing is a/the language the children know well rather than a second language.

Figure 4 (next page) presents in graphical form the same data found in Table 11 for the benefit of those who find a graphical format easier to follow.

Figure 5. Comparison of performance on test of language arts for children in Class 3.

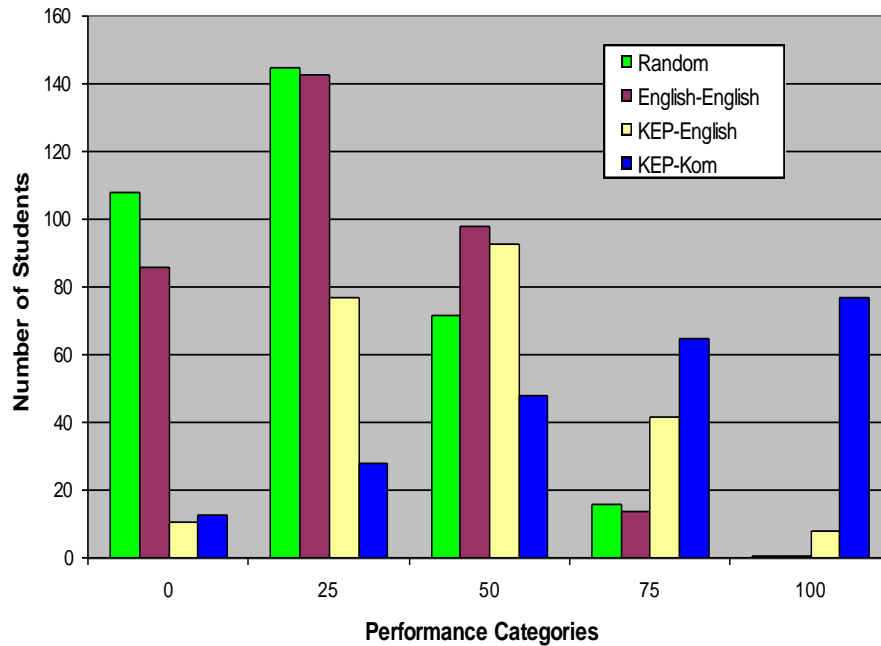


⁶ The economic analogy is not intended as a cost-benefit analysis but simply another mechanism for describing the contrast between the two levels of achievement.

Reading Comprehension in Class 3

Since reading comprehension is a fundamental instructional objective of early primary education, we examined more closely the evidence of progress towards being able to read with comprehension in the three different test modalities resulting from the assessment design for Class 3. We begin with a graphical rendition of the results as shown in Figure 6

Figure 6. Proficiency in reading comprehension demonstrated in three different assessment modes by Class 3 children.



Again, the three test modes are: (1) English-medium children being tested for reading comprehension in English; (2) Kom-medium children being tested for reading comprehension in English (exact same test); and (3) Kom-medium children being tested for reading comprehension in Kom.

The nature of the Class 3 assessment made it convenient to divide the data into 5 quintile-like distributional points. The number of children from each testing mode as described in the preceding paragraph is plotted in each of these 5 groupings (which have no independent conceptual justification in terms of phases in the process of learning to read). What IS clear is the fact that each mode produces a distinctive distribution in terms of skill in reading comprehension. In mode 1 (English-medium instruction, English reading comprehension), the distribution is clearly weighted towards the lower end of the scale. Distributionally, the performance of the children in this mode is very similar to that of a random guessing model (the green bars or the bars on the left-hand side of each cluster of bars).

At the other end of the scale, the blue (the darkest) bars reflect the performance of KEP children when reading in their first language and their language of instruction (mode 3). The distributional profile is most similar to that seen in criterion referenced testing in which children have, in fact, met or come close to meeting the criterion which has been specified.

Finally, the distribution of readers in the second mode (KEP children reading in English) approximates a normal (though flattened) distribution. The theoretical significance of this distribution for these children is not obvious. Pragmatically, it appears to be the result of applying a skill (reading) which has been largely mastered to a language (English) which has not.

In an effort to match the distributional profiles in Figure 6 to the more pedagogically defined idea of actually becoming a reader, the distributional data can be helpfully clustered into three groups: (1) non-readers (those who demonstrate zero comprehension on the reading tasks contained in the assessment), (2) incipient readers (those who scored above zero but not higher than 50 percent on the reading comprehension task); and (3) passable to good readers (those who scored above 50 percent on the reading comprehension tasks. The resultant profile of readership is set forth in Table 12.

Table 12. Progress towards becoming readers in Class 3.

Instructional/testing combination or mode		Non-readers (zero comprehension)	Possibly incipient readers (greater than zero but not more than 50 percent)	Passable to good readers (well over 50 percent comprehension)
Mode 1	English-medium, English test	25.1 percent	70.5 percent	4.4 percent
Mode 2	Kom-medium, English test	4.8 percent	73.6 percent	21.6 percent
Mode 3	Kom-medium, Kom test	5.6 percent	32.8 percent	61.5 percent

Under a more optimistic interpretation of the data, a little over 4 percent of the children in standard Class 3 classrooms are reading with passable comprehension. Taking a more pessimistic perspective, it is possible to interpret the data (applying the probabilities of random guessing) to mean that 0 out of 314 children tested in this mode are able to read with comprehension. That is, choosing answers based on tossing a four-side die would also have resulted in a similar number of children achieving the higher level of performance.

When we compare the first and second rows in Table 12, the big difference is the small percentage of tested students in the 'zero' column of the second row (just 4.8 percent) while we find 21.6 percent of students in the last column (passable readers). The educational difference between the children in the first and second rows is that the children in the second row were taught to read in Kom and were able to transfer that skill to English without focusing excessively on learning to read in English. Their comprehension, however, is still relatively low since they were being tested in a second language.

