

The Lubuagan Mother Tongue Education Experiment (FLC)

A Report of Comparative Test Results

Stephen L. Walter, PhD, SIL International

Diane E. Dekker, MA, SIL International

presented to the

Committee on Basic Education and Culture

Committee on Higher and Technical Education

House of Representatives

Quezon City, Republic of the Philippines

February 27, 2008



Languages Matter! Celebrating International Year of Languages 2008.

Introduction

The Philippines is among the top ten countries in the world in terms of multilingualism according to Dr. Ricardo Nolasco, Komisyoner ng Komisyon sa Wikang Filipino. Multilingual diversity brings with it a depth of wisdom and beauty though it is often seen primarily for the accompanying challenges, particularly in education. Multilingual societies have the advantage of learning from multiple perspectives, sharing wider ranges of knowledge and wisdom. The Philippines as a multilingual society is changing rapidly as people of different languages move around, learning more languages and making different choices for language use within the home.

Within this multilingual setting the Philippines has chosen two languages that each Filipino needs to master in order to participate in wider society and global issues. Those languages are Filipino and English. The need to strengthen the acquisition of these two languages was the impetus behind the Lubuagan First Language based educational innovation. First Language or Mother Tongue based language approaches have proven beneficial in situations where only one second language is added in school. In the Philippines the need to add two languages and the desire to add both languages from the beginning is a unique difficulty which has received little attention from researchers. Typically where a third language is added it is only introduced later in the elementary years, rather than from the beginning.

This paper presents the *initial* findings of the longitudinal research project on a Mother Tongue based Multilingual Education innovation in Lubuagan Kalinga. The study paid particular attention to the educational implications of the early addition of two second languages to the curriculum.

For approximately ten years, the Lubuagan district has been host to an educational innovation in language of instruction known as the First Language Component (FLC). In this innovation, children receive instruction via their first or primary language for the first three years of school. During this period, children also receive instruction in oral English and oral Filipino as preparation for learning to speak and read in these languages as well. After three years, children in this experimental program merge with their peers in control classrooms into a classroom setting where English is the primary language of instruction.

For purposes of documentation and research, the FLC program has been launched in only a subset of local schools or classes. Each of these has been matched with a school or class assumed to be approximately equivalent in terms of SES, level of community support, and size.

Formal testing for research purposes has begun only recently with the first standardized test being administered in 2006-2007.

Description of the test and testing procedure.....

A standardized (or common) test was developed by a group of teachers and administered to students in control and experimental schools for all three years. The test was developed and administered primarily in English since children in the control schools had not had any experience in reading the local language. This format would seem to have been slightly biased in favor of English-medium students and schools but it seemed the most feasible way of testing all students.

The tests developed tested knowledge and skill in three areas—English, Filipino, and Math. The content was drawn from the curriculum common to all schools and classes. In Grades 2 and 3, test areas were further nuanced to distinguish between English speaking,

English listening, and English reading (the same for Filipino). The Grade 1 test included 60 items, the Grade 2 test 74 items and the Grade 3 test 85 items.

For Grades 1 and 2, the test questions were read by the test administrator to remove reading skill as a constraining variable. Grade 3 students read all questions themselves. Whenever possible, a multiple choice format was used to reduce subjectivity in scoring.

This preliminary report contains results for all children in both programs for Grade 3 and for children in all three grades in the only school where both programs exist side by side (because the school has two sections for each grade).

There are several reasons for this selective report. First, the researchers believe that later outcomes matter more as those better reflect how a given child will actually benefit from the educational experience. Second, the emerging body of evidence on the longitudinal impact of first language education suggests that the benefits are cumulative over time so that later measures are more telling than early measures. Third, in this as in most first language educational experiments, the early results are more sensitive to local variation such as teacher and school effects, significant differences in the ability and backgrounds of children, and uneven application of the respective models (both control and experimental).

More comprehensive results¹ will be available in the future as the program becomes more stable and the evidence from multiple years of testing begins to accumulate. The results from this current report should be taken as being suggestive of what can be expected in first language education programs. The reader is invited to engage publicly and intellectually with these findings as the country engages with the educational challenges entailed in a complex multilingual environment.

English Test Results

Worldwide, even in the developed countries, any proposal for an educational program in which children begin instruction in their first language (L1 or mother tongue) is greeted with public concern that such children are going to be educationally deprived. Much of this concern focuses on the supposition that such an approach will surely compromise a child's learning of the language of higher education and career opportunity. In the case of the Philippines, this concern is typically expressed in terms of learning English. There is, after all, a lot of intuitive logic to the assumption that the more a child is exposed to English in school, the better he or she will learn English. This may be another form of the conventional wisdom that if "some of something is good, more of that something has to be better."

