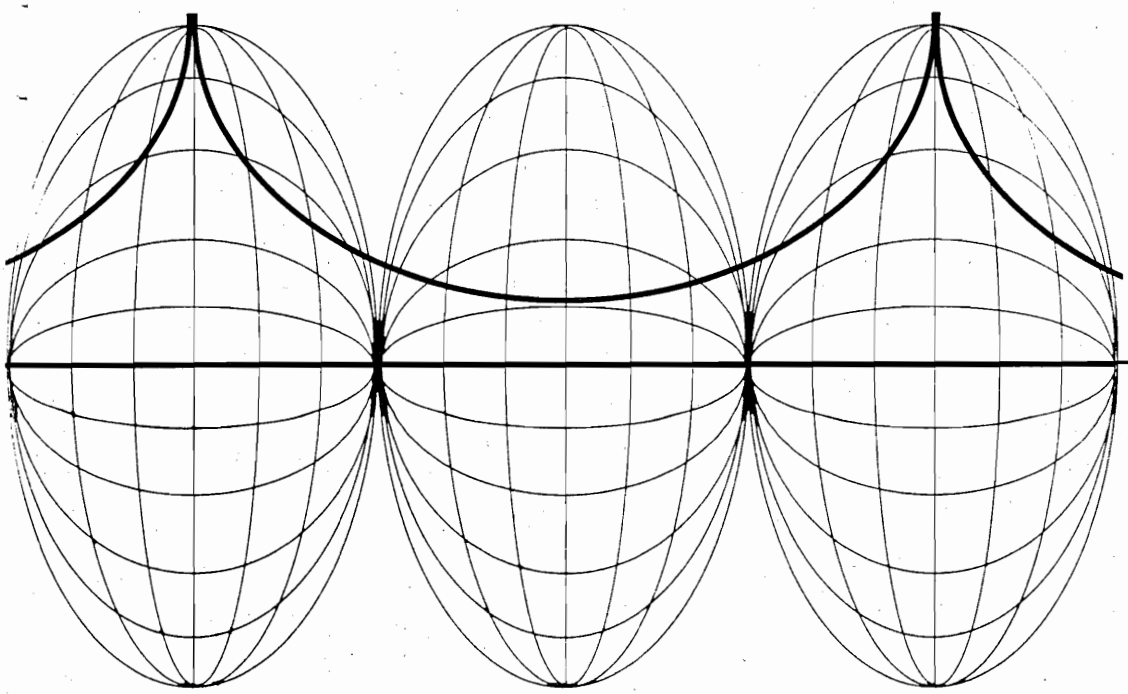


B • R • I • D • G • E • S

Basic Research and Implementation
in DevelopinG Education Systems

CASUAL PAPERS

April 11, 1990

Empirical Results and Conventional Wisdom: Primary School Effectiveness in Burundi

Thomas Owen Eisemon
McGill University

John Schwille
Michigan State University

Robert Prouty
Adventist University of Central Africa

Francois Ukobizoba and Deogratias Kana
Centre de Perfectionnement et de Formation
en Cours d'Emploi (CPF)
Burundi

Gilbert Manirabona
Ministry of Primary and Secondary Education
Burundi

Acknowledgements

The research reported in this paper was supported by the USAID BRIDGES Project administered by Harvard University (DPE 5824-A00-5076-00). The research was planned and carried out by the Centre de Perfectionnement et de Formation en Cours d'Emploi (CPF) in collaboration with Michigan State University and McGill University. Miss Heather Usher and Mr. David Williams, graduate students at McGill, assisted in the data analysis. Many valuable suggestions and comments on particular aspects of the analysis were made in consultation with Drs. Michael Abrahamowicz, Carl Frederiksen and Socrates Rapagna, McGill University, and William Schmidt and Richard Houang, Michigan State University. The authors take responsibility for the remaining shortcomings. The views expressed are not necessarily those of the participating institutions or funding agencies.

Abstract

This paper examines the effects of language policies, school management and instructional practices on student achievement in primary schools in Burundi. Data were obtained from a probability sample of forty-seven schools in twenty-four school directorships in rural areas, and from about two thousand (1,946) grade six students in 1989. School directors and teachers were surveyed and the students tested in language arts, mathematics and science-agriculture. Subsamples of students took these tests in French while other subsamples were administered the same tests in their mother tongue, Kirundi.

Language of assessment was found to affect the measurement of achievement, especially for students of high ability. The type of school a student attended (i.e. whether it was a core or a satellite school) had little influence on student test scores. Several school and instructional characteristics were shown to relate to student performance, and a model of school effectiveness is presented. Relationships between frequency of teacher supervision by the school director, instructional practices that increase opportunity to learn and measures of teacher skills are investigated and shown to affect student performance in various ways. Among student characteristics, repetition in grade six was strongly and positively related to student performance, especially on those tests administered in French that were most comparable to the secondary school entrance examination.

The findings of the paper are in some ways a vindication of conventional thinking about what makes for effective primary schooling in Burundi. They attest to the value of close teacher supervision by school directors, the positive effects of repetition, and the importance of such indicators of

teacher engagement as punctuality. But in other respects the findings deviate from conventional wisdom about how language of assessment influences the measurement of individual differences or about how variations between core and satellite schools affect student achievement. The paper concludes with a call for focused experimentation to identify effective strategies for raising levels of student achievement, improving monitoring of student learning, and increasing internal effectiveness so that the progress already made in expanding access to primary schooling can be consolidated and reinforced.

Introduction

While a great deal of attention has been given to documenting the poor quality of African primary schools and the need to improve them (World Bank 1988), the supporting evidence is often drawn either from international studies of educational achievement or from national research which uses a metropolitan language to measure achievement. African students, it seems, have not learned very much at primary school or even in secondary schools which select students on the basis of rigorous national examinations that are generally administered in a metropolitan language used for instruction at the upper primary level in thirty-two of the thirty-nine countries surveyed in a recent World Bank report (World Bank 1988,154-156). What such results indicate is unclear. According to Berry, two types of language related learning problems occur in African schools; type A problems reflecting poor metropolitan language proficiency, and type B problems that result from the "distance between the cognitive structures natural to the student and implicit in his mother tongue and those assessed by the teacher" (Berry 1985, 20). Type A and type B problems are, of course, reinforcing since much of the knowledge transmitted through instruction in a metropolitan language is incongruent with the knowledge obtained from social experience. Measurements of student achievement in a metropolitan language and in the mother tongue have been shown to produce different estimates of learning outcomes in African countries (e.g. Zepp 1982). More proficient metropolitan language learners possess more linguistic and substantive knowledge to assimilate new information and carry out information processing tasks, contributing to higher level performance measured in the metropolitan language. Greater metropolitan language proficiency is usually thought to facilitate the transfer between the

mother tongue and the metropolitan language. However, previous research suggests that there may be little integration of knowledge (Lemon 1981) and cognitive skills (Zepp 1982) because of mother tongue "interference." A better understanding of how language may affect learning outcomes and measurement of individual differences is central to the objectives that have guided the research reported in this paper.

Much has also been said about poor quality as a consequence of primary school expansion. There are essentially two ways through which access to primary schooling may be dramatically increased given the financial constraints faced by most African countries. One involves rapid expansion of the educational infrastructure with inevitable deterioration of the more conspicuous indicators of school quality: large class sizes, poor school facilities, shortages of instructional materials, a high proportion of untrained teachers, infrequent supervision and inspection, and so on. Another approach seeks to maintain some semblance of school quality by slowly expanding facilities and resources while making more efficient use of the infrastructure through, for example, introducing double shifts to increase enrollment. This provides greater access to primary schooling without a proportionate expansion of instructional inputs at the cost of reducing instructional time and increasing teachers' and school directors' workloads. Burundi has followed the latter strategy and increased the rate of primary school participation by 256% in the six years following 1980-81 (Eisemon, Schwille and Prouty 1989).

In order to manage such expansion, the Ministry of Primary and Secondary Education, like some others in Africa, has grouped or clustered schools for administrative purposes. In other words, school directors, particularly in

rural areas, are responsible for several schools. The school director is based at a core school which offers a full cycle of instruction. Satellite schools, which provide incomplete or complete programs of instruction, may be located at considerable distance from the core school. Differences between core and satellite schools in instructional resources and in implementation of ministry policies are thought to have a powerful influence on student outcomes. The impact of these school characteristics on student performance will be examined.

We will be concerned as well with what contributes to effective instruction at the classroom level. Effective teaching is, of course, difficult to define for any educational system. But for countries that use national examinations to allocate educational opportunities, a teacher whose students have a record of success in examinations is viewed as successful. This formulation may ignore many things though it captures what is significant for students and parents as well as many teachers and educational administrators. Teacher qualifications and experience may be as important for successful examination preparation as are, perhaps, inservice teacher training and inspection (Mwamwenda and Mwamwenda 1989). Since these teacher characteristics and administrative practices can be manipulated by educational policies, they are of much interest to policymakers (Alexander and Simon 1975). Yet, the effects of gross manipulations of policies affecting, for instance, years of teacher training on student achievement are uneven and often ambiguous (Heyneman and White 1986, 49). Not surprisingly, efforts to correlate achievement with more subtle indicators of instructional quality such as use of time, curriculum coverage and monitoring of learning suggest that "some elements of effective teaching are common cross-culturally, while

others may be culture specific" (Lockheed and Komenan 1989, 15). For African countries with influential examination systems, teacher understanding of assessment tasks, sources of student errors and strategies for remediating errors should also be considered important features of instructional quality.

Finally, we will investigate the impact of student repetition on achievement. Depending on how internal efficiency is defined, repetition may either enhance or erode it. Repetition increases the number of years it takes a primary school to produce a graduate and, thus, wastes instructional resources. On the other hand, repetition may increase student achievement and in doing so, bring more students up to a level of learning expected of primary school graduates. In Burundi as in many other African countries, repetition is largely voluntary, i.e., students repeat to improve their chances of educational success. Some authorities (cited in Schwille, Eisemon and Prouty 1990) assert that repetition simultaneously wastes resources and may have only a negligible impact on student achievement. That is the position of many African ministries of education and of donors to African education. High repetition rates are often represented as an indicator of poor educational quality: "poor quality teachers, too few textbooks and other relevant learning materials, inadequate guidelines for assessing learning, too few national examination(s), and too few well-trained educational managers and school directors" (Heyneman and White 1986, 48). Student repetition is prohibited in some African countries like Kenya, and strongly discouraged in others. Nevertheless, according to a World Bank (1988) report, 16% of primary school students in Sub-Saharan African countries were repeaters--23% in Francophone countries. In Burundi, it is not unusual for students to repeat the sixth grade two, three or even more times.