The third row (Kom-medium children being tested in Kom) presents yet a different profile. The children in this row were taught to read and write in their mother tongue AND they were being tested for reading comprehension in the same language. Note that 61.5 percent of the children tested are reading with comprehension. In fact 30 percent of the children tested in this mode read with 100 percent comprehension compared to 1 student (out of 342) in the first mode (English-medium, English test) and only 3.5 percent in the second mode (Kom-medium, English test).

The apparent conclusions to be drawn from these results are the following:

1. Most children in the Kom-medium schools are readers by the end of Class 3 (based on the evidence from row 3 AND a comparison of rows 1 and 2 in Table 12);
2. By having learned to read in their first language, the children in the Kom-medium schools are able to apply those skills to reading in a second language even though their mastery of that language is limited (based on a comparison of rows 2 and 3);
3. Depending only on a second language to teach children to read appears to be a highly inefficient model (based on a comparison of rows 1 and 3). Unfortunately, this is the prevailing model in most developing countries.

Gender in Class 3

In reports from previous years it was noted that the variable of gender produced somewhat uneven and surprising results in the Kom data especially given the prevailing view that education in Africa seems to strongly favor males. Table 13 reports the findings from 2010 for Class 3.

Table 13. Comparison of Class 3 performance by Gender.

Metric of Comparison	Males	Females	statistic
Entire test population			
Language Arts	30.8	31.1	t = 0.23; P = 0.820
Math	31.0	27.9	t = 1.85; P = 0.065
Overall	30.9	29.9	t = 0.76; P = 0.450
English-medium schools (English test)			
Language Arts	24.0	24.4	t = 0.27; P = 0.789
Math	22.7	19.7	t = 1.89; P = 0.059
Overall	23.5	22.6	t = 0.76; P = 0.447
Kom-medium schools (English test)			
Language Arts	40.1	41.9	t = 0.89; P = 0.376
Math	42.4	41.1	t = 0.49; P = 0.622
Overall	41.0	41.6	t = 0.32; P = 0.752
Kom-medium schools (Kom test)			
Language Arts	64.4	68.2	t = 1.31; P = 0.192

In the data from 2011 for Class 3, there is no consistent pattern in terms of the relative performance of males and females. Boys generally outscored females on math while females tended to outscore boys in language arts with the biggest advantage manifested in Kom reading. None of the differences was statistically significant. The results from 2011 contrast somewhat with those from 2010 in which females generally outscored males on all measures though not by large margins.

Fulfulde-speaking students in Class 3

A common issue raised in discussions about the feasibility of mother-tongue education is the problem of "mixed communities." If there are members of two different linguistic communities living together or interspersed one among the other, then, it has been argued, a mother tongue model is not possible. The only option is to educate everyone in a second language.

To our knowledge virtually no hard data exist describing likely outcomes under this condition of "mixed languages." The Kom research setting provides some data which can be presented to shed some light on this situation of "mixed languages." In some parts of the Kom area, one finds a sizeable population of Fulfulde speakers. Since there are no separate schools for the children of this language community, they attend whatever school is nearby. The 2011 data set tracked whether the children being tested were speakers of the Fulfulde language. The majority of these children were in the English-medium program giving a fair amount of reliability in statistical terms. Table 14 reports our findings on this sub-population. Because this is not a large population (approximately 23 in total), it is not possible to draw major conclusions from the data..

Table 14. Test results for Fulfulde-speaking children in both control and experimental schools.

Metric of Comparison	Fulfulde	Non-Fulfulde	statistic
Entire test population			
Language Arts	25.5	29.5	P = 0.178
Math	29.0	31.1	P = 0.371
Overall	27.7	30.5	P = 0.167
English-medium schools (English test)			
Language Arts	28.0	24.0	P = 0.107
Math	22.7	20.9	P = 0.507
Overall	26.0	22.8	P = 0.109
Kom-medium schools (English test)			
Language Arts	36.0	41.1	P = 0.390
Math	44.4	41.7	P = 0.690
Overall	39.2	41.3	P = 0.120
Kom-medium schools (Kom test)			
Language Arts	31.0	66.8	P = 0.001

In the 2010 data, the Fulfulde children outscored the non-Fulfulde children on virtually all measures with the exception of the Kom language reading test. This year, the data are more ambiguous leaving us uncertain as to whether this year's or last year's data are/were unusual.

Looking at the entire test population, we note a small advantage on all measures for the non-Fulfulde population though none of the differences is statistically significant. However, when we look just at those cases where Fulfulde children were attending English-medium schools, we find the Fulfulde children scoring slightly higher than the non-Fulfulde children though, again, none of the differences is statistically significant.

However, when the Fulfulde children attended a Kom-medium classroom (there were only 3 cases in the data this year), they performed slightly below their Kom-speaking counterparts with the exception of math. Furthermore, the children in the Kom-medium schools outperformed their Fulfulde counterparts in the English-medium schools by approximately 50 percent.

The one measure on which the Fulfulde children fared significantly worse than the Kom-speaking children was that of language arts when the test was administered in the Kom language—31 percent versus 66.8 percent. This contrasts sharply with the data from 2010 when the difference between the Fulfulde and Kom-speaking children on the Kom language reading

assessment was only about 2 percent. We will have to have more data before we can come to any conclusion as to what is normal in this situation.

What do we learn from this data? (1) Instruction in English serves the Fulfulde-speaking children no better than it does the Kom-speaking children. (2) The Fulfulde-speaking children attending a Kom-medium school still benefit from this model almost as much as the Kom-speaking children. Why? The answer is almost certainly the fact that they have learned at least some of the Kom language from living in the community—at least enough to be effectively instructed in that language. However, the evidence this year is not nearly as positive as that from last year.

Late entry into KEP (data from 2010)

This section is included from last year. This year (2011), the data gatherers did not track this variable so we cannot provide additional evidence.