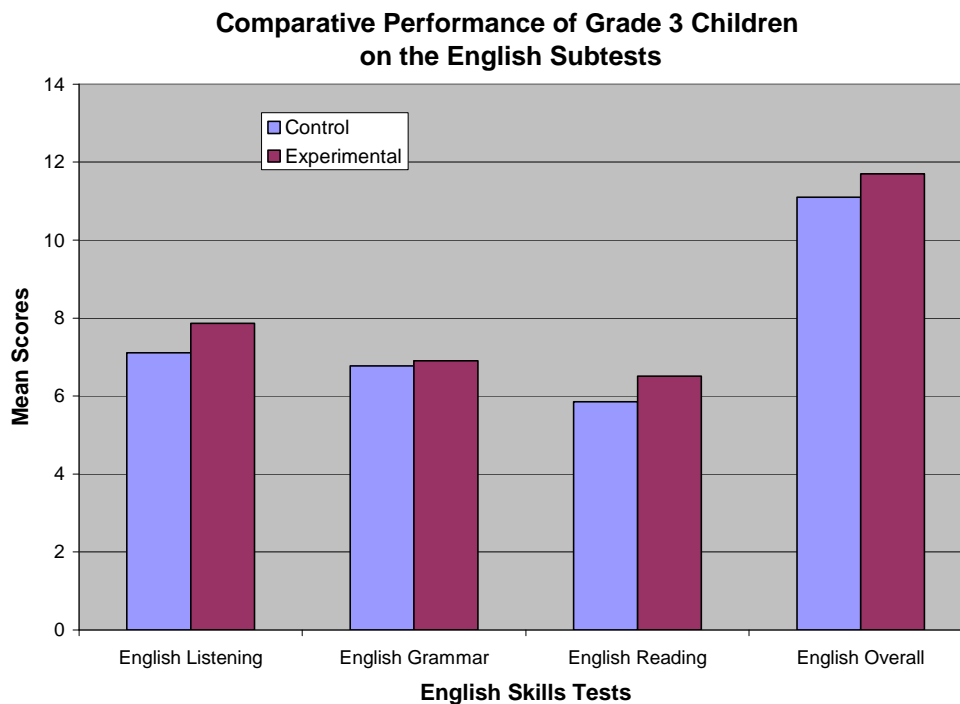
Educators and researchers interested in the "language of instruction" issue have compiled and are continuing to compile evidence that this popular wisdom is not valid in the case of language development in basic education. Large scale research carried out both in the US and in Canada during the last 30 years has provided compelling evidence that the critical variable in second language development in children is not the *amount* of exposure, but the *timing* and the *manner* of exposure. Thomas and Collier (1997), for example, tracked 42,000 children in the US who entered school not speaking English. These children had been placed into one of several programs which varied extensively in *how much* English language instruction children were exposed to and *when* they were exposed. In one of these programs, non-English speaking children were placed directly into English-medium classrooms thereby receiving ALL of their

¹ Statistical detail has been kept to a minimum in this preliminary report as it was prepared for a more 'popular' audience. Those who desire a full set of statistics may contact either of the authors through SIL Philippines for further information.

education in English. In other programs children received 3 years of initial instruction via their first language (with preparatory English language instruction) before moving on to an English-medium classroom. In the most extreme (and innovative) type of program, children received up to 6 years of instruction in their first language and while receiving, at the same time, 6 years of preparatory instruction in English before being inserted into English-only classrooms for middle school.

To the surprise and consternation of many educators (and parents), those children who received ALL of their education in English learned the least amount of English and scored the most poorly on nationally normed and standardized tests of academic achievement finishing, as a group, at the 11th percentile nationally (the bottom ten percent). They were also the most likely to drop out of school before finishing (see Thomas and Collier, 1997 for full details). Conversely, children participating in one of the 6-year Mother Tongue based programs actually completed their secondary education scoring well above the national norm for their **native English speaking peers**.

Because of the sharp contrast between "popular belief" and the findings or predictions of educational researchers, we tested Grade 3 children participating in the Lubuagan First Language Experiment to measure their English language development. The results of this testing are summarized in Figure 1.



Three components of English language learning were tested—English listening, English grammar, and English reading. A summary or composite score is also reported for English (English Overall). The purple or dark bar represents the mean score of all children in the experimental (FLC) program while the blue or lighter bar depicts the mean score of children in the control program (all instruction in English or Filipino). In the Lubuagan experiment, there are 3 schools in each program with approximately 120 children total in each program.