The Policy Context

In the 1960s when African countries were becoming independent, school expansion was presented as an instrument through which new states might acquire legitimacy and demonstrate capacity to govern (Almond and Coleman 1960). In making primary schooling widely available, governments could rapidly bring the polity within the influence of the modern state and, simultaneously, provide evidence of its ability to deliver public goods. Unlike most African countries, Burundi moved slowly to expand primary schooling. When the government did begin to rapidly expand primary schooling in 1982, it demonstrated political capacity in increasing rates of school participation and, significantly, in maintaining the quality and control of expenditures on instructional inputs. Nevertheless, these accomplishments have created many problems which threaten to erode much of what has been achieved.

These problems stem from the discordance between the objectives that have guided the development of primary schooling since independence and the means adopted to expand schooling and raise levels of educational attainment. The objectives set out in the 1973 educational reform continue to be cited in support of government policy at the primary level. The reform presented a plan to make six years of primary schooling universal and to orient instruction to the teaching of the vernacular language and skills to improve rural life (Greenland 1974). Educational practices inherited from the colonial period were criticized as elitist and alienating (Ministry of National Education 1988a, 6). Government schooling, reserved formerly for an elite which received academic training in French in preparation for entry into colonial administrative or technical cadres, was to be made available to all.

That meant diminishing the role of voluntary organizations in the provision of schooling. Church schools had provided instruction in Kirundi for religious participation as well as practical training in modern agriculture. This was now the responsibility of government schools. The Kirundisation and ruralisation of primary schooling were key features of the 1973 reform (Ministry of National Education 1988a).

Ethnic conflict and other obstacles delayed educational expansion until the early 1980s when the government moved swiftly toward universal primary schooling. It might have implemented expansion through policies intended to give communities more responsibility for financing schooling. Indeed such a strategy was recommended in the 1973 reform which envisaged primary schools becoming community institutions and foci for rural development much like the ujamaa schools in neighboring Tanzania. However, the government, while committed to expanding access to primary schooling, has not yet succeeded in implementing the community school concept.

The introduction of double shifts was combined with a policy of collective promotion to increase educational attainment. As a consequence, school enrollments grew dramatically. Another consequence was that the curricular changes proposed in the 1973 reform could not be easily implemented (see Ministry of National Education 1988b).

The policies of Kirundisation and ruralisation were predicated on two unrealistic assumptions. The first was that access to schooling could be increased without raising aspirations for academic rather than practical training, and for instruction in French, the language of secondary and higher education, instead of Kirundi. Many parents, it turned out, did not send their children to government schools to learn vernacular languages and become

better farmers. They wanted an education for their children that would afford them a much better life off the farm, preferably in government employment. That usually required a secondary education and knowledge of French.

Second, it was assumed that the curricular objectives of the 1973 reform would continue to guide instruction despite less instructional time, higher class sizes, and heavier teacher workloads that resulted from the introduction of double shifts and collective promotion. But in fact the scope of the school curricula had to be reduced. The teaching of agriculture was one casualty. The amount of agriculture instruction was cut to one period of thirty minutes per week (Eisemon 1989). The teaching of agriculture in Burundian schools now consists chiefly of supervised work in the school gardens. Mulching, weeding, and other routine agricultural skills are emphasized, most of which are already familiar to students (Eisemon, Schwille and Prouty 1989).

Kirundi was intended in the 1973 reform to become the medium of instruction at the primary level and, eventually, of lower secondary schooling. However, French has been retained as the medium of instruction for grades five and six due largely to fears that students would be unable to cope with French in secondary school (Ministry of National Education 1988b). In 1989 lack of student facility in French led the government to begin instruction in the metropolitan language in the first grade. Previously, French was taught as a subject from the third grade.

Educational quality has been equated with achievement in academic subjects which are assessed mainly in French in the rigorous <u>concours national</u>, the national secondary school entrance examination. The small proportion of students admitted to secondary schools on the basis of their

examination results, less than 10%, has not increased since the introduction of double shifts, though the number of entrants has increased in absolute terms. The examination orients instruction at the upper stage of the primary cycle (Eisemon, Schwille and Prouty 1989). The concours national is, to reiterate, a secondary school entrance examination rather than a primary school leaving examination. Primary school teachers and school directors do not participate in the selection of examination questions or the marking of answers, and until 1989, individual student scores were not announced. Successful students were simply notified that they had been awarded a place in a secondary school. Nor were aggregate score or item statistics published. These measures have ensured that the examination is viewed as an almost lottery-like event, generating keen interest in participation on the part of students and their parents, and strong pressures on teachers and school administrators to prepare students for the examination.

The concours national consists of examination papers in French, Kirundi, mathematics and a combination of science and social studies. Two hundred points are awarded. French and mathematics are the most important papers, accounting for 80% of the possible marks. Examination papers are constructed to identify the most able students, not to measure what most students have learned. The mathematics paper, for example, includes items relating to topics in geometry that many North American high school students would find difficult, such as this one from 1987: "The perimeter of a rectangular garden measures 1,890 meters. Knowing that three fifths of its length equals three quarters of its width, calculate its area in ares [hundredths of a hectare]."

A consequence of the selectivity of the concours national is the high reported rate of student repetition in grade six: 51% in 1988-89. Ministry policy has called on schools to limit repetition to an overall 10% average across all grades. But the overall national repetition rate for primary school in 1988-89 was 25.7% (Proceedings 1989). As students are not given unique inscription numbers when registering for school, repetition cannot be easily monitored. Students wishing to repeat grade six can enroll in other schools. Collective promotion of students reduces wastage in the lower stage of the primary cycle but increases repetition in the final year.

The absence of a strict control over repetition in grade six is curious in an educational system that is otherwise highly managed and well controlled. The explanation may have to do with the fact that the examination is the principal determinant of a student's economic future. No student is prevented from repeated failure, from experiences likely to convince parents and students that the causes of poor performance have to do with individual effort and ability rather than with a selection mechanism that does not afford sufficient opportunity to succeed.

Of course, the quality of instruction students receive also affects their performance on the concours national. At the school level, instructional quality may be related to an important organizational characteristic, namely, whether a school is a core or a satellite school within a school cluster (or directorship as it is locally known). In 1988, almost two thirds (65%) of primary schools were satellite schools (statistics furnished by Ministry of Primary and Second Education, 1989). Core schools usually have two or more satellite schools attached to them; sometimes up to six schools. A small proportion (10%) do not. Many satellite schools do not

offer a complete program of instruction though, typically, at least one satellite school in each school directorship offers six years of instruction.

Core schools are often older schools (or new well constructed schools) with office space for the school director, staff residences or rented modern housing, water, and in many cases are in close proximity to a tarmac road. In brief, core schools are the most desirable places to live and work. Satellite schools may be located at considerable walking distance from the core school, often between six to twelve kilometers. School directors are responsible for supervising instruction at satellite schools. According to Ministry guidelines, they are supposed to make one hundred and fifty classroom visits per year to core and satellite schools, and organize inservice teacher training. Textbooks, teachers' guides and other school supplies are to be distributed to schools on the basis of need. Yet the clustering of schools for administrative purposes gives much scope for wide variations within the cluster in the availability of inputs, inservice training and especially teacher supervision. These advantages of core schools may be exacerbated by patterns of teacher recruitment that are likely to favor core schools. These schools, in turn, have greater access to substitute teachers to cover absences. Teacher absenteeism is a serious problem, particularly in isolated satellite schools where it is more difficult to replace teachers on maternity leave.

The desire to expand enrollments prompted the government in the early 1980s to introduce double shifts which decreased the amount of instruction students receive to three and one half hours a day. Double shifts increased efficiency in the utilization of staff and facilities but also increased requirements for other resources that support teaching, including

instructional materials. School fees of about US $1.70 are assessed for purchase of student textbooks. These are available only for language arts subjects, i.e. French and Kirundi. Teacher guides (fichiers) are used for other subjects even in the upper stage of the primary cycle. They are not intended simply as resource materials. The fichiers provide detailed prescriptions for each lesson and include the examples to be used for explanation, questions to be asked of students and exercises that teachers are to assign. They are designed to reduce lesson preparation, encourage class participation to develop students' language skills and to ensure uniformity in the instruction students receive.

Although all primary schools are given a common curriculum to follow and teaching is circumscribed by the fichiers, there is still opportunity for alteration of the program of studies, especially in sixth grade when students are preparing for the concours national. For example, teachers and principals may increase in several ways the amount of instruction students receive in the subjects tested in the concours national. They may reduce the number of lessons given in subjects like agriculture. Teachers may combine morning and afternoon shifts, teaching both groups together all day. Or they may increase the length of instruction by beginning classes earlier and/or ending later. Some of these measures deviate from the guidelines school directors and teachers are given to follow. Enforcement is problematical given the importance that is placed on student performance in the concours national.

What emerges from the foregoing are many unreconciled objectives, policies and practices. Access to primary schooling has been expanded through double shifts and educational attainment raised by collective promotion. But this has not been accompanied by a proportionate increase in opportunities for

secondary education. While primary school is the terminal stage of schooling for most children, instruction in the upper stage of the cycle is oriented to gaining admission to secondary school. Selection mechanisms and double shifts do not support the curricular objectives of the 1973 reform. Although emphasis is placed on efficient utilization of instructional resources, the selectivity of the concours national fosters grade repetition. Policies adopted to facilitate school expansion produce qualitative variations in instruction notwithstanding efforts to ensure uniformity, and so on.

Study Design

The research reported here expands on the findings of pilot studies of three schools carried out by the BRIDGES project in 1988. Interviews with school directors and teachers as well as classroom observations drew attention to many school, classroom and student characteristics that may affect student achievement. These include syllabus coverage, the use of French as the medium of instruction, the amount or quality of examination preparation students receive, and student repetition rates (Schwille, Eisemon and Prouty 1989; Eisemon and Schwille in press; Eisemon 1989). Official grade six repetition rates in the three pilot study schools ranged from 30% to 70%. In addition, although the three schools included in the pilot studies were core schools, some information was obtained from satellite schools suggesting that the qualitative variations between types of schools might be larger than those between core schools. The pilot studies further involved some testing of student knowledge of science and agriculture as well as comprehension skills in French and the vernacular language, Kirundi. The results indicated that student performance should be assessed under different language conditions to disentangle outcomes of instruction from the effects of language proficiency.