Despite instructions and guidance from the research staff, there are students who entered the experimental program the second and third years having missed entirely the first year(s). While this phenomenon compromises the maintenance of a "clean" program for research purposes, it also offers an opportunity to observe the consequences of late entry. The relevant data are summarized in Table 14.

Table 14. Data on late entry into the experimental program (data from 2010).

	English-medium	Late entry KEP	Full KEP	Statistics
Language Arts	19.91	31.45	41.51	F = 142.27; P = 0.000
Math	21.84	37.66	50.70	F = 128.38; F = 0.000
Overall	20.60	33.68	44.81	F = 172.06; P = 0.000

There were 22 such students identified in the testing done in Grade 3. All of these were students who entered at the beginning of the third year. The most obvious comment to be made is that the late entry students clearly outperformed the English-medium students. Interestingly, the mean performance of this group of late entry students is about 55 percent of the way towards the performance of the full KEPP students at each of the three points of measurement. On a preliminary basis, we can make the broad generalization that the one year of Kom-medium instruction has definite value in terms of educational progress though it is not enough to make up the gap created by being in the English-medium schools for Classes 1 and 2.

Age

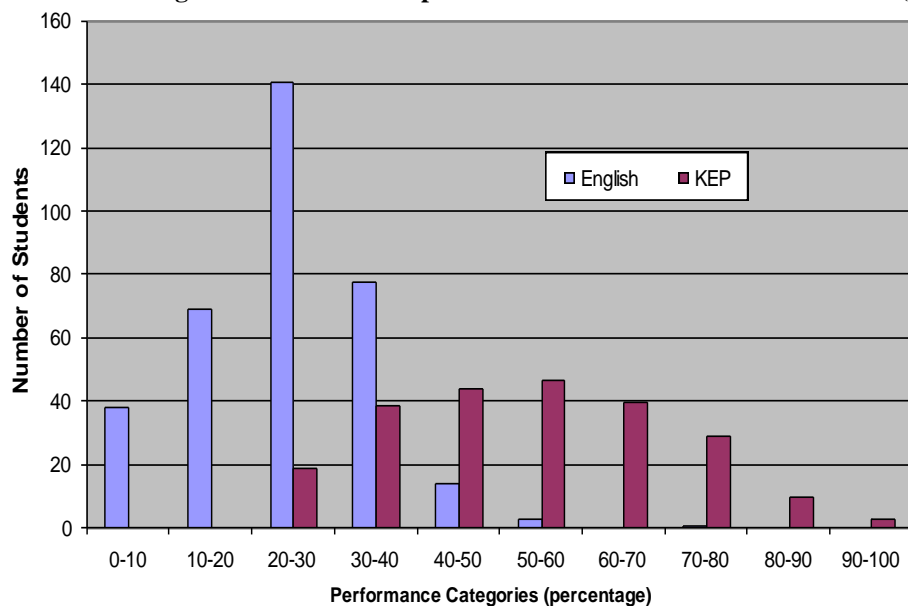
The variable of age was examined for evidence of impact on any of the outcomes. For children in Class 3, only two trends were noted the most significant of which was that 6 year old students (of which there were only a few) did markedly worse on all measures of performance and especially in the English-medium program. The data suggest quite convincingly that parents are wasting their money sending 6 year-old children to Class 3. Further, it is likely, though we lack formal evidence to say so, that their presence is also an impediment to teachers and other students in the learning process.

The other trend was similar to that noted for Class 2. In the Kom-medium program, there is a small correlation between age and performance with older children tending to score higher. The trend, however, was not statistically significant.

Performance profile for Class 3

The performance profile for Class 3 (2011) was constructed by using the assessment scores for the English-medium program straight from the data. However, the data for the KEP children is a composite of their performance on the common math test (given in English) and their performance on the Kom reading assessment rather than the English reading assessment. We felt this would produce a profile more consistent with the instructional experience of these children.

Figure 7. Performance profile for Class 3 based on overall scores (2011).



The distributions for both groups more closely approximate a normal distribution than we saw in the profile for Class 2. The primary difference is that the distributions have radically different shapes. The English profile is centered on 20-30 percent while the KEP profile is centered on 50-60 percent. Those in the KEP program are clearly exhibiting substantial learning while those in the English program are still characterized by performance which is not very different from that generated by random guessing. In fact, there are still more scores in the two lowest categories than one would predict from a random guessing model, but fewer than seen for Class 2. A slightly cynical interpretation would be that the children in Class 3 have become better at random guessing than those in Class 2. A more positive interpretation is that we are seeing the faint outlines of actually being able to figure out the content of the test but little evidence of actually being able to read.

It is reasonable to ask of the KEP program why there are still a substantial number which lag behind in the range of 30-50 percent. Our best guess at this point is that we have a case of a potentially effective instructional model which is running into the various systemic limitations of rural education in Cameroon—inadequate time on task, weak pedagogy, under-age children in class, lack of supervision, etc. Improvement in educational outcomes dependent directly on the

intervention are evident, but there are constraints on the overall reach of this intervention due to the broader systemic limitations of the educational system in place.

Interpretative comments on the findings for Class 3

In short the assessment results indicate that children in the English-medium program are still not readers, as a group, by the end of Grade 3. These data clearly reflect the common observation about Africa that it takes 4 to 5 years for children to learn to read. The evidence is quite clear, however, that this statement needs to include the addendum that this is true IF such children are being taught in a second language. IF they are being taught in their first language with even a modicum of support, most are passable readers by the end of Class 2 and reasonably good readers by the end of Class 3.

Proficiency levels for the KEP children in both math and reading at the end of Class 3 are still a little below the expectations of the national curriculum (with those in the English-medium program barely visible on the horizon). Assuming that the national curriculum is paced appropriately (a debatable issue), we would have to conclude that other educational weaknesses such as those already alluded to likely account for these results.