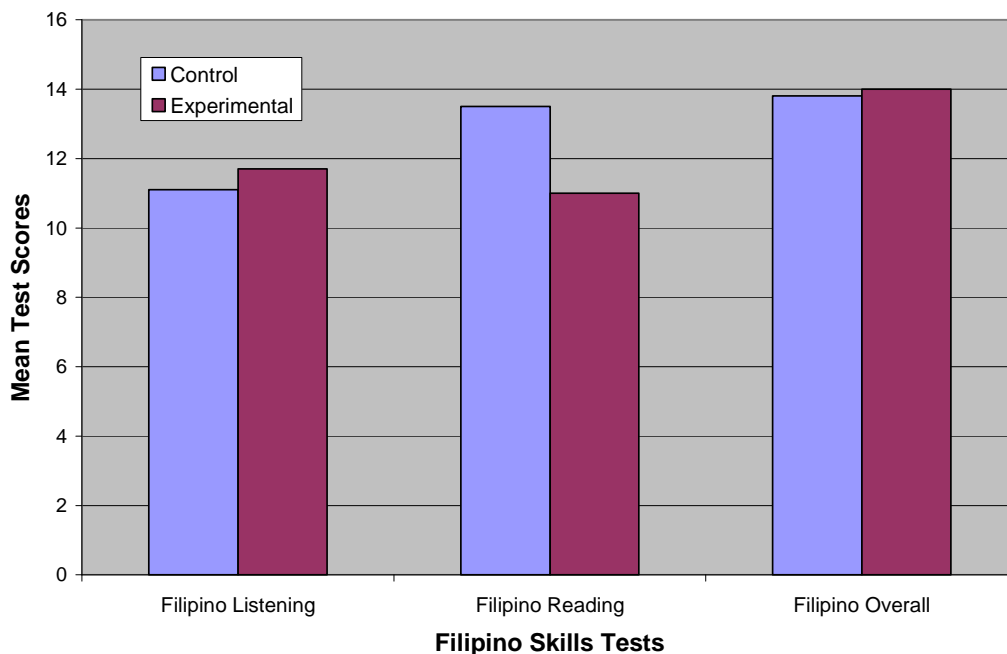
The test results for English language development show a small but consistent advantage for those children in the FLC program (in which most initial instruction in reading is in the mother tongue). In technical terms, none of the differences is statistically significant.

The test results clearly run contrary to the fear and concern of many parents (and educators) that children beginning in the first language are going to suffer educationally as a result. Not only are these children learning to read in their own language, they are actually learning more English than their peers who are receiving ALL² of their instruction in English. This finding is consistent with most of the carefully done research on multilingual education. Further, according to the findings of Thomas and Collier, IF the Lubuagan FLC experimental program were to be extended to grades beyond Grade 3, this advantage would continue to grow.

The Filipino Test Results

Since Filipino is also an important language in the Philippine context and a required feature of the curriculum, testing for Filipino language development was included in the research design. The results are reported in Figure 2.

**Comparative Performance of Grade 3 Children
on the Filipino Subtests**



The results of the testing on Filipino are generally comparable with those for English although, it seems reasonable to conclude that the two models are probably closer to being equivalent for Filipino language development. The Experimental group had a very slight advantage in Filipino listening while the Control group had a larger advantage for Filipino reading. The

² A small caveat needs to be made explicit on this point. In the experimental design, all of the children in the control program (the English-medium classes) are supposed to be receiving all of their instruction in English and Filipino. Informal feedback indicates that teachers do considerable code-switching in this program so that the model is probably not a true English-only program.

composite score (Filipino overall) showed almost identical results. Again, none of the differences is statistically significant.

Only tentative conclusions can be drawn from these mixed results. Probably the most defensible conclusion is that the use of the mother tongue as a medium of instruction is NOT compromising the ability of Lubuagan students to learn Filipino as both groups scored almost identically. (Whether either group scored "well" in Filipino is another and legitimate question to ask but one which depends upon someone supplying a definition of "well" before the question can be answered.) The slight but noticeable difference in advantage between Filipino listening and reading seems best explained by slight variation in teacher skill and emphasis during the instructional time given to Filipino.