Target Population and Sampling

To examine the issues raised by the pilot studies, a larger study was planned with a more representative target population. Sixth grade students and teachers in predominantly rural areas of the country were the focus. Of the 31 school cantons in the country, 3 were excluded as predominantly urban, 2 were excluded because they were still disrupted from recent ethnic violence and 5 were excluded because they were judged to contain too many inaccessible schools. It seemed preferable to exclude such cases from the target population from the beginning rather than to draw a probability sample which could not be surveyed without undue replacement. In 1987-88, these excluded cantons accounted for 18,480 of the 64,910 sixth graders in the country.

From the remaining 21 cantons a multistage cluster sample was drawn by probability methods. In the first stage, eight cantons were selected from a list of cantons stratified by region with probability proportionate to size of the sixth grade population. In the second stage, from each of the 8 cantons, 3 directorships (i.e. school clusters) were drawn with probability proportionate to size from a list of directorships, stratified by the sixth grade repetition rate in the core school. Within each directorship, in order to be able to contrast core and satellite schools, one core class and one satellite class were selected, using a random procedure whenever there was more than one school or class to be chosen. However, in five of the directorships there was no satellite school and in these cases, two classes from the same school were chosen except in one small school which had only one sixth grade class. The sampling was executed virtually according to plan; replacement was limited to one satellite school which turned out not to belong to the expected directorship.

All students in the sampled classes were tested. Thus, a total of 47 classes in 24 directorships were surveyed during a two-month period (mid-March through mid-May 1989) after pretesting in 3 other schools and prior to administration of the concours national. The number of units drawn at each stage of sampling was a compromise between efforts to make the data collection manageable and our attempt to acquire all the data needed for the envisaged analyses. In the analyses of student data reported below, weights have been used to compensate for the fact that the sampling was not self-weighting in the final stage.

The data collection was planned and carried out with the collaboration of the Centre de Perfectionnement et de Formation en Cours d'Emploi (CPF) and with the support of the Ministry of Primary and Secondary Education in Burundi. Research assistants spent approximately 6 to 10 days at each of the schools selected for the fieldwork.

Characteristics of the Study Populations

Information was collected from all students in the selected sixth grade classes who were present at the time of data collection, nearly two thousand (1946) in all. The majority (59%) of these grade six students were males. The sample ranged in age from eleven to nineteen years. More than half (54%) of the students were fourteen or fifteen years of age. Only 10% started school at age six and progressed to the sixth grade without repeating. A majority (55%) were repeating the sixth grade and half (50%) had repeated previous grades.

Nearly three fourths (72%) of the students' fathers had been to school; 41% completing upper primary schooling and 12% receiving some form of post primary education. A smaller proportion of mothers had been to school;

nearly half (48%) had no schooling. Less than a third (31%) completed five or more years of formal instruction.

All but one of the school directors and most (83%) of the grade six teachers were men. Given the lack of substitute teachers for staff on maternity leave, few women are assigned to teach the final year when students are preparing for the concours national. Nor were there any untrained teachers in the sample. Many (40%) of the teachers were graduates of a one-year teacher training program offered to students who had completed four years of academic secondary schooling. This program was abolished in 1982 and teacher training lengthened to two years. A large proportion (41%) had been teaching two years or less. Only a third (36%) of the teachers had five or more years of professional experience. Almost half (49%) of the school directors had three or fewer years of experience in that position.

Development of Student Tests

Two sets of test instruments were developed for administration to students, one assessing performance in the domains of text comprehension and production, and the other dealing with problem solving in mathematics, science, and agriculture. The tests were prepared in English and then translated, first into French and from French into Kirundi. This translation of test items raises fundamental questions about whether tasks can be made nominally equivalent under different language conditions. The Whorf-Sapir (1956 & 1963) linguistic relativity hypothesis asserts that tasks will be comprehended and performed differently in different languages. This has been represented as a powerful critique of the theory of nominal equivalence which is central to cross-cultural testing (Cole and Means 1981). The implication is that conventional measures of cognitive performance do not elicit universal

skills. African educational literature provides many rich illustrations of problems in constructing comparable tasks in metropolitan and indigenous languages. Consider, for example, the seemingly simple task of finding the circumference of a circle give its radius. In the languages of Eastern and Central Africa, there is no word for centre that is semantically equivalent to this term, the nearest equivalent being "navel." There are no equivalent terms for radius or circumference either (Mamari 1975). Thus, we make no assumptions about the nominal equivalence of test instruments in French and Kirundi. Nevertheless, we do assume that differences in student performance on translated tests are meaningful enough to require explanation of the sort offered below. While translation may explain some differences in student performance on the tests, differences in student exposure to the two languages offer a potentially powerful explanation as well.

Reading comprehension tests. In the case of the French comprehension tests, two versions were developed. One employed standard French vocabulary (including scientific terms), similar to the vocabulary contained in the *fichiers* prepared and distributed for classroom use by the Bureau of Rural Education in the Ministry of Primary and Secondary Education. The second resulted from a Bureau suggestion that test instruments also be developed to examine comprehension using a simplified, colloquial French vocabulary based largely on a back translation from the Kirundi version.

The comprehension test included two texts, one describing the filtration of water and the other telling how to construct a water filter from ordinary household materials. These two texts were designed to contrast processing of narrative/descriptive and procedural information. Previous research has shown that performance on comprehension tasks varies with text genre (Bracewell,

Frederiksen and Donin-Frederiksen 1982; Eisemon 1988). Generally, school children process narrative/descriptive texts more easily than the procedural texts which are more often encountered outside school (for example, instructions for using pharmaceutical products and agricultural chemicals).

The Bureau of Rural Education has promoted use of home-made water filters through health instruction in home economics classes. Water filtration is also dealt with in science classes in the context of water treatment and waste disposal. Consequently, students should have had instructional experiences from which they could acquire prior knowledge relevant to processing the texts.

The comprehension tasks elicited skills involved in locating and recalling information as well as making inferences from text propositions or prior knowledge. The tasks were constructed in the multiple choice format used for the concours national except that four possible responses were given instead of five. In total for the two texts, fifteen questions were asked. In comparison to the French and Kirundi text comprehension papers in the 1987, 1988 and 1989 concours national examinations, a higher proportion of the questions required inferences for correct responses. Moreover, the texts constructed to measure comprehension skills on the concours national are narrative texts. The secondary school entrance examination does not test comprehension of those genres in which literacy would be most useful to daily life in rural areas of Burundi (see Eisemon, Schwille and Prouty 1989).

Written composition test. The concours national examines knowledge of French and Kirundi production rules requiring students to select, from multiple-choice options, answers that are grammatically, syntactically, or semantically correct. It assesses whether students can recognize the correct

answers, in other words, that they have the ability to make inferences from a knowledge of production rules when the stimulus includes correct and incorrect applications of these rules. There is a great deal of controversy regarding how well objective tests measure knowledge of language production rules; testing organizations like the Educational Testing Service in the United States have claimed that objective test scores are so highly correlated with ratings of student compositions that there is little justification for analyzing student productions (for a critique, see Frederiksen, Bracewell, Breuleux and Renaud 1988). However, knowledge of production rules may have little to do with being able to apply them to composition tasks, particularly if the method of assessment discourages teaching of composition.

A picture task was constructed to stimulate production of a narrative composition in French or Kirundi. The pictures were presented and students asked to relate a story based on the characters and events described: theft of a woman's purse in a village market and apprehension of the culprits. Nine pictures were used for the story composition task in an effort to increase the amount of text students produced for subsequent analysis. The task, according to the Bureau of Rural Education, is similar to classroom activities prescribed in primary school for teaching composition skills in French and Kirundi. But, to reiterate, it is <u>not</u> the way knowledge of language production rules is assessed in the <u>concours national</u>.

Students were informed that their compositions would be marked for technical accuracy in the use of language, the fluency and coherence of their stories, and the overall quality of their productions. Three scores were thus derived, based on ratings by CPF researchers trained for this task. The first and second dimensions correspond to the IEA characterization of writing

competence as having two aspects, "text-producing or inscribing competence and discourse-structuring or meaning-making competence" (Purves, Gorman and Takala 1988, 43). The third captures production characteristics such as imagination and originality but, of course, incorporates the two others. As in the IEA studies of written productions, criteria were refined through ranking and scoring sample compositions in rating sessions, a process also intended to increase inter-rater agreement (de Glopper 1988). To monitor consistency, rated compositions were randomly sampled and re-rated. To lessen the possibility of rater bias influencing the results for different classes, compositions from each class were subdivided among raters. Two raters were employed for this purpose.

Mathematics test. A test of mathematics was developed measuring computational skills and problem solving. Four of the nineteen questions had to do either with computations or operations involving numerical transformations (for example, representing common fractions as decimal fractions), and required little or no text processing for successful performance. Most, however, were story problems with one or two step solutions. They test application of text processing and problem solving skills imbedded in practical tasks requiring numeracy and literacy, such as calculating correct application rates for an agricultural pesticide. Many were similar to tasks used to assess functional literacy and numeracy of school leavers in the BRIDGES pilot study of 1988.

All mathematics problems were examined by the Bureau of Rural Education (BRE) for curriculum validity, i.e., whether they measured skills that were supposed to be taught in mathematics in the upper stage of the primary cycle. The mathematics items were viewed as relatively easy by BRE staff. Moreover,

the problems cover only a small portion of the mathematics curricula and exclude many topics covered in the concours national. For instance, there were no problems having to do with calculating areas of complex geometric figures. Questions on such topics frequently appear in the mathematics paper of the concours national and probably have high discrimination indices. Nevertheless, they do not assess rules that are likely to be used in daily life.

Science-agriculture test. Another test was developed for measuring learning in science, health and agriculture--subjects which are covered in different parts of the school syllabus. Science and health are taught in etude du milieu, a course of instruction which includes geography, history, and civics as well. Agriculture is a practical subject and unlike etude du milieu, it is not a separate examination paper in the concours national. Other questions dealt with nutrition which is part of the home economics syllabus and which likewise is not examined in the concours national.

Most test items, nineteen altogether, measured students' ability to make observations or assemble scientific information in such a way that inferences could be drawn consistent with a scientific principle. Some questions elicited knowledge of physical laws and their application in problem solving: for example, a question on how much force will be needed to lever an object of known weight, given the distance to and from the fulcrum. Others presented a scientific schemata for analysis. One question concerned a food chain that can be observed in Lake Tanganyika, the task being to select a likely outcome given a hypothetical disrupting event, requiring prediction from an understanding of the schemata. In contrast, most science questions in the

1987 and 1988 concours national were constructed like those for history and geography, and measured recall of factual information imparted by instruction.