Results for Class 4

The results for Class 4 are of intense interest as this group includes a cohort of children who completed 3 full years of Kom-medium instruction and are now fully immersed in an all English environment along with those who have been in an English-medium environment from Class 1.

As in the past, the assessment instrument for Class 4 was based on content of the curriculum. The evidence is that some teachers/schools covered most or all of the curriculum during the year while others did not. We noticed, in processing the data that children in some schools consistently did well on certain tasks or questions while others often did poorly.

The testing format used this year for Class 4 was a little more challenging than in past years, especially for the reading section of the assessment. Following assessment strategies currently being employed in the US, we selected longer reading passages and then asked questions about these passages which probed not merely basic comprehension, but also more general comprehension, ability to draw inferences from the passage, ability to generalize about the passage, as well as knowledge of specific lexical items encountered in the various passages.

In addition, the multiple choice format was made more challenging by making use of more subtle distinctions than in past years. (For example, a question about a particular outcome of a story might include response options which were culturally plausible, but not correct in terms of the content of the passage.) We expected the assessment to be quite challenging for Class 4 students and a good test of their strengths and weaknesses as readers.

The math section of the assessment included 31 items while the language arts section had 23 items. The details of what these items covered will be reflected in the subsequent tables which present the results from the assessment.

Overall results

The assessment for Class 4 included just two major components—language arts (reading) and math. The results are presented in Table 15. In the table, we will continue to use the labels standard or English to designate children coming from this program and KEP to refer to children who came through the experimental program (though they are no longer in a mother tongue program). In Class 4 ALL children are now in English-medium classrooms receiving the same instruction in the same language. An examination of the data revealed very little movement or mixing of students in schools so the cohorts by classroom continue largely as they were in the earlier classes.

Table 15. Summary results from Class 4, 2011.

	Standard			KEP			Gain (%)	Statistics
	N	Mean	SD	N	Mean	SD		
Language Arts/Reading	361	28.1	11.4	204	38.7	13.5	37.7	T = 9.48; p = 0.000
Math	361	27.3	15.9	204	40.1	19.2	46.9	T = 8.09; p = 0.000
Overall	361	27.6	11.5	204	39.5	14.7	43.1	T = 9.93; p = 0.000

Several observations can be made about these data. First, the children coming from the standard program continue to score at the level of random guessing or just slightly above that level. Second, the scores of the children coming from the KEP program no longer show the larger spreads or differentials we saw in the earlier classes. Third, even receiving identical instruction and in a second language, the children who participated in the KEP program still show substantial carryover benefits from having been in that program. Overall, these children scored 43 percent higher than did the children from the control group. Fourth, neither of the groups scored well. We will explore possible explanations in the interpretive section which comes later.

A more detailed look at the data

Table 16 examines in more detail the content of the assessment and the relative performance of the two groups of Class 4 students.

Table 16. Detailed comparison of Class 4 students on the sub-skills of the assessment (2011).

	Standard Program			KEP Program			Gain (%)	Statistics
	N	Mean	SD	N	Mean	SD		
Language Arts/Reading	361	28.1	11.4	204	38.7	13.5	37.7	T = 9.48; p = 0.000
Comp. factual	361	32.5	25.3	204	46.4	26.8	42.7	T = 6.07; p = 0.000
Comp. holistic	361	28.5	31.9	204	42.4	37.5	48.8	T = 4.45; p = 0.000
Conceptualization	361	24.1	24.7	204	30.6	26.8	27.0	T = 2.82; p = 0.005
Word knowledge	361	25.2	19.4	204	24.5	18.6	-.03	T = 0.39; p = 0.698
Punctuation	361	22.6	23.3	204	46.1	32.1	104	T = 9.15; p = 0.000
Grammar 1	361	35.6	33.5	204	49.8	35.3	40.0	T = 4.66; p = 0.000
Grammar 2	361	33.4	26.5	204	40.8	27.4	22.2	T = 3.13; p = 0.002
Math	361	27.3	15.9	204	40.1	19.2	46.9	T = 8.09; p = 0.000
Calculations	361	28.9	20.2	204	42.0	23.9	45.3	T = 6.57; p = 0.000
Fractions	361	18.6	20.8	204	25.7	25.4	38.2	T = 3.40; p = 0.001
Geometry	361	28.3	24.7	204	45.1	27.6	59.4	T = 7.19; p = 0.000
Story problems	361	27.5	23.6	204	39.4	25.3	43.3	T = 5.52; p = 0.000
Overall	361	27.6	11.5	204	39.5	14.7	43.1	T = 9.93; p = 0.000

First, we note that the performance pattern for those coming out of the standard programs continues to be in the range of 20-30 percent, what we have so far characterized as the level of random guessing. At the same time, we see some glimmers of upward performance gains with scoring on three sub-skills above 30 percent—comprehension of factual information in a text, grammar 1 (the proper use of the English comparative), and grammar 2 (use of tense/aspect in English). There is also an increase in the number of mean scores which are in the high 20s approaching 30 percent.

On the KEP side the major finding is that scoring has dropped from the 60 percent range which characterized Classes 1 and 2 and Class 3 when assessment involved reading comprehension in the Kom language. In Class 4 we find performance hovering in the low 40 percent range with a few sub-skills dipping below the 40 percent range.

Several of the sub-skills stand out as being atypical of the overall pattern. Knowledge of English vocabulary was very low for both groups as was ability to handle fractions. Students from both programs also had difficulty thinking broadly about a text (main point, possible title, summary statement). The very low performance level on fractions is actually rather common in developing countries. Obviously, fractions are hard to teach and to learn. An examination of individual test papers indicated that students definitely wrestled with the problems on fractions but simply had not mastered the key conceptual notion of a fraction and its relationship to a whole nor the processes for dealing with operations involving fractions. This is an area which clearly needs more attention in instruction and more practice on the part of students.