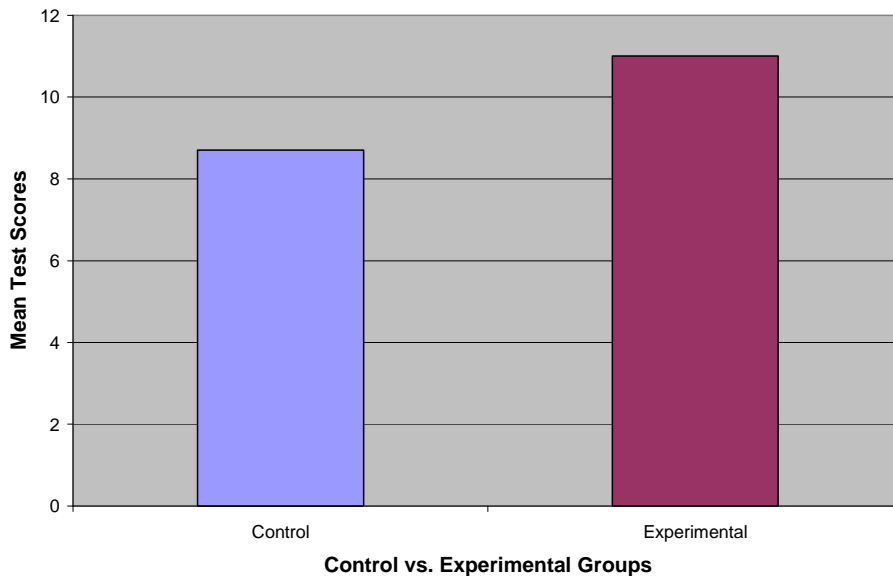
Math Test Results

A math test was also administered to both groups of students/schools. The results of this testing is probably of greater practical as well as theoretical interest given both the importance of math as an academic subject (and practical life skill) and the nature of the subject matter. Conceptually, math concepts are more abstract and "non-intuitive" than is language learning. In general, educators find that math is harder to teach and more cognitively demanding than is language.

Math is also of theoretical interest to researchers on the Language of Instruction question. Thomas and Collier have suggested that one of the fundamental virtues of first language instruction lies precisely in its superiority as a vehicle for dealing with (i.e., teaching and learning) more complex and subtle concepts (of which math would be just one example). The logic is simple. Grappling with subtle, nuanced, or complex material requires well-developed language skills for communicating shades of meaning as well as new concepts and models which often have to be introduced by means of analogy.

The content of the math test was taken directly from the Grade 3 curriculum. No effort was made to test reasoning skills about math concepts beyond the knowledge and skills taught in the national curriculum. The math test was not subdivided by domain so is presented as a single score in Figure 3.

Comparative Performance of Grade 3 Children on the Test of Math



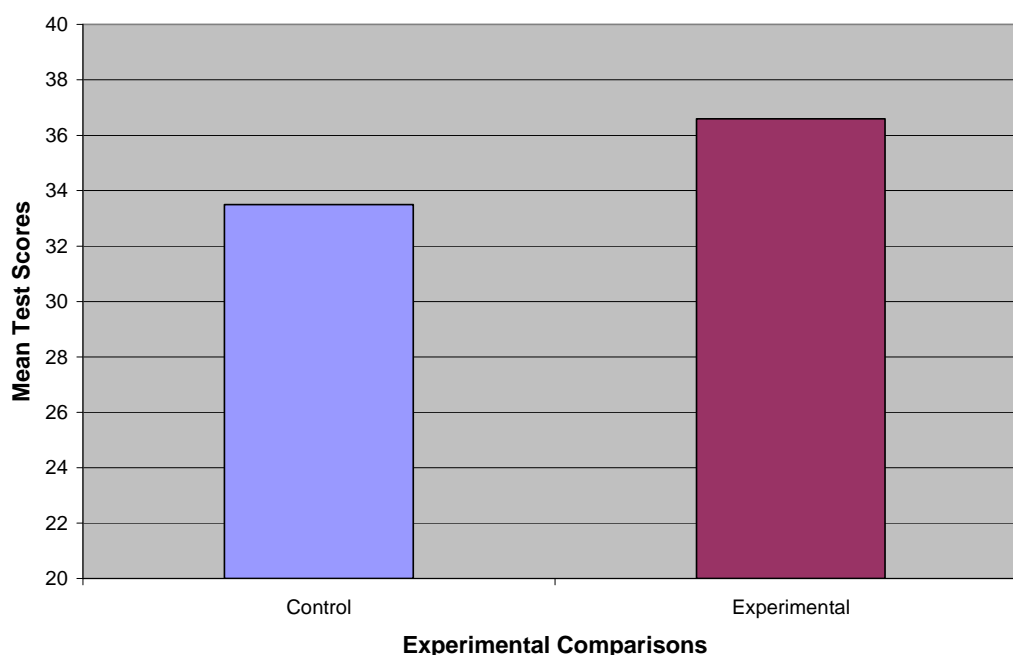
The comparative test of results on the math test between the control and experimental programs is slightly more marked than is the case for English or Filipino. The mean score for the control group was 8.7 (SD = 3.67) while that of the control children was 11.0 (SD=6.98). The difference is statistically significant ($T = 2.26$; $p = .026$) though the level of significance is relatively modest. The high standard deviation for the experimental children reflects some extremes in performance (generally due to an unfortunate classroom situation in one school).