All science and agriculture items were reviewed by the Bureau of Rural Education for curriculum validity, i.e., for relevance to the syllabus in etude du milieu or in practical agriculture and home economics. About half the items required integration of knowledge from two subjects which are usually taught and assessed separately. In other words, the test largely measures students' abilities to make inferences from one domain of knowledge to another in ways that are not promoted by the organization of the school curricula, even though the policy statements from the Ministry give importance to developing skills in applying scientific knowledge to practical tasks.

Rotated test versions. Again, each test was translated into French and Kirundi. In all, therefore, five versions of the tests were produced and administered to subsamples as follows: French comprehension and composition (with standard French text) administered to approximately one-eighth of the sample; French comprehension and composition (with simplified, colloquial French text), one-eighth; Kirundi comprehension and composition, one-fourth; French mathematics, science, and agriculture, one-fourth; Kirundi mathematics, science, and agriculture, one-fourth. Within the tested classes, these five versions were distributed in a predetermined order to avoid bias in the selection of the subsamples taking each version of the tests. Students were given two hours to finish one of the five versions. Pretests established that the testing materials could be completed by most students in this time.

Development of Instruments for School Directors and Teachers

The school director and teacher instruments provide most of the information we will use to analyze the effects of school and teacher

characteristics on student achievement. The school director instrument, which was administered in French as a questionnaire, was constructed to examine possible differences between the core and the satellite school selected for study in regard to: 1) the number, qualifications, experience, and supervision of staff, size of instructional groups and incidence of repetition in the sixth year; 2) variation in implementation of double shifts, coverage of the school curricula and application of the policy of using French as the medium of instruction for most subjects in the upper stage of the primary cycle; and 3) practices involving preparation of students for the concours national such as lengthening the number of hours of instruction.

The teachers' instrument, which was also administered in French as a questionnaire, consisted of three parts. One, using an adaptation of IEA opportunity-to-learn ratings, required the teacher to review each question in the mathematics and science-agriculture tests given to students and determine if the information needed to answer the question had been taught. Another section of the teachers' instrument obtained information on how much of the prescribed curricula teachers had been able to cover, and about their examination preparation strategies. Particular attention was given to combining shifts, especially for instruction in academic subjects, deletion of nonacademic subjects from the weekly schedule, and other ways of lengthening the amount of instruction in subjects tested in the concours national. As well, teachers were asked about the extent to which they used Kirundi in lessons that are to be given in French, and about student difficulties in following instruction in French.

The third section, especially developed as an explanatory and new type of instrument for this survey, contained six types of questions measuring

teachers' understanding of the knowledge and skills elicited in assessment tasks similar to those in the student instruments. Three types of questions related to assessment tasks involving French comprehension or knowledge of production rules. Teachers were presented with sample student errors or with examination questions, and were asked to analyze them. For instance, teachers were given examples of student production errors that either resulted from misapplication of French grammatical rules or reflected first language interference errors in applying them. Teachers were requested to categorize these and propose strategies for improving student performance. This instrument presumes that, on the one hand, grammatical errors are best corrected through drill exercises which impart production rules and, on the other hand, correction of first language syntactical or grammatical interference errors requires replacement of one set of production rules by another appropriate for the second language. The latter is most likely to be achieved by practice in oral and written communication to develop fluency in the use of new production rules. In other language questions, teachers were provided with examples of examination items and asked to characterize differences between questions requiring inferences and those demanding recall of information for the correct solution.

In addition, one pair of questions dealt with mathematics and two other short sets with science. They were similar to those described above. Teachers were shown mathematics story problems that cause many students difficulty and asked to indicate sources of errors in problem representations and ways to teach students to avoid them. Some science questions involved identification of strategies for demonstrating principles such as the relationship between proper spacing and plant growth in the school garden.

Teachers' responses to these questions were rated according to a scheme for analyzing how well the responses articulate what specific questions were intended to elicit. For instance, in the example relating to student language production errors mentioned above, the best answer distinguished between kinds of errors and proposed different activities for remediating them. The resulting ratings for the six types of questions are intended to measure the insights teachers have into what is expected of students for successful performance on difficult assessment tasks, why they make mistakes, and what can be done to correct them and/or to develop the necessary knowledge and skills.

Classroom Observation and Experiment

In order to better examine the influence of language on student performance, an experiment was designed in which teachers were asked to teach, partly in Kirundi and partly in French, a lesson on calculation of simple interest. This topic is introduced in the final months of the sixth grade and is frequently examined in the *concours national*.

The national sixth grade teacher's guide (*fichier*) for the topic provides a definition of simple interest, sample calculations which teachers are to demonstrate on the blackboard and review with students, and exercises to be worked out in class as well as some to be assigned for desk work and correction by the teacher. Two characteristics of this format require comment. First, the lesson script, as for most academic subjects, is meant to be presented *in toto*; teachers are not encouraged to make selections, vary the organization of lesson material or make other innovations. Second, in contrast to the way simple interest and related computational procedures are presented in textbooks used in certain other countries, the lesson guide is

not designed to provide an explanation of procedures from which students are to infer rules for problem solutions. It is directed, instead, to imparting the rules through repeated demonstration and application. Explanatory understanding is seen to follow from rather than to facilitate competent performance of procedures and their generalization to novel situations.

Guides for this lesson were made available in both French and Kirundi to all of the teachers, the Kirundi guide being a manuscript translation from the French that had been prepared by the Bureau of Rural Education in anticipation of the eventual use of Kirundi to teach all academic subjects. The blackboard demonstrations and related exercises were divided into two parts, and each of the 47 teachers was asked to give both parts, one in Kirundi and the other in French. Half of the teachers were randomly selected to begin the lesson in French while the others began with Kirundi. Research assistants were trained to make notes on the teacher's presentation and teacher-student interaction and to expand these notes after the lesson into as complete a narrative as possible. In addition, sixteen lessons were tape recorded and transcribed. After the lesson, the teachers were interviewed about difficulties in preparing and presenting the lessons in French and Kirundi, and about their perceptions of how well students understood the subject in the two languages.

Plan of Analysis

Subsequent sections report the results of the initial analyses of the data collected in 1989. Findings relating to the effects of language on the measurement of student performance are presented first. French and Kirundi test scores will be compared, and the sources of performance differences investigated with analyses of question items. In this connection, we will

also summarize the findings of the observations of the mathematics lessons given in French and Kirundi.

Next, we examine correlates of achievement with particular attention to the effects of repetition and the comparison of core and satellite schools. The impact of teacher and instructional characteristics will then be considered. Analysis focuses on teaching practices thought to increase student performance, such as combining shifts to lengthen instruction in academic subjects and teachers' understanding of assessment tasks and sources of student errors. Finally, we will present a path model estimating the influence of important school and teacher characteristics on student test scores.

Results of Analyses

Effects of Language on Test Performance

Language of assessment, French or Kirundi, profoundly influenced the measurement of achievement in most of the subjects tested, as Table 1 indicates. The mean number of correct answers was significantly higher for the Kirundi versions of the language comprehension, composition and science-agriculture tests. The variations in performance under the different language conditions were greatest for scores on the science/agriculture test and the three ratings of the student compositions. Only in mathematics were the results for the French and Kirundi tests similar.

Insert Table 1 about here

The lowest scores were obtained for the student compositions. For instance, the mean score (M=2.27) for the overall quality of the French compositions represents an evaluation of "poor." The Kirundi compositions received higher scores, but the mean scores were also low (M=4.11 for overall quality). Ratings of the overall quality, coherence and use of language were highly and significantly inter-correlated under both language conditions with the correlation coefficients ranging from r=.87 to r=.96 ($p < .001$) for the French compositions, and between r=.78 to r=.94 ($p < .001$) for those in Kirundi. In both languages, students experienced the most difficulty in constructing a coherent narrative, suggesting that they have little practice with such tasks.

An outstanding French composition was written by a fifteen-year-old girl who was repeating the sixth grade. The composition received a near perfect score of nine for overall quality. It describes and relates the events depicted in the pictures in a clear, coherent fashion. The principal figures in the story are well characterized, especially the fruit seller. The theft of the mother's purse does not figure as prominently in the plot of this story as it did in many of the best student compositions. The student, instead, developed a more interesting story, innovating on the information she was given for the task.

At the Market

[Translated from the French original]

> One day, a mother decided to go to market to buy some fruit. Her little girl went with her. When they reached the market she decided to choose the best fruit. In the market there was fruit of all kinds: ripe bananas, avocados, pineapples and also other kinds of fruit. The fruit vendor was a fat man, with a stomach like a blown up balloon; he had a round nose and his face was also round. He wore a watch on his arm, his arms were too fat like a

> tree trunk. He was therefore very terrifying but he had a smile on his face which made him attractive. The mother carried on choosing fruit. She let herself be tempted by the biggest avocados. Next to the market there was a group of boys who were playing ball. Among them [there was a boy] who was a thief; he had taken the mother's bag and he had run. The mother and her daughter were asking help. The other boys catch him, grab the bag from his hands and then give him to the mother. As punishment they beat him so hard that he asks their forgiveness. The mother and her daughter go off to continue buying their fruit. The mother starts to negotiate the price with the fearsome vendor. The vendor did not want to agree as the mother gave him little money. Seeing that the vendor does not agree, the little girl starts to cry. As the vendor was generous, he ended up yielding as he did not want the child to cry. And the mother of the child gave the money to the merchant, then she and her daughter go in direction of the house. Once at the house, the little girl said to her mother, Well, I would like to taste those avocados. How good they are, she said! And yet the vendor gave us some at a very low price. If that is it he is just as good as these avocados!"
>
> That vendor was really very fearsome, they thought that it was an ogre.
>
> So he was ugly, but he had a good heart.

But most French compositions were very short and revealed a poor grasp of French vocabulary and production rules which in the concours national are examined with multiple choice questions. Students have little opportunity to acquire facility in written French. While students did better in Kirundi, the scores for written communication in both languages were very poor.

The comprehension test mean score for standard French (M= 6.04) was only slightly lower than that for simplified French (M= 6.53). Both were well below the score for Kirundi (M= 8.05 or about 53% of the maximum score). Interestingly, the differences in total scores for the French and Kirundi tests result more from answers concerning the narrative text on water filtration than from those concerning the procedural text. The responses of students to the questions pertaining to the procedural text on how to construct a water filter did differ significantly by language of assessment,

but the absolute differences were small. Procedural texts are less often used for instruction and are more difficult for children to comprehend. Unfortunately, comprehension of procedural texts is perhaps a better measure of functional literacy.

Scores for the comprehension and composition tests were substantially inter-correlated for the two language conditions. The standard and colloquial French comprehension tests correlated with the composition scores for overall quality, for example, at r= .59 and r= .58, respectively ($p < .001$).