Looking at the math section, it is clear that even basic math facts have not been mastered by the end of Class 4. Students commonly did addition when the indicated symbol (and written instruction) was to multiply. Similarly, students did multiplication (usually incorrectly) when division was the indicated task. In our judgment, in addition to the limitations imposed by

language of instruction, students do not receive enough instruction or enough practice to consistently master even basic math tasks.

A few comments about English and English reading are also worth noting at this point. First, the KEP children have clearly learned more about the formalisms of language such as punctuation and grammar. It is also evident that their previous experience with Kom reading had imparted some basic skills in terms of how to attack a text for comprehension purposes. The KEP children were more adept at finding information in a text whether literally given in the text or expressed in an indirect manner. They were also somewhat better at thinking globally or holistically about a text.

One of the striking findings of the assessment is that both groups have a poor grasp of English vocabulary beyond the basics. Neither group was able to use text and context to determine the sense of a word which they might recognize but which was being used in a more nuanced way. This lack of knowledge, in turn, significantly reduced reading comprehension (Droop and Verhoeven, 2003).

In the final section, we will provide some additional reflections both on the meaning of the data and the reasons for the performance of the Class 4 students.

Age

There is one interesting age-related anomaly in the data; the average age of children who went through the KEP program is 1.3 years less than those who came from standard English-medium schools. At this point we lack the data to determine whether this is because of much lower attrition and repetition rates or to a tendency to start younger since language of instruction is not a hindrance. The same pattern was not evident in the Class 3 data.

The problem of under-age children is apparent again in Class 4 with 19 seven year olds and 70 eight year old children. An examination of the data by performance, however, showed little apparent impact of age on performance.

Gender

Performance by gender is virtually identical for the major test sections so will not be further investigated in this report.

Fulfulde children

The pattern of performance of Fulfulde children in the Class 4 data is different yet again from any of the patterns reported previously. Quite likely this is due to the small sample size—14 such children coming from the standard program and 9 from the KEP program. We will report the findings but refrain from drawing any major conclusions since the sample size is so small.

Table 17. Comparison of the performance of Fulfulde and non-Fulfulde children in Class 4.

	Coming from the Standard Program		Coming from the KEP Program	
	Fulfulde	Non-Fulfulde	Fulfulde	Non-Fulfulde
Math	23.0	27.4	44.4	39.9
Language arts/Reading	24.2	28.3	49.8	38.2
Overall	23.5	27.8	46.7	39.2

Curiously, the Fulfulde children coming from the KEP program outperformed their KEP peers by sizable amounts on all sections of the assessment. Due to the small sample size, however, none of the differences is statistically significant. On the other hand, the Fulfulde children coming through the standard program, scored below their peers in this program. We can offer no obvious explanation other than the likelihood that this is a sampling problem.

Because of observably large differences in the performance of the Fulfulde children between programs, we tested these comparisons statistically with the following results.

Table 18. Comparison of Fulfulde children in Class 4 coming from Standard vs. KEP programs (2011).

	Coming from the Standard program			Coming from the KEP program			Gain (%)	Statistics
	N	Mean	SD	N	Mean	SD		
Math	14	23.0	11.7	9	44.4	23.1	93.0	T = 2.57; p = 0.028
Language arts/Reading	14	24.2	8.65	9	49.8	17.5	104	T = 4.06; p = 0.002
Overall	14	23.5	7.60	9	46.7	18.1	98.7	T = 3.64; p = 0.005

Both in terms of measured performance and statistical comparisons, the Fulfulde children in Class 4 who came through the KEP program scored much higher than did those coming through the standard program. Overall, the former group scored just under twice as high on the exact same test and after having received the exact same instructions. As before, however, we must caution that we are dealing with a very small sample size. There is also the possibility that those coming through the KEP program might have gone to the better KEP schools though a scan of the data indicates only a minor difference in the ranking of the originating schools within their respective programs.

School and Program effects

A program effect has already been observed in terms of the relative performance of children coming from the KEP MLE program versus those coming from the standard program. In this section we further investigate what we will somewhat imprecisely term “the school effect.” Past reports have demonstrated that schools/teachers vary quite substantially among themselves in terms of the level of performance of the children attending such schools. This has consistently been true for both programs.

What we want to begin to investigate here is the extent to which there is any kind of carryover effect once children move from the KEP MLE program to all-English instruction. Does this

transition put all children and all schools back at the same starting point in terms of instructional impact or might it be the case that schools taking in students who come with a stronger academic foundation are able to more effectively build on that foundation and take them even further academically? Previous research on early exit multilingual education has indicated that there will tend to be a “regression to the mean” on the part of such students. That is, they will tend to lose the momentum gained in the MLE program and slowly return to the level of those who were educated only in an immersion program—the standard English program in this case.

Table 19 is a ranking of all schools included in the Class 4 assessment. The identity of the schools is not given here but can be provided upon request.

Table 19. Ranking of all schools participating the assessment of Class 4 students (2011).