This test result for math coincides nicely with the predictions of research on mother tongue education. Children cope better with the introduction of cognitively complex material when it is presented in their first language. Therefore, they are likely to do better on a test of math skills since math becomes more and more cognitively demanding as one progresses in school.

All Subjects

Probably the best general measure of relative performance is that of a combined score for all tests taken. Overall measures such as a combined score tend to neutralize differences due to teacher interest in a particular subject or uneven time given to a particular subject in the curriculum. Figure 4 provides the broadest available measure comparing the academic development of children in the two models.

Comparative Performance of Grade 3 Children On the Combined Measure of All Subjects



On this combined measure, the children in the experimental program scored somewhat higher than did the children in the control program. The relevant statistics are given in a footnote below.³ The graph shows a somewhat higher performance for the experimental group, a finding which is generally consistent across the curriculum for Grade 3 students. As indicated in the footnote, the difference, while real, is not statistically significant.

What conclusion, then, can be drawn from the testing done in this program? With a small number of schools in each program and a small scholastic population, one needs to exercise caution in making broad generalizations or assertions based on this preliminary data set. The reader is advised to view these results as suggestive, not definitive. On the one hand, the results are consistent with the general body of research on mother tongue education. The foundation for learning created by getting a good start via the first language begins to manifest itself in a fairly consistent way in Grade 3. On the other hand, the differential is not so great that one is immediately swept away by the overwhelming superiority of one approach over another. Nonetheless, it is important to keep in mind that the weight of the research evidence accumulated elsewhere suggests that this differential in performance would grow as the curriculum becomes more demanding.

The two conclusions which we feel to be best supported by the data are as follows. First, children are not compromised in their educational development by beginning in first language contrary to widely expressed concerns that this is apt to be the case. Second, the subject for which the greatest benefits are derived by the first language program is the one most

³ The mean for the control population = 33.53 with a SD of 9.55. The mean for the experimental population = 36.63 with a SD of 15.28. The population size is 60 in both cases. The difference in performance is not statistically significant ($T = 1.33$; $p = .185$). The test was done assuming equal variance.

cognitively demanding—math. Given the strong national concern about lagging performance in this area, the evidence from this experimental program should be encouraging.

A Within-School Comparison – the Evidence of the Lubuagan Central School (LCS)

In developing countries, it is not uncommon to find extreme variation in educational outcomes from one school to another. Very often, these extremes are due to variation in the quality of instruction though other variables are often present as well including patterns of attendance, attitudes of parents, extreme linguistic diversity, etc.

In the experimental FLC program being presented in this report, there is one school which is large enough to have two sections for each lower grade. This school has agreed to enroll one section in the experimental program and keep one in the traditional program. This school is the most established of the schools in the research site and is considered the standard bearer for education in the area. It also has the most experienced teachers and the best equipment. In addition, the school administration has agreed to randomly assign incoming students to the two sections though it takes into consideration parental requests when these are made.

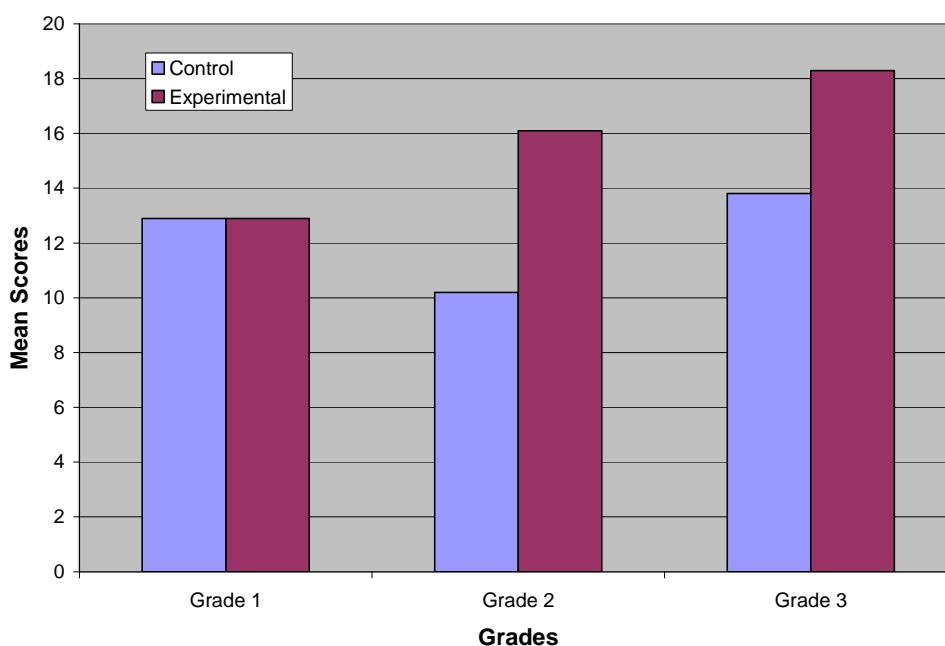
Under these circumstances, we have decided to report results for all three years of testing from this school. This is being done not because the results are especially good or bad for one program or the other but because we believe the results of the testing within this school best exemplify the "normal" pattern of what one can expect from implementation of a first language component in the Philippine context. In doing so, we must emphasize again that sample size is very small (approximately 20 children per class) so the results have to be treated as being illustrative, not definitive. They suggest what it is reasonable to expect, but cannot be considered as proof of anything.