The variations in the mean scores for the French and Kirundi tests of knowledge of science and agriculture were the largest for the four sets of tests. Still, the mean score for Kirundi (M= 9.31) was less than half (49%) of the highest possible score. For the French test of science and agriculture, students selected about a third (38%, M=7.16) of the correct answers. The poor results may have to do with the fact that, although the knowledge tested in this examination is supposed to be taught in <u>etude du milieu</u> or in the practical subjects of agriculture and home economics, the students' teachers reported that only half (53%) of the questions had been covered in class. Most of the questions that had not been covered dealt with agricultural topics. Nevertheless, most (68%) of the teachers claimed that the questions on this test were identical or very similar to the kinds of question items selected for the <u>concours national</u>.

The results for the mathematics tests in French and Kirundi were nearly identical. Mathematics was the only subject in which students scored slightly higher in French. Most teachers had covered the topics examined in class (89% of the questions) and most felt that the test items were similar to those appearing in the <u>concours national</u> (95% of the questions). Nevertheless, the

mean scores for the French and Kirundi versions of the mathematics test were only 46% and 43% of the highest possible score, respectively. Though the mathematics test had the highest instructional validity and the lowest differences in performance under the two language conditions, these favorable factors do not lead to much better results. The French mathematics and science-agriculture test scores were correlated as were those for the Kirundi tests (for both $r = .45$, $p < .001$).

Responses to the question items included in the reading comprehension, mathematics and science-agriculture tests were examined using monotone regression splines analysis to determine how language of assessment affects performance for students of different ability. An explanation of this new and still developing technique can be found in a paper by its originators, Ramsay and Abrahamowicz (1989), who carried out this analysis as an exploratory application. In this analysis, ability was estimated from the student's response pattern for each test, and the estimate compared to the student's response to individual questions. No independent estimates of student ability were obtained such as scores on related tests administered in the student's mother tongue. Consequently, since test difficulty affected the ability estimate, the developers of this methodology have suggested that these initial results could <u>understate</u> language effects (personal communication).

Figures 1-3 present aggregate results for the three tests. The vertical axis represents the probability of success expressed as the expected proportion of correct answers (corresponding to the probability of answering the individual item correctly in the case of the individual item graphs). Ability is estimated along the horizontal axis for the groups which were administered a particular test. The aggregate results sum the item analyses

which used a one-knot spline function to produce the item characteristic curves. A question that is highly discriminating for students of varying levels of ability should have a curve that rises steeply with higher estimated ability and probability of successfully answering the question.

The curves representing the French and Kirundi results for each item and total test scores indicate how well they measured the students' knowledge and skills. Assuming that the groups of students tested did not differ in ability--they were selected randomly from the same classes for testing in either French or Kirundi--the curves suggest how performance is affected by the language of testing. Different curves for French and Kirundi reveal underlying differences in measurement. An important feature of monotone regression splines analysis for this study is that it permits identification of which kinds of students are most affected by being tested in their mother tongue or in a second language.

In Figure 1, these curves are represented for the expected scores on the standard French, simplified or colloquial French, and Kirundi comprehension tests. At the lowest levels of estimated ability, the curves follow a similar trajectory and then begin to diverge. At the highest levels of ability, the curves for the Kirundi and colloquial French tests are close together. The curve for the standard French test is much lower and flattens out at the high end of the ability range, meaning that the results for this test give lower estimates of the proportion of correct answers among the most able students. A similar finding was obtained from comparison of the curves for the French and Kirundi science tests which were not nearly as successful in discriminating among students in either language (see Figure 2).

Insert Figures 1 and 2 about here

Because Figures 1 and 2 represent the summation of item results, they necessarily conceal wide variations in responses to individual items and tend to suppress language effects. The item analyses revealed that several items on the French science test were not well translated into Kirundi. Such difficulties could have either increased or decreased the mean differences in student performance between the two language conditions. For instance, the second question asked students to select the most desirable rotation of food crops in order to control soil borne plant diseases. The correct answer to this question required an inference from a complex set of premises regarding the vulnerability of crops to particular diseases. This was a good question in the French version of the test. In the Kirundi test, student responses peaked at a low level of estimated ability and levelled off. The most able students were not more likely to get the correct answer in Kirundi probably because they did not understand the question. On the other hand, for many questions, the Kirundi curves were substantially higher than those for French. This was particularly true of the questions in the comprehension tests requiring high level inferences. An example is a question requiring students to select an appropriate title for the procedural text (see item 15 in Figure 4). For this question, the curve for standard French peaked at the 30% level of probability of successfully answering the question and then flattened out. The curve for colloquial French reached 45%. The Kirundi curve was near perfect, reaching 100% probability of success for students at the highest level of ability.

The principal finding derived from the analyses presented in Figures 1 and 2 is that comprehension skills and knowledge of science and agriculture are more poorly measured in French than in Kirundi. Of particular importance is the fact that it was the performance of the most able students that was most affected by being tested in French. Conversely, testing in the mother tongue did not increase the performance of students with less ability who had great difficulty with the tests in either language. And to reiterate, an independent Kirundi measure of student ability would probably magnify the language effects for these tests. Given that the etude du milieu section of the concours national is given entirely in French and is used to identify students at the very highest level of ability, this finding deserves special attention.

Insert Figures 3 & 4 about here

Figure 3 presents a very different finding. There were no language effects for mathematics. The French and Kirundi curves follow the same paths, indicating that the two tests performed similarly in discriminating among students. Since most of the test questions were story problems, the similarity in the French and Kirundi results cannot be explained simply in terms of the fact that the items required little language comprehension for correct solution. Linguistically demanding problems such as one requiring students to solve an ordering problem using a set of "bigger/smaller than" premises, sometimes did produce differences in favor of Kirundi. But many problems of comparable difficulty requiring students to process a great deal of textual information for accurate problem representation (such as

calculation of area problems necessitating multistep solutions) did not. The domain of mathematics, it seems, has little in common with the other subjects tested in one important respect: the underlying knowledge and skills normally acquired and tested in French in the upper stage of the primary cycle are not better measured in Kirundi even when the cognitive tasks involved necessitate complex representations of information for problem solving.

Certain kinds of computation and mathematical problem solving may involve little cognitive interference when carried out in French because equivalent symbols, structures or concepts now exist in Kirundi, having been imported into that language from French. The base-ten number system is an illustration. It has no traditional equivalent in Kirundi (Zaslavsky 1973, 240-41) or in many indigenous African languages which use a base-five system (Gay and Cole 1967; Reed and Lave 1979). Consequently, complex computations are often performed in Kirundi with a number system and operations for manipulating numbers derived from French and taught in Kirundi in the lower stage of primary school. This may facilitate mathematical thinking in French and in part explain differences between performance in mathematics and the other domains.

An Experiment for Teaching Mathematics in French and Kirundi

Teachers were asked to teach a grade six lesson on the calculation of simple interest in two parts, one in French and the other in Kirundi. More than three fourths (77%) of the teachers said that the lesson was easier to teach in French than in Kirundi because they and the students were more familiar with the necessary vocabulary. Kirundi words have been invented for French terms like rate of interest inyungu, which means interest and rate as

well as rate of interest. But these terms are often ambiguous and seldom used in ordinary speech.

The number of student/teacher interactions and the length of each class were observed in order to determine whether the language of instruction influenced the amount of student participation. However, language of instruction had no effect on the mean number of student/teacher interactions (M= 1.23/min and M= 1.18/min., F= 1.29, NS). Only about 20% of the students participated under either language condition.

Sixteen lessons were transcribed and the number and kinds of language switching recorded. An equal number were given in French and Kirundi. Forty-six examples of language switching were noted, eighteen when the class was given in French and the rest when Kirundi was used. Almost all (89%) were initiated by teachers. When teachers switched to Kirundi during a class given in French, the purpose was usually to introduce a new topic or repeat instructions (14 of 18 instances), as in these examples:

> Teacher (begins reading from the French lesson plan): ...The first problem: Someone puts 12,000 FBU into a savings bank in Burundi at a rate of 3%. Calculate the interest at the end of nine months. The second problem: A merchant puts 500,000 FBU at the rate of 8% during six months. Calculate the interest received. Nyarutsa [work quickly on] the two exercises. Read them carefully. Distinguish the questions from the answers. Si ibintu bigoye, abatazi irya facon yo kugira, abatazi irya facon yo guca baronka operations buno nyene barashobora gukoresha [They are not difficult, those who do not know the way, those who do not know the way of getting straight away to the problem can use] the example...[It is] easier you understand muravyumva. Hein abatari ...[Those who are not] capable yo gukoresha yo kubona [of solving] the problem directly, you can use the example...Mugabo ushobora guca abona [But if you can immediately understand] the problem buno nyene ushobora kuyikoresha [you can solve it immediately]...
>
> Teacher : This, that is not the correct answer! You have to redo it subiramwo urabe [redo the answer].

Kirundi was seldom used in the French classes to simplify, clarify or provide explanations. An exception drawn from the same lesson is given below:

> Teacher: Ah, you mixed up the rates with other problems! Urabona, benshi mukunda kwihenda mufata ivyo mwamye mukoresha mukaba ar'ivyo mukoresha. [You see, many make mistakes by taking things they have been using elsewhere and using them again.] Here we have been given in the problem the rate of 8%, but Michel had used the rate of 3%. Iyo ukoresha agapfindo, uzoba uri [If you had used the correct rate, it would be] perfect. Naho ukoresha amaformules yo ntushobora gutora ntuze [Even if you used the right formula, you cannot find the right] answer because you made a mistake in the rate.

In the parts of the lesson taught in Kirundi, teachers often (14 times) switched to French to prompt students to answer questions they apparently had difficulty following in their mother tongue. French was also used for simplification and providing feedback to students. Student-initiated language switches usually involved using French words when a student was unable to find the appropriate Kirundi term.

To summarize, changing the medium of instruction in mathematics from French to Kirundi presented many difficulties to students and teachers, and it did not increase student/teacher interaction or encourage expository teaching except in the sense that it obliged teachers to make unfamiliar Kirundi terms better understood by switching to French. If the findings of this experiment and the test results reported above are indicative, important changes in teaching and cognitive outcomes are not readily obtained by offering mathematics in Kirundi. In other subjects, the advantages could be greater.

Relationship between BRIDGES Test Scores, Secondary School Entrance Examination and Background Characteristics

The BRIDGES tests and the concours national measure somewhat different learning outcomes. The scope of the tests is different as is the language of assessment and marking. The BRIDGES tests were given in both French and Kirundi and for most, the scores follow a normal distribution. The high "floor" for success in the concours national is determined by the number of

places in secondary schools; as noted previously, about 10% of the students pass the examination.