Rank	Mean Score	Program from which students came	Rank	Mean Score	Program from which students came
1	50.56	KEP	14	33.11	KEP
2	49.66	KEP	15	30.91	English
3	41.78	KEP	16	29.63	KEP
4	40.74	KEP	17	27.78	English
5	40.40	English	18	27.50	English
6	40.04	English	19	26.56	English
7	38.37	KEP	20	26.37	English
8	38.27	KEP	21	26.09	English
9	36.73	KEP	22	25.84	English
10	36.35	KEP	23	25.41	English
11	35.88	KEP	24	23.88	English
12	33.80	English	25	22.69	English
13	33.61	KEP	26	20.60	English

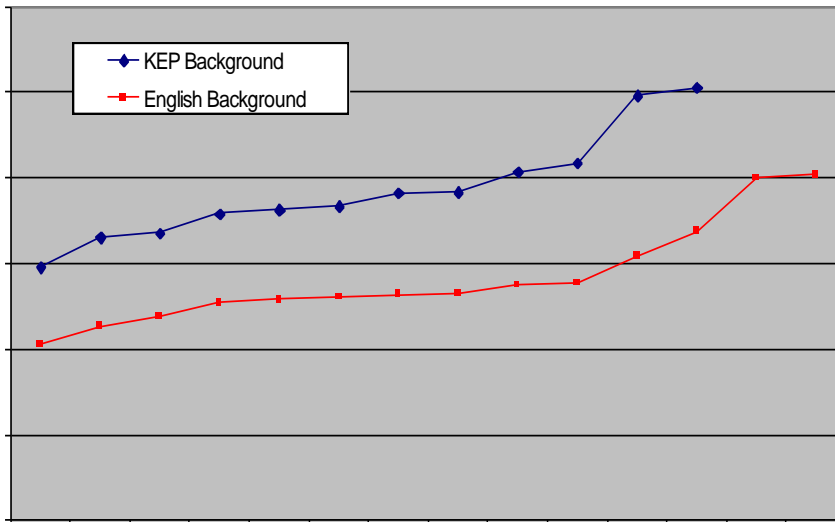
From Table 19 it is clear that there is a program effect that extends beyond the end of the program. The top 4 schools are KEP schools and 9 of the top 12 are KEP schools. At the same time the 10 poorest performing schools are those whose students came from the standard program (mainly from earlier classes in the same school). The two highest performing schools in Class 4 are public government schools which were also high performers within the MLE program (Classes 1-3). The two high-ranking English schools are small private schools—one Catholic and one Presbyterian.

The distribution of schools by their ranking was tested with the non-parametric Mann-Whitney test and found to be statistically significant ($W = 125$; $p = 0.0011$) providing further evidence that the effect of the MLE program does extend into performance at the level of Class 4.

Figure 8 plots the performance of all schools based on their Class 4 mean scores by program. As observed in previous years and reports, the plot lines for the two sets of schools are very similar. Both programs show 4 schools which would appear to be somewhat superior, 3 schools which seem to be low performers, and a cluster of ‘average’ schools which appear to make up the mid-range of schools. Note especially that there is quite a consistent 10 percentage point gap between the two sets of schools at both the high and low end of the performance curve. This is

the visual correlate of the reported 43 percent performance gain based on the KEP background of the students coming from the experimental program.

Figure 8. Plot of the performance of the two sets of schools based on the origin of their students—KEP or standard program. Based on Class 4 data from 2011.



One additional test was applied in an effort to measure the impact of previous schooling on performance of children in Class 4. We tested the relationship between previous school rankings and their (the students’) performance in Class 4. We found the following.

Table 20. Analysis of the impact of school quality in previous years on the performance of children in Class 4.

School rank	Standard schools			KEP schools		
	R-sq	p	Coefficient	R-sq	p	Coefficient
Year 1	4 %	0.000	.71	0.5%	0.164	.46
Year 2	1 %	0.039	.37	10.0%	0.000	1.2
Year 3	0%	0.764	.06	5.6%	0.000	1.1
Mean rank	1.8%	0.009	.70	12.3%	0.000	2.4

The numbers in the column “R-sq” indicate how much of the observed variation in overall student performance in Class 4 is explained by the ranking of the school of the student(s) in the three previous years of the research program. The numbers in the “p” column indicate whether the value in the R-sq column would be judged to be statistically significant or not. The number in the column labeled as “Coefficient” indicates how much the mean performance of students in a given school increase with each increase in school rank.

What Table 20 tells us is that children coming from standard schools are much less impacted by the quality (or rank) of the school(s) they attended in the earlier classes than are the children who come from KEP schools. We can use the evidence of the last row labeled “Mean rank” to illustrate this effect. “Mean rank” refers to the average ranking of the school(s) attended by a child for the first three years of attendance. Since schools (or teachers) may vary in performance

from one year to the next, “Mean Rank” provides us with a general measure of a school’s performance over a three year period of time. In the case of students coming from English-medium schools, their performance improved only .7 points for each increase in rank. This means that a child attending the highest ranked school could expect to see his or her performance exceed that of a child coming from the lowest rank school by a total of about 8 points as a result of having attending a top school rather than the lowest ranked school.

In the case of the KEP schools, a child attending the top KEP school could expect to see his/her performance in Class 4 increased by 29 points as a result of attending the best KEP school rather than the poorest performing KEP school.

What does this mean and why did it happen? There are several possible interpretations so we have to be careful in making too strong of a claim. The most amenable interpretation is that the KEP schools, as a group, have a greater educational impact upon their students so their quality is reflected in the educational development of their students. This effect also exists for the standard schools but is only about 30 percent as strong.