English

In most first language programs like the Lubuagan program, it is normal for children in Grade 1 receiving instruction primarily in English to score better on measures of English than children receiving all or most of their instruction via the first language—Lilubuagan in this case. In some first language programs, children receive no instruction at all in the second language—English in this case. Not surprisingly, such children do very poorly on a measure of learning achievement about subject matter to which they have had no exposure.

In the case of the Lubuagan program, the Grade 1 curriculum puts a heavy focus on learning to read in Lilubuagan but does not ignore the second language. Children receive considerable instruction in oral English and even some instruction in learning to read English—perhaps more than is appropriate in a first language model. Because of the substantial time devoted to English in Grade 1—even in the first language component—these children were tested for achievement in English. The results of this testing along with the testing done in subsequent grades is presented in Figure 5.

English Language Development In Lubuagan Central School



Several points can be made—albeit carefully—about these results. First, there is no difference in English development between control and experimental children in Grade 1. This is contrary to the normal finding that control children would be considerably ahead of experimental children after Grade 1 since the experimental children normally receive little or no instruction in the second language in Grade 1. This somewhat unusual result probably reflects the greater than normal amount of attention given to English even in the experimental program.

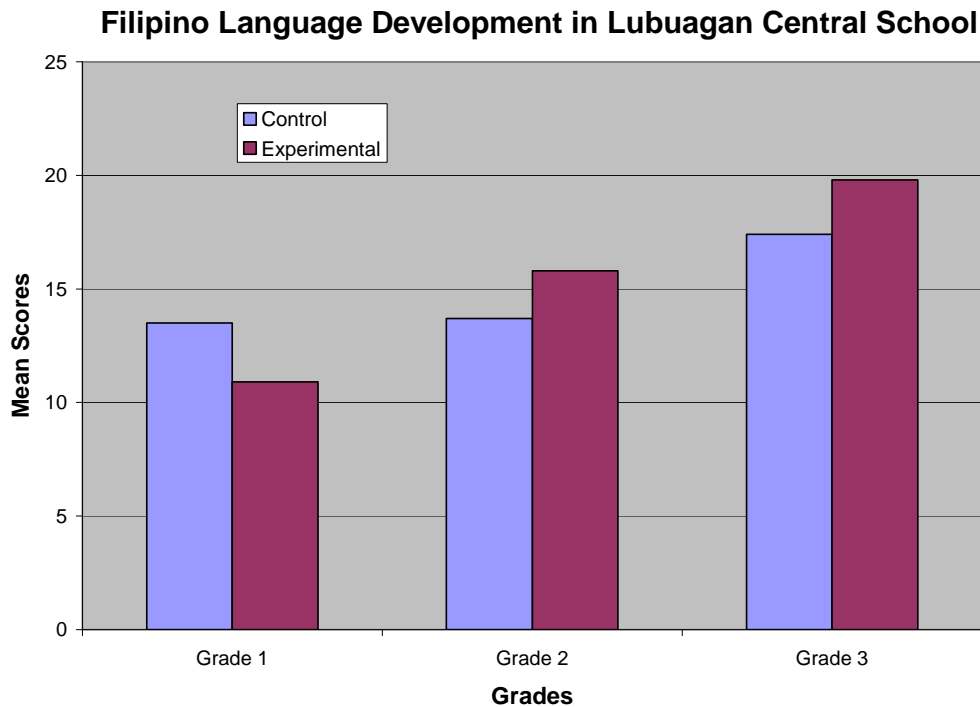
Second, after Grade 1, children in the experimental program begin moving ahead of their peers in the control class **even though** English is not the language of instruction for those children. Again, we suspect this to be due to the heavy attention given to second language development in the national curriculum.

Third, we note what appears to be a steady, positive growth in English language development in the experimental children while, for the children in the control program, the growth path is somewhat erratic. The large differential in Grade 2 is not consistent with the results of most similar programs. We believe this difference to be largely artifactual (an accident of circumstances) even though it is statistically significant ($p < 0.000$). The Grade 3 result is more consistent with theory and with what we expected to happen. Curiously, there is some visual evidence in the graph that in Grade 3, the control students are beginning to catch up with the experimental children in English language development. Since there is no Grade 4 in the program, we cannot verify the possible validity of this observation.