Recent results for the concours national examination were obtained for each of the schools in the BRIDGES sample. With these schools as the unit of analysis, concours national success rates for two years (1987 and 1988) were inter-correlated, but not at a particularly high level ($r = .32$, $p < .05$). The various BRIDGES test scores were also correlated with the scores students in these schools received on the 1989 concours national examination. For instance, the school mean scores in BRIDGES for comprehension of standard French and Kirundi were correlated with mean school level results for the French and Kirundi papers from the concours national ($r = .37$, $p < .01$ and $r = .32$, $p < .01$, respectively). The correlation between results on the concours national mathematics paper and the scores for the BRIDGES French test of mathematics was somewhat higher ($r = .48$, $p < .001$).

Several student biographic characteristics were weakly associated with performance on the various BRIDGES tests. Father's schooling, for instance, was slightly correlated with seven of the test scores, six of them scores for French tests, statistically significant correlations including comprehension of standard French ($r = .14$, $p < .01$) and the three scores for the French composition task ($r = .13$, $p < .01$, $r = .14$, $p < .001$, and $r = .12$, $p < .01$). Males scored slightly but significantly higher than females on every test except, interestingly, the French and Kirundi tests of knowledge of science and agriculture and the test of Kirundi composition skills. The largest sex differences were for the French and Kirundi tests of mathematics ($M = 9.06$ for boys in French, $M = 8.07$ for girls, $F = 5.444$, $p < .05$; $M = 8.68$ for boys in Kirundi, $M = 7.50$ for girls, $F = 8.850$, $p < .05$, respectively).

Effects of Grade Repetition on Student Performance

Of all the variables examined, whether a student repeated grade six had the most impact on the BRIDGES test scores. The differences between the mean scores of repeaters and nonrepeaters were both large and highly significant as is evident from an inspection of Table 2.

Insert Table 2 about here

Repeaters scored higher on all tests, especially in mathematics. In general, the differences were greatest for the French tests. In the French mathematics test, for instance, the difference between repeaters and nonrepeaters in the proportion of correct answers was 55% compared to 32%. Nevertheless, the language effect appears in large part to compensate for the repetition effect. In other words, on certain tests the nonrepeaters did as well or almost as well on the Kirundi version as the repeaters did on the French version. This can be seen in the comprehension test (nonrepeaters 6.79 in Kirundi vs repeaters 6.94 in standard French), the general quality of composition (nonrepeaters 3.24 in Kirundi vs. repeaters 3.33 in French) and especially in the science/agriculture test (nonrepeaters 8.79 in Kirundi vs. repeaters 7.82 in French). In addition, whereas the repeaters did better on the French version of the mathematics test than on the Kirundi version (10.46 vs. 9.69), the nonrepeaters did slightly better in Kirundi than in French (6.40 vs 6.06).

Effect of Type of School on Student Performance

and Concours National Success Rates

Whether the school a student attended was a core or a satellite school was hypothesized to affect learning outcomes. However, for only one of the BRIDGES tests were the scores significantly different. In the mathematics test given in Kirundi students in core schools with no satellites had the highest scores ($F = 5.4659$, $p < .01$).

The relatively little variation in test scores across school types is all the more surprising in light of the fact that, within the BRIDGES sample, student reported repetition rates in grade six, which strongly affect student performance, are higher in core schools without satellites (see Table 3). Almost two-thirds (62%) of the grade six students in core schools without satellites were repeaters compared to less than half (48%) of the students in satellite schools. Controlling for repetition eliminates the difference for type of school on the Kirundi mathematics test.

Insert Table 3 about here

Levels of performance of students on the various tests were examined to determine whether use of mean scores to make these comparisons obscured performance differences at other points in the distributions. The results (not shown) indicated that the type of school attended was not related to extreme levels of performance, i.e. student placement in the upper or lower quartile of test scores. Students in schools in the same directorship were compared to determine whether the variability among directorships in, for example, the distance between core and satellite schools, was responsible for

these results. The results (also not shown) showed no consistent relationship between distance and the size of the difference between core and satellite schools.

These findings are consistent with analyses of 1989 concours national student scores for the schools studied. Comparison of mean scores for the different types of schools did not reveal significant differences. For instance, the mean total score of students at satellite schools (M= 58.68) was only slightly below that for core schools (M= 61.33 for those with satellite schools attached to them; M=61.15 for core schools with no satellites, F= .3497, NS). Unfortunately, we were unable to compare pass rates because the minimum passing scores needed for admission to secondary and postprimary technical schools had not been announced at the time these analyses were carried out.

In contrast, the pass rates of students in core schools on the previous year's concours national examination of 1988 were higher than the rates at satellite schools. In fact, students at the core schools with satellites were twice as likely to obtain a pass. The success rate at the satellite schools was 5%; at core schools with no satellites, 8%; and at core schools with satellites, 10%.

Surprisingly, grade six student repetition rates estimated at the school level had a slight negative correlation with concours national pass rates (r=-.11), which is the opposite of what we found for the relationship between repetition and the broader measure of learning outcomes provided by the BRIDGES student test scores. Also, in contrast to Table 3, which is based in self-report by students, the 1988 grade six repetition rates officially reported by school directors for core and satellite schools were nearly

identical; 42% for satellite schools and 41% for core schools with no satellites. These reported repetition rates ranged from 25%-53% for core schools and from 21% to 69% for satellite schools.

Many factors could contribute to lower success rates of students in satellite schools on the concours national. Satellite schools are smaller schools with fewer grade six students and teaching staff. Their average total enrollment, for instance, was 478 compared to 659 students in core schools with satellites and 892 students in core schools with no satellites. Facilities at satellite schools are somewhat poorer. According to school director reports, more than a third (36%) of the satellite schools are not constructed of durable materials compared to a fourth (25%) of the core schools. Most (62%) of the core schools have good or acceptable sanitary facilities for students and staff while only one-third (33%) of the satellite schools do. Almost all (89%) of the satellite schools had no running water, compared to half (48%) of the core schools. Satellite schools also have less access to substitute teachers. None of the satellite schools sampled had substitutes available to replace absent teachers compared to about a quarter (26%) of the core schools.

Moreover, the satellite schools had fewer experienced teachers, especially those with experience in grade six. For example, most of the teachers interviewed at satellite schools had about three years of experience in grade six compared to more than seven years for teachers in core schools with satellites. As well, a higher proportion had obtained older teaching qualifications. Teaching experience was correlated with 1988 concours national pass rates ($r = .27$, $p < .05$), but level of teacher qualification was not ($r = -.18$, NS) since certain levels of teacher qualification represented

completion of programs that have now been phased out. Teachers interviewed in satellite schools placed less importance on preparing students for the concours national examination. Only one third (33%) ranked it most important, compared to almost half (46%) of the teachers at core schools. Nevertheless, there were few differences between core and satellite schools in examination preparation practices. Satellite schools were as likely as core schools to combine classes in academic subjects, skip instruction in practical subjects and lengthen the school day prior to the administration of the concours national.

Thus, it remains something of a mystery why there is not more difference in test scores between core and satellite schools. This issue deserves more analysis.

School and Instructional Characteristics Related to Student Performance

Measures of factors affecting school effectiveness in preparing students for the tests developed for this study were grouped according to whether they related to: (1) student social and educational background; (2) school management and teacher supervision; (3) implementation of the national curricula and instructional practices that influence opportunity to learn; and (4) teacher experience and skills.

Student social and educational background Two important factors relating to student social and educational background have already been mentioned: father's schooling and student repetition of grade six. While primary schools in Burundi are not supposed to be selective, school directors and teachers do make decisions that affect student repetition. In grades 1-5 school directors are authorized to allow repetition for students whose end-of-year grades do not exceed a certain minimum (40%), whereas at the end of sixth

grade, repetition is authorized for better students who are thought deserving of additional tries at the secondary school entrance examination. Normally, permission to repeat sixth grade is contingent on the student passing the examination given in the final semester (which is different from the concours national). Among the schools sampled, school-level pass rates on the 1988 school leaving examination varied from 18% to 94%, with a mean of 56%.

Teachers and school directors may adopt a permissive policy toward student repetition if they feel that it increases the proportion of students who succeed in the concours national examination. Many (61%) of the school directors reported that concours national success rates strongly influence how their work is evaluated by the Ministry of Primary and Secondary Education. Fewer (39%) felt that controlling student repetition is very important. Repetition rates are, however, monitored by the Ministry. That may explain the seemingly low repetition rates reported by some school directors.

The fact that, in Burundi, sixth grade repeaters did better than nonrepeaters runs against the grain of much that has been written (see Schwille, Eisemon and Prouty, 1990, for references) and hence it is tempting to discount such findings on methodological grounds. For example, one objection is to say that the comparisons are inappropriate because repeaters and nonrepeaters are drawn from such different populations. But in our view, the reason for making these comparisons is precisely to raise such questions as: How do these repeaters differ from nonrepeaters at the same grade level and from those who get promoted to the next grade level? What is the variation in the nature and effects of repetition across schools and countries?

Our Burundi data do not answer all these questions, but they do suggest that the social meanings and consequences of grade repetition may differ greatly across settings. Many Burundi repeaters at the sixth grade level are well motivated and scholastically capable children who are taking extra time to prepare for additional attempts to pass the very selective secondary school entrance examination. The contrast is great with a view of repetition in which repeaters are by and large students who are doing very poorly at school and who are held back because they are judged not to have learned enough to meet the minimum standards for the grade in question.

In considering the selectivity of repeaters versus nonrepeaters, we examined several background factors to the extent possible with data that are limited in these respects. The proportion of repeaters who were male was nearly the same as the overall proportion of males in grade six. The fathers of repeaters had, on the average, a little more schooling than the nonrepeaters ($M=4.26$ years vs. $M=3.82$ years, $F=8.779$, $p < .05$). There was no such difference in mothers' education. A more important difference derives from the fact that nonrepeaters were the first class to go through the double shift system which was adopted progressively one grade at a time since 1982. Due to this fact repeaters had more advantaged school conditions in their earlier years.

School management and teacher supervision. This group of variables includes the many activities associated with monitoring instruction to ensure compliance with ministry policies. One of the main responsibilities of school directors is to visit teachers, observe their classes, and provide guidance to them. School directors, we have noted, are required to make one hundred and fifty classroom visits a year. School directors submit written reports on

these visits to school inspectors. Information was obtained, first, from the school directors as to the number of times they had visited the teachers in the core and satellite schools in the sample and, secondly, from cantonal school inspectors on the number of classroom visit reports received from school directors for the first and second trimesters of the 1988-89 school year. It is evident from the records of school inspectors (who themselves do little classroom supervision) that few school directors are able to visit one hundred and fifty classes per year, fifty in each trimester. In one of the cantons sampled, for example, of the six school directors, none had completed one hundred school visits by the end of the second trimester. There was a great deal of variation among school directors in the number of class visits they reported. One director had made ninety visits and another fifty-one, or half of the number of visits required. Distance between schools and lack of transportation sometimes make it difficult for directors to visit satellite schools as often as core schools. Three of the school directors in this canton made substantially fewer visits to satellite schools than to classes in core schools. In one case, only sixteen visits to satellite schools were reported, compared to forty-six to the core schools.