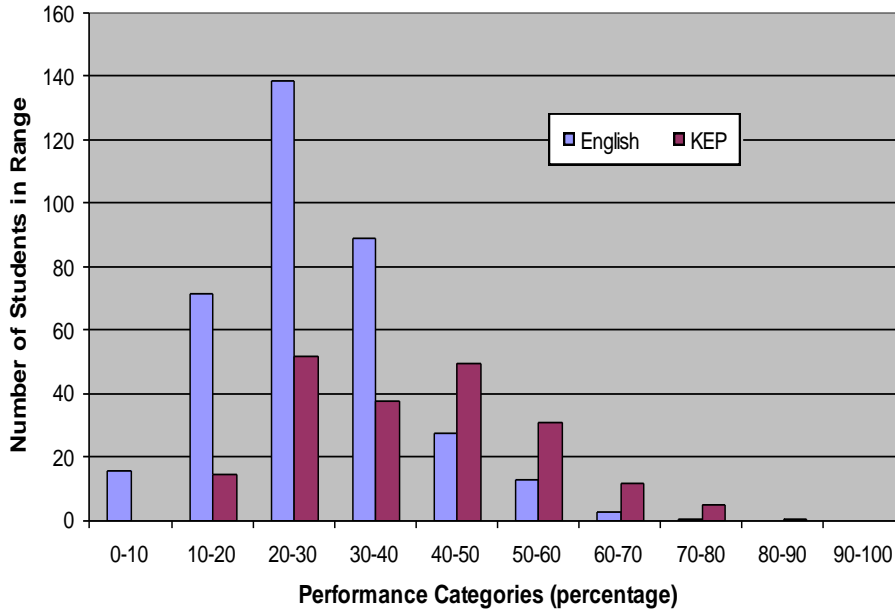
An alternative interpretation consistent with the data is the possibility of a “high-end” effect; that is, the observed effect is mainly true for the high performing schools and not so much for average or poorer schools. There is at least a little support for this interpretation both in Figure 8 as well as the data plot (not included in the report) which showed a somewhat more marked effect at the high end of the rankings.

Yet another explanation may be the problem of language of instruction. If children are attending a school where they do not understand what the teacher is saying, it matters less (or maybe little at all) whether the school is a high-performing school or a low-performing school—little is going to be learned by most children.

A performance profile for Class 4

Figure 9 presents a performance profile for all students tested in Class 4. We draw two primary conclusions from the profile. First, there has been a significant reduction in the distinctiveness of the profiles seen for the earlier classes. The two profiles are now closer together (in their mean) and more similar in their distribution. Secondly, the profile of the students originating from standard schools is still centered in the 20-30 percent category indicative of random guessing.

Figure 9. Performance profile for children in Class 4 (2011).



It is not difficult to conclude from this profile and the other data presented above that children in Class 4 whose previous educational experience was in the standard program have either still not learned to read or have not learned enough English to understand what is being taught in the classroom (or both). Their performance on the math section of the Class 4 assessment (which requires almost no language knowledge) certainly supports a conclusion that minimal learning is taking place whether because of the language of instruction or because of other factors in the educational environment.

The profile of the KEP group is much less positive than in previous classes. Two or three possible conclusions are possible. (1) As many have suggested, the early exit from a more supportive mother-tongue instructional environment has negative educational consequences as reflected in the relative similarity of the two profiles. (2) None of the children are truly ready for English-medium instruction. The smaller observed advantage for the KEP children is merely a carry-over from the more favorable instructional experience in Classes 1-3. (3) The lower performance in Class 4 by the KEP students is temporary. Once they begin to get a handle on English, the educational foundation laid earlier in Classes 1-3 will enable them to make stronger progress in school in Class 5 and beyond.

We probably will not know until the end of Class 5 which is the better interpretation of the Class 4 results.

Interpretive comments on Class 4 results

Considerable interpretation has already been given so we will try not to be redundant. First, we know from the assessments done in earlier classes that children coming from the KEP schools come to Class 4 knowing how to read and having general familiarity with the nature and structure of writing and of texts. At the same time, even though they showed a 37 percent advantage over the children coming from standard schools on the language arts assessment, the Class 4 assessment seems to indicate fairly clearly that their proficiency in English is insufficient to fully benefit from instruction in English at the level of Class 4. A more fine-grained

assessment of English proficiency is needed before a more definitive statement can be made as to the true level of proficiency of both groups.

It is equally evident that children coming through the standard schools have not learned to read with comprehension in English by the time they arrive in Class 4 and still have minimal proficiency at the end of Class 4. Because they cannot read reliably in English, it is hard to assess their actual proficiency level in English from the Class 4 assessment. If the data from the math part of the assessment is any indication, it would appear that their English is no better than that of the KEP children. It will be interesting to see to what extent such proficiency emerges in Class 5.

A review of the response patterns of children in Class 4 coming from English-medium schools indicates that, for many such students, it would be better to characterize their performance as “systematically poor guessing” rather than “random guessing.” Let me explain. When pure random guessing is taking place, students select answers to multiple choice questions in such a way that each response option receives approximately the same number of guesses. So, if there were 4 possible answers and 100 children taking a test, each of the four options—we will call them A, B, C, and D—would be selected approximately 25 times as the correct answer. That is not the pattern seen for the reading comprehension questions. Rather, some answers were far more likely to be selected than others. Some options, in fact, were hardly selected at all while others were notably over-selected. Furthermore, possible answers which were over-selected tended to be either (a) culturally plausible answers or (b) semi-correct answers in the sense that actors and actions occurred somewhere in the text but not in the way stated in the distractor(s).

These observations lead to the following tentative conclusions. First, a fair number of Grade 4 students coming from English-medium schools may be able to read well enough to identify key actors and actions in a story or to locate relevant details, but not well enough to reliably use the information thus gathered. Often it appeared that students read the questions and the possible answers but not the story (or not all of it). They tended to select reasonable, but not necessarily correct, answers including actions or events not even referred to in the stories in the assessment. This impression was reinforced by performance on the first item in the reading test. By accident the first item was a repeat of the practice item which appeared on the same page immediately above the first test item. Performance on this item was just over 50 percent for those from English-medium backgrounds and over 80 percent for those from Kom-medium backgrounds. Even though it appears that the level of English proficiency is approximately the same between the two groups, the fact that the Kom-medium children could read quite reliably led to the difference in performance on this item.

Second, it is also possible to argue that the results reflect a poor level of English mastery rather than a low level of reading comprehension. This hypothesis draws support from the observation that both groups performed better on test items which only required locating factual details in the text versus showing overall comprehension of the text.