Fourth, the most important observation is that made earlier—the children in the experimental program are not being compromised in their English language development. They are actually doing **better** than their peers in the control classes.

Filipino

Children in most schools in the Philippines are actually being expected to learn two second languages simultaneously. While English is a more prestigious language in the country and more widely used in higher education, Filipino has also been designated a national language and is a required part of the curriculum all over the country. In the Lubuagan area, English is used frequently but Filipino usage is rare (Ilocano is the lingua franca of the area). Therefore, we expected to find a somewhat different pattern in the learning of Filipino in the classrooms of Lubuagan. The results of the testing done are presented in Figure 6.

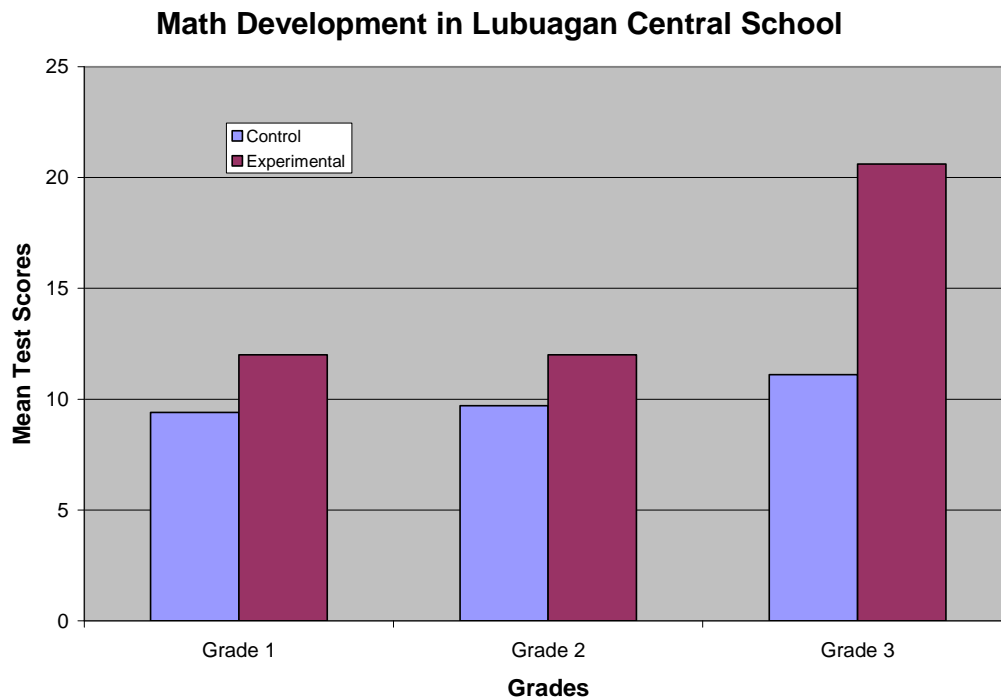


The profile of results for Filipino is actually closer to what we would have expected to find for English language development. In Grade 1 there is a clear advantage for children in the control program. In Grades 2 and 3, there is a slight advantage for the experimental children. Only in Grade 1 is the differential statistically significant ($p = 0.029$). We do not have a lot more to say about these results other than to observe again that the data do not support the fear of parents and others that beginning instruction in the first language is going to compromise development of ability in a second language.

Math

As indicated elsewhere in the paper, the subject of math is a good one to test hypotheses about the need to have a solid language foundation in order to deal with the more cognitively demanding material of the higher grades. Thomas and Collier (1997) have proposed that academic development, language development, and cognitive development *must* proceed in parallel or in lock step. The typical child who receives most or all of his/her educational instruction in a language not well understood never develops the cognitive tools needed to master more cognitively demanding curricular content. Therefore, the argument goes, children receiving first language instruction in a language they already know are building that

foundation and, all things being equal, are more likely to do well in cognitively demanding subjects. The relevant test data for math instruction in the Central school are set forth in Figure 7.



The basic pattern demonstrated in these data is a slight advantage for the experimental children in Grades 1 and 2 and a large advantage in Grade 3. The differentials in Grades 1 and 2 were statistically significant but not strongly so. The difference in Grade 3 is markedly so ($t = 6.89$; $p < 0.000$). Grade 3 children in the experimental program scored almost twice as high as their peers in the control program.

The results in Figure 7 provide vigorous support for the hypothesis that a good learning foundation (via the first language) better prepares children for dealing with cognitively demanding curricular content. It would be nice if we had further evidence from a Grade 4 or Grade 5 but, unfortunately, the program does not extend to that level.