When school directors visit classes, they are supposed to evaluate many aspects of a teachers' performance: whether the lesson is being taught at the time specified in the calendar given for each academic subject, how well the teacher follows the fichiers, and so on. In addition, the school director observes the language of instruction and provides advice to teachers in the upper stage of the primary cycle on the use of French and Kirundi. Ministry policies call for all subjects except Kirundi, even agriculture and home economics, to be taught in French in the fifth and sixth year. Unfortunately,

many students cannot follow instruction in French. Teachers estimated that less than one third (30%) of their students understood French well enough to be taught in that language without frequent recourse to Kirundi. School directors were asked what advice they gave to the teachers on the use of French and Kirundi in the fifth grade when the transition to French occurs. Such information constitutes a useful measure of the extent to which they are encouraged to use French exclusively. This variable was correlated with the proportion of students in the sixth grade classes sampled who were perceived to be able to follow instruction in French ($r = .29, p < .05$).

National curricula and instructional practices. Practices that may be effective in increasing learning outcomes include some that involve adherence to ministry policies as well as others that require deviations from them. Successful implementation of the school curricula is predicated on fidelity to the *fichiers* provided for academic subjects and to the schedule for teaching lessons. Teachers cannot easily accelerate coverage of the curricula through, for instance, selecting some topics for independent study. In part, this is because there are no student textbooks for most academic subjects. The *fichiers* supplied to teachers are the only published instructional materials for subjects like mathematics and *etude du milieu*. Moreover, double shifts have reduced instructional time, necessitating adherence to the teaching schedule to cover what may be examined in the *concours national*. Only two to three weeks is provided for review for the examination. This period cannot be increased without skipping topics. Many topics of examination questions are introduced in the lesson schedule in the period immediately preceding testing, calculation of simple and compound interest, for example. The highly variable syllabus coverage of the *concours national*, together with the factual content

of many examination questions, discourages teachers from being very selective, especially in teaching academic subjects.

However, as noted above, teachers can increase the amount of instruction students receive on important topics in other ways. A majority (55%) of the teachers admitted to often combining morning and afternoon shifts during the school year, and almost all (88%) had done so in the past two weeks for the teaching of French. This deviates from the double shift policy which schools are expected to follow. But the practice of combining shifts is related to school success on the concours national examination. Schools which combined classes had higher success rates in 1988 ($M = 9\%$ vs. $M = 5\%$, $F = 6.662$, $p < .05$).

Teachers also combine different grade levels for instruction in practical subjects. Most (60%) did this for teaching practical agriculture. A fourth (25%) of the teachers said that they skipped teaching practical agriculture often or occasionally, and about the same proportion (27%) reported skipping home economics. In the two weeks prior to the time the teachers were interviewed, most teachers had skipped one or more lessons in agriculture (61%) and home economics (57%), usually using the period to teach French or mathematics instead. Another strategy to increase instruction is to lengthen the school day.

Teacher skills and experience. Insightful measures of teacher skills were obtained from administration of an instrument measuring teachers' understanding of assessment tasks, sources of student errors and strategies for remediating them. These were described earlier. Their purpose was to facilitate identification of teachers who are most effective in preparing students to answer the kind of questions included in the instruments measuring student achievement.

The teachers showed a great deal of variability in the scores received for this instrument, in terms of total scores and between subscores as well. Primary school teachers in Burundi and other countries are not subject specialists; they must teach all areas of the school curricula. Some teachers scored highly on all measures, and some scored very poorly overall. Most received variable scores.

In educational systems like Burundi's that are highly selective and in which teaching in the upper stage of the primary cycle is oriented to preparing students for the national secondary school entrance examination, one way to measure effective teaching is in terms of student results and the skills implicated in examination success. Teachers' previous experience in grade six and their success in preparing students for the concours national may be considered indicators of teacher skill. A third indicator we used is the number of times the teacher has given demonstration lessons to colleagues during in-service training.

In the instrument concerned with teacher understanding of assessment tasks, teachers were given the following examples of errors students often make in spoken and written French:

1. Je ne vois lien (the word rien is often mispronounced by Kirundi speakers, becoming lien and transforming its meaning: "I don't see link" for "I don't see anything").

2. Je vais venir hier ("I will come yesterday").

Asked about the source of these errors, one of the teachers said: "It's phonetics and [semantics] that are the problems, as well as translation from the mother tongue to a foreign language which creates interference. For example, ejo in Kirundi means at the same time yesterday and tomorrow in

French. This distinction between the past and the future is not the same as in French." In order to assist students to correct these kinds of errors, the teacher described several oral exercises to improve pronunciation and increase students' awareness of tense structures in the second language.

This teacher also exhibited much insight into the difficulties that many students have in mathematics. The teacher was shown two story problems: (1) Twice as many boys as girls participated in a marathon; if 200 girls participated, how many boys participated? (2) Twice as many boys as girls participated in a marathon; if 400 boys took part, how many girls took part? The teacher responded that multiplication problems are easier for most students than division problems because they involve application of algorithms. He knew that the correct solution depended on an accurate problem representation: "to help students better resolve such problems, one has to first explain to them that to resolve (them), one has to first have understood it, that is, which operation to use, to know when to do it..." He received a score of fifty-six of a possible sixty points for all of the items in this instrument.

Ratings for the language arts and mathematics-science items in the teacher skills instrument were correlated with years of teaching experience in grade six ($r = .33$, $p < .01$; and $r = .36$, $p < .05$, respectively). They were also correlated with demonstration teaching ($r = .26$, $p < .05$; and $r = .29$, $p < .05$). Unexpectedly, level of teacher training was negatively related to the two measures of teacher skills ($r = -.25$, $p < .05$; $r = -.23$, $p < .05$). This anomaly may be explained by the fact that level of teacher training is negatively confounded with teaching experience in grade six. Graduates of the older

teacher education program which provided only one year of training after four years of academic secondary education have more experience at this grade level.

Correlates with Student Performance

Correlations between the mean student achievement scores in language arts and mathematics-agriculture for each school and various teacher and instructional characteristics are presented in Table 4. The number of class visits in the past two weeks to sixth grade classes as reported by the school director had the highest relationship with test scores. The correlation was moderately strong and significant for all of the tests. The number of times the teachers studied were supervised ranged from two to twelve. Two thirds (65%) of the teachers received four or fewer visits from the school director. Reported teacher punctuality was correlated with three of the mean test scores. Application of the language policy was correlated with the two language arts tests. Most measures of opportunity to learn such as the number of combined shift classes given in the past two weeks were, with few exceptions, not significantly related to student class mean scores. Among the teacher characteristics and measures of teaching skills, teacher experience in grade six and demonstration teaching were related to the Kirundi test of mathematics and science; previous success in preparing students for the concours national, to the Kirundi language arts and French mathematics-science tests; and teacher assessment skills in mathematics and science, to the associated Kirundi test.

Insert Table 4 about here

A Model of School Effectiveness

The relationships among the supervisory and instructional characteristics mentioned above, and their contribution to variations in student performance for the classes sampled were investigated using a model for examining linear structural relations among variables (as estimated by LISREL, see Joreskog 1978; Joreskog and Sorbom 1986). This method allows investigation of the relationship of variables to underlying theoretical constructs, as well as among variables whose causal relationships are specified *a priori*. The latter feature is of more importance for this analysis which tests a structural model of how school and instructional characteristics might influence student achievement.

A path analysis model was developed to describe relationships among variables likely to affect student performance in language arts and mathematics and science assessed in French. For exploratory purposes, student level variables such as fathers' schooling and grade six repetition were excluded and class mean scores were used for the data analyses. That is, information about supervisory and instructional characteristics pertains to classes rather than to individual students. This has several implications. Using mean scores for the achievement tests obscures variations in performance within classes. The effects of student background characteristics are ignored. In addition, the number of observations is reduced very substantially, potentially inflating correlations and controverting some assumptions of the LISREL method which normally require large sample sizes to satisfy. Our intentions, however, are modest. We wish to explore whether a general model of the impact of supervisory and instructional characteristics

on student achievement at the school level is consistent with the data for the French tests in language arts, mathematics and science.

The general model contained four exogenous variables: the number of class visits to sixth grade classes in the past two weeks as reported by the school director (NUMVIS), the advice given by the school director on the extent to which French should be used exclusively (LGADVD), the number of extra teaching hours added to the class schedule in the past two weeks (NUMHRS), and the number of years the teacher has taught in grade six (YRSIN6). Endogenous variables included reported teacher punctuality (PUNCT), the number of times shifts have been combined for instruction in the past two weeks (MIXTOT), the number of lessons skipped in the less important subjects over the past two weeks (TOTSKP), the average predicted number of days required to finish the school syllabus in various subjects (TOTEND), the number of demonstration lessons given by the teacher in the past two years (DEMTCH), the teacher's success in preparing children for the <u>concours national</u> examination the previous year (EST7SC), the teacher's assessment skill ratings in language arts (SET123), and the French test scores for language arts (LARTSF), the latter being the sum of the comprehension and composition scores. This model is shown in Figure 5.

Insert figures 5 & 6 about here

Arrows indicate the direction of relationships. For example, teacher supervision, i.e. the number of visits made by the school director (NUMVIS), was considered to "cause" teacher punctuality (PUNCT) which, in turn, was viewed as "causing" higher student achievement (LARTSF). Similarly, number of

director visits (NUMVIS) was expected to have a negative relationship to skipping lessons (TOTSKP), and through it, to achievement (LARTSF).

In the general model, several instructional practices were regarded as being affected by teacher supervision in different ways. Closer teacher supervision might be expected to foster greater adherence to ministry policies, resulting in less lesson skipping, fewer instances of combining shifts and fidelity to subject-matter syllabi which call for completion of subjects not earlier than two to three weeks before the administration of the concours national examination. School directors' advice to use French exclusively was posited to directly influence student achievement in French.