Conclusion and Discussion

Overall, the data from 2011 are consistent with the pattern observed and reported in previous years (2008 – 2010). In general, children in the Kom-medium program show a collective gain in learning efficiency of about 125 percent over their peers in the standard English-medium program in the first 3 classes.

The Class 4 data give us a partially different profile. While the children from the standard program are still scoring largely at the level of random guessing, the children from the KEP program have seen a significant decrease in their educational achievement. Their mean scores demonstrated an advantage of only 40-50 percent compared to the 125 percent (and more) advantage demonstrated in earlier classes. Furthermore, the mean performance of children from a KEP background slipped below 50 percent for the first time. The overall findings lead us to the following tentative conclusions.

1. The data indicate that the Kom-medium children are quite proficient in reading Kom-language materials after three years of schooling. This contrasts sharply with the performance of children from the English-medium program who still show very little reading skill after four years of schooling.
2. The Kom-medium children have consistently demonstrated a higher level of proficiency in mathematics indicating that mastery of this subject matter is also enhanced by use of the mother tongue for instructional purposes. This advantage persists into Class 4 though it is also evident that math content taught during Class 4 has not been well mastered.
3. The data indicate rather clearly that the educational foundation developed in Classes 1-3 in the KEP program is not being further built upon by reverting to exclusively English-medium instruction in Class 4.
4. The Class 4 data indicate VERY CLEARLY that few students are mastering the content of the curriculum established for Class 4. This is probably for a cluster of reasons including language of instruction, lack of textbooks, inadequate time on task, low quality instruction, etc.
5. The data about the performance of the Fulfulde children living in the Kom area provides some perspective on possible educational outcomes and options for non-native speakers of Kom such as the Fulfulde children. The data reported here clearly indicate that Fulfulde children did much better in Kom-medium classes than they did in English-medium classes. The sociolinguistic likelihood is that such children have learned some or, perhaps, much Kom from living in the Kom community. For this reason, they are minimally compromised (in most cases) by participating in the Kom-medium program. It seems very likely, of course, that this would not have been the case if they had not learned the Kom language.
6. Gender differences are not much in evidence in the Kom area.
7. There is a small but persistent problem with under-age children especially in Classes 1-3. These children consistently under-perform indicating that they are not prepared for the level of work encountered in any of the classes. The problem is less evident in Class 4.

Finally, we will reflect briefly on some of the major issues encountered so far in terms of both lessons learned as well as reasons for continued concern for the state of education in the Kom area (Boyo Division). First, we find ourselves continually contending with the issue of time on task (total number of effective instructional hours in an instructional year). In its optimal form, the school year consists of 720 contact hours in the classroom. Several lines of research as well as direct personal experience make it clear that the effective number of contact hours in Boyo (and throughout the North West province) is more like 500 contact hours per year. This contrasts unfavorably with approximately 1,100 contact hours in the US educational system (primary level) and well over 1,500 hours in some Asian school systems. For this reason alone we would expect to see reduced learning outcomes.

Further exacerbating the whole time on task issue is a hugely overcrowded curriculum. At last count, the Class 3 curriculum mandated 16 subjects, many of which have little real relevance for children—gender issues, music theory, homemaking, environmental issues, gardening, crafts, etc.

Second, the educational weaknesses of an educational model dependent on instruction via a second language for both learners and teachers continue to be emphasized in our research data. This has been especially evident in the Class 4 data as we observed the low performance on a reading task (in English) of children who clearly know how to read. It is also starkly evident in the results of new topics in math which had been taught in English in Class 4 (fractions). Clearly, the children have not learned this material. The results lead one to wonder whether some of the teachers even fully understand the subject matter.

Third, in going through individual exam papers, we were struck over and over again by the evidence that children worked anxiously even desperately to find answers to math questions but lacked the procedural skill to do so. We could envision the frustration felt at trying to perform tasks one had not been adequately prepared to tackle.

The continuing low performance, even at the level of Class 4, on very basic math facts and operations suggests that students get very inadequate drill and practice in this regard. This seems to be a systemic issue having to do with time on task, teacher training, lack of materials, etc.

The results of the assessment of Class 5 next year will help answer many of the questions regarding the overall effectiveness of an early exit instructional model especially under the educational conditions prevailing in rural Cameroon.

References

- Bialystok, E. 1988. Levels of bilingualism and levels of linguistic awareness. *Developmental Psychology* 24(4), 560-567.
- Cummins, J. 1976. The influence of bilingualism on cognitive growth: a synthesis of research findings and explanatory hypotheses. *Working Papers on Bilingualism* 9, 1-43.
- Droop, M. and Verhoeven, L. 2003. Language proficiency and reading ability in first- and second-language learners. *Reading Research Quarterly* 38(1), 78-103.
- Galambos, J. and Hakuta, K. 1988. Subject specific and task-specific characteristics of metalinguistic awareness in bilingual children. *Applied Psycholinguistics* 9, 141-162.
- Thomas, W. and Collier, V. 1997. *School effectiveness for language minority children*. http://www.ncela.gwu.edu/files/rcd/BE020890/NCBE_School_Effectiveness.pdf Last accessed 14 July 2010
- Thomas, W. and Collier, V. 2002. *A National Study of School Effectiveness for Language Minority Students' Long-Term Academic Achievement*. Center for Research on Education, Diversity and Excellence, Santa Cruz CA. http://crede.berkeley.edu/research/crede/research/llaa/1.1_final.html Last accessed 5 August 2010
- Walter, S. and Trammell, K. 2010. The Kom experimental mother-tongue education project report for 2010. Unpublished research report. Available from the authors or SIL Cameroon.
- Walter, S and Trammell, K. 2009. 'The Kom MLE report for 2009', Unpublished research report. Available from the authors or SIL Cameroon.
- Walter, S. and Trammell, K. 2008. 'Results of the First Year of the Kom Experimental Education Project', Unpublished research report. Available from the authors or SIL Cameroon.