The cautionary note which must be kept in mind when interpreting these results is the fact that the two classes involved were small. It is possible that a group of outstanding students or an outstanding teacher account for this result. Evidence from other testing done (reading in first language), however, tends to disconfirm this possible interpretation as the Grade 3 students from LCS were very similar to their peers in other schools on a test of mother tongue reading achievement.

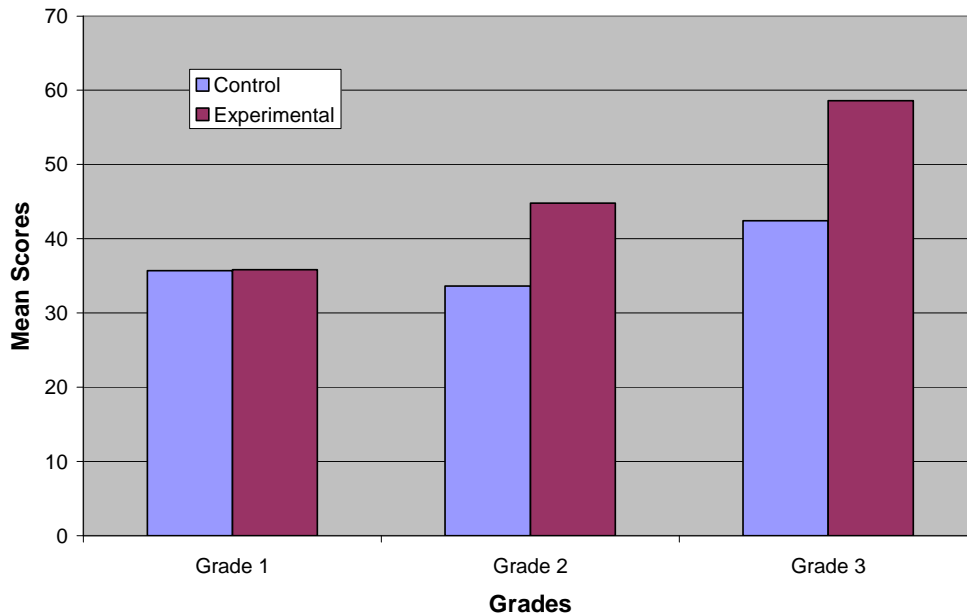
Overall or combined achievement

As in the case of the Grade 3 results reported earlier, we are probably on safer ground using a composite of educational achievement when trying to evaluate the extent to which the first language component may be contributing to improved educational outcomes in this program. The potential for testing bias is greater when a single measure (a single subject) is used as a basis of comparative performance. With a broader and more encompassing measure

we increase the likelihood that we are looking at a "model effect" rather than a "teacher effect" when comparing results.

Figure 8 presents a composite measure⁴ for each of the three grades of LCS and will be the final piece of evidence presented in this report.

Composite Measures of Educational Achievement for the Three Grades of the Central School



Notably, in Grade 1 performance was almost identical. In Grade 2 there is a strong advantage for the experimental students with this advantage increasing even further in Grade 3. The differential advantage is statistically significant in both Grades 2 and 3 ($p < 0.000$ in both cases).

Most first language projects have reported somewhat more ambiguous results. More typically, control children in Grades 1 score more highly than experimental children on subjects other than first language reading. Typically, that difference is reduced in Grade 2 and perhaps reversed by a small amount in Grade 3. The performance of children in this one school is more positive with parity in Grade 1 and a significant advantage for experimental children thereafter.

Since Figures 5-8 present results from a single school, we remind the reader again that we see these results as illustrative rather than definitive based on just one years worth of test results. Nonetheless, the results for this one school, when juxtaposed with those for the entire program should provoke useful discussion among educators about the practical and conceptual merits of first language instruction. As researchers, we are keenly aware of the limitations of a constrained data set. Nonetheless, we have agreed to make these results public in the interest of provoking dialog within the country on the important issues surrounding policy choices about language of instruction in the nation's schools. We believe that informed debate proceeds on a more rational track when empirical data is available to test positions rather than basing one's arguments on conjecture, anecdotes, and subjective experience.

⁴ The composite measures were computed by simply adding together the scores of each of the separate tests given. No weighting or any other transformation was used in the preparation of this measure.

This limited data set should not and can not be the sole basis for quick and radical policy adjustments. Similar research can be and should be done elsewhere in the country to test the validity of our findings. We believe that sound policy should rest on sound research.

References

Thomas, Wayne and Virginia Collier. 1997: School effectiveness for language minority children.
www.ncela.gwu./pubs/resource/effectiveness.