Teacher characteristics and skills constitute a submodel within the general model. Years of teaching experience in grade six, an exogenous variable, was represented as influencing teacher skills, demonstration teaching, and previous success in examination preparation. It is shown in Figure 5 as indirectly influencing student achievement. Teacher characteristics and skills were not connected to particular instructional practices in the model as they were regarded as being conceptually independent of them. Nor were most linked to teacher supervision for the same reason. Closer supervision, for instance, was not predicted to directly affect demonstration teaching though it was thought to influence teacher success in preparing students for the concours national. The number of extra hours of instruction was not seen as a policy variable affected by teacher supervision, an outcome of teaching experience or necessarily an indicator of teacher skill, but instead was thought to be related directly to student performance.

Results of the Path Model

The LISREL method requires testing of the assumptions of a model through chi-square tests or other indices of fit. Using the measure of student achievement in French language arts, the full model was tested as depicted in Figure 5. It yielded a chi-square of 38.32, with 40 degrees of freedom ($p = .546$), a moderate fit to the data indicating that the model cannot be rejected. The model was also estimated after removing the direct path linking number of extra hours of instruction (NUMHRS) to the outcome measure. This second model produced a chi-square of 35.57 with 33 degrees of freedom ($p = .348$).

Figure 6 presents the results of the latter model for the French language arts scores. Among exogenous variables, the number of director visits has the highest coefficients, particularly for the paths to teacher punctuality and previous teacher success in examination preparation, and the one leading directly to scores on the French language arts test. The connection between teacher supervision, teacher punctuality and achievement was the strongest indirect relationship. Instructional practices that were predicted to be related to teacher supervision, such as combining shifts and skipping lessons, had relatively weak relationships to supervision. However, the signs of these relationships are of interest. Combining classes and skipping lessons were positively influenced by director supervision. Moreover, these practices were positively related to student performance. It seems that schools which are supervised more closely and, presumably, held most accountable for the performance of their students, are those which in certain respects exhibit less not greater adherence to ministry policies. While the path coefficient for the link between director visits and combining classes is weak, the

coefficient connecting combining shifts to student scores is more substantial. Combining shifts did seem to increase achievement.

The number of director visits is strongly related to the measure of previous teacher success in examination preparation. But we cannot conclude that closer teacher supervision necessarily causes successful teaching as the coefficient for the path leading from concours national success to the outcome measure is relatively low. Moreover, our measure of supervision was collected after not prior to the measure of teacher examination success. Nevertheless, the direct effect of teacher supervision on teaching success is noteworthy.

Among the variables in the submodel, the relationships between years of teaching experience and demonstration teaching, and especially the measure of teaching skill in language arts, are for the most part stronger than the connections between these variables and achievement. The endogenous variables seem to "absorb" the effects of teaching experience and do not themselves consistently and strongly influence student achievement. Demonstration teaching is the one variable linked to teaching experience that has considerable impact on achievement. The effect of the teacher skill measure is disappointingly small.

In brief, the most powerful features of this model of school effectiveness have to do with school director visits, its direct impact as well as its indirect impact on learning outcomes through teacher punctuality. Similar results (not shown) were obtained from testing the revised model with the French mathematics-science scores.

Discussion

Five conclusions are supported by the various data analyses. First, students exhibited low levels of achievement in all of the subject domains

tested. Testing in the mother tongue improves the measurement of learning outcomes. But it mainly enables a more precise estimate of levels of performance for better students. Although few students performed well under either language condition, the mean level of performance by repeaters in Kirundi was high enough to give one the sense of how much the Burundi system would accomplish under better conditions.

The low levels of performance in the language arts tests should be of particular concern to policymakers. Most students, for instance, had difficulty comprehending the kinds of procedural texts in French and Kirundi which they are likely to encounter in daily life. The ability to use information from such texts is central to functional literacy. Students' skills in written communication were even poorer, particularly in French. That, of course, has serious implications for secondary schooling.

The performance of students was relatively strong in the test of knowledge of science and agriculture when they were examined in Kirundi. This might be seen as consistent with the objectives of the 1973 reform which stressed practical studies and mother tongue teaching. Unfortunately for this explanation, practical studies currently have an inconsequential place in the primary school curricula, largely separated from science and other cognate academic subjects, and they are to be taught in French and not in Kirundi. Moreover, the content of the science and agriculture test was generally not covered by the students' teachers. In other words, the test results provide little evidence of the effectiveness of present practices; in fact, the tests may be considered to measure learning outcomes that are important but not likely to be learned in school.

Second, the policy of switching to French as the medium of instruction in grade five profoundly influences how much students learn. Even in grade six, according to the teachers interviewed, only a small proportion of students can follow instruction in French. This depresses test performance in both languages since French is used as the medium of instruction for most subjects taught in grades five and six. The exception is mathematics. Although the mean scores in mathematics were a little below fifty percent of possible marks, some students obtained perfect scores. Student performance in mathematics was least affected by the use of French as the language of assessment, and teaching mathematics in Kirundi, our experiment indicates, poses many problems for students as well as for teachers. The implication is that language policy should be more sensitive to the subject domains which are to be taught in French and the mother tongue.

Third, some instructional practices that increase opportunity to learn like providing extra hours of instruction have a relatively weak impact on learning outcomes both directly, and indirectly as estimated in the path models. That may be because the BRIDGES tests measure cognitive skills that are less sensitive to practices which many teachers feel are useful for preparing students for the concours national. Examination "cramming," though it may increase the likelihood of concours national success, may not be particularly effective in developing high level cognitive skills.

Fourth, school management has a strong direct and indirect influence on instruction and learning. Teacher supervision is an important means through which the amount and "quality" of instruction is influenced. School directors also do substitute teaching, organize inservice training, and have other responsibilities such as liaison with school committees and ministry and party

officials. As a result, many may have too little time and too few resources to carry out their primary responsibility which is to supervise instruction in several schools.

Fifth, repetition has more effect on student performance than any other social background, school or instructional characteristic measured. A threshold effect cannot be established from this study since we did not collect information on the number of times a student repeated. Nor, more pointedly, do we know how repetition affects success in the concours national. Nevertheless, students who repeat grade six do benefit from the additional opportunity to learn. High repetition rates are an indicator of limited opportunity to learn what is expected of students for entry into secondary school. It must be reiterated that repetition is largely self-selective; students voluntarily repeat in order to obtain a better examination result. Those who repeat are likely to be the most able students.

Conclusion: Strengthening Basic Education

Strategies for change

The findings summarized above, together with those obtained from the pilot studies of rural schools, suggest many strategies for raising levels of achievement, improving monitoring of learning and reinforcing school management as well as for increasing the internal efficiency of primary schooling.

Increasing opportunity to learn. The primary school curriculum in Burundi and especially the standards of performance expected for success in the concours national are probably well beyond what is expected of grade six students in many other African countries, and in many Western countries as well. The scope, content and level of item difficulty of the concours

national and the tests developed for this study, for instance, appear comparable to the Kenya National Certificate of Primary Education Examination administered at the end of Standard 8 to students who have received not only at least eight years of instruction, but also two hours of instruction per day more than their Burundian counterparts (Eisemon and Schwille in press).

The double shift policy has reduced the school day to little more than three hours in order to expand school participation. There is strong anecdotal evidence from teachers and school directors that the quality of education has been adversely affected. Many teachers combine shifts and classes to increase the amount of teaching and curriculum coverage of academic subjects during the school year, and especially prior to the administration of the concours national. Combining classes at different grade levels for practical instruction (e.g., agriculture or home economics), for example, allows teachers to spend more time on lesson preparation and marking while combining shifts increases the number of periods of instruction students receive in mathematics and other subjects examined in the concours national.

Combining classes in different grade levels for practical instruction is not likely to adversely affect learning or seriously strain school facilities and resources in light of the way such instruction is presently conceived and organized. Much of this practical teaching takes place outside school buildings in school gardens and involves no formal instruction or use of equipment apart from the tools students bring with them to school.

However, combining shifts for instruction in academic subjects has more serious implications for learning, though not the same implications for all subjects. In the upper stage of the primary cycle, student textbooks are available only for language arts subjects, French and Kirundi. Doubling class

sizes for these subjects makes it difficult to teach reading and written communications skills which are best developed through study and exposition of printed texts. On the other hand, teachers guides (fichiers) are the only instructional materials for mathematics and etude du milieu. This lack of student materials may not be desirable for teaching independent study skills, but it affords opportunities for increasing class sizes without a proportionate reduction in the amount of learning material that a teacher can cover.

Changing the curricula and language of instruction. Teachers reduce the scope of instruction by neglecting the teaching of practical subjects. This problem is not likely to be remedied through closer supervision and inspection since teachers are responding to demands to better prepare students for the concours national by, for instance, skipping lessons in agriculture. There is not enough time to teach these subjects. The introduction of double shifts was not accompanied by sufficient modification of the school syllabus or expansion of the concours national to ensure that compulsory subjects are taught.

One way to address the problem would be to integrate instruction in certain scientifically-oriented subjects and to cover all of them in the concours national. This includes biology, health, nutrition and agriculture that are now, for the most part, taught as separate subjects. They could be profitably combined into a single subject, integrated science. This is also an implication of the pilot studies which showed how little knowledge of modern science was used by school leavers in practical problem solving tasks or in understanding procedures associated with modern production technologies.

The separation of these subjects in the school curricula does not foster integration of knowledge that is necessary to increase productive capacities.

Designing an integrated science program for grades five and six will require rethinking of the purposes of instruction in agriculture, health and home economics. Instruction in practical subjects lacks scientific content. For example, work in the school garden usually involves application of cultivation skills that students in rural areas learn from home. Most agricultural instruction is not intended to demonstrate what is taught about biology in etude du milieu. Yet the pilot studies suggest that it is in the teaching of the scientific bases of modern agriculture that schooling is most effective in transforming the ways that individuals think about and practice farming (Eisemon 1989). Similar observations may be made about how schooling may change health and nutritional practices.

The separation of scientific from practical instruction in the program of studies is reinforced by the use of French to teach etude du milieu while Kirundi is used for most practical subjects. Though practical subjects are officially to be taught in French in grades five and six, they are in fact usually taught in Kirundi since they are not examined in the concours national (Schwille, Eisemon and Prouty 1989). On the BRIDGES tests, students exhibited little understanding of the scientific information that the question items elicited in French, but great improvement when Kirundi was used. This contrasts with the results of the mathematics tests where language effects were less pronounced. Much French mathematical vocabulary has been absorbed into Kirundi, perhaps because it is used for instruction in mathematics prior to grade five. The use of French for teaching most science in primary schools

may inhibit the development of Kirundi as a language of ordinary scientific discourse and, in consequence, the dissemination of modern science.

Progress toward Kirundization, an important objective of the 1973 educational reform, has been impeded by concern for the French language skills of students who are admitted to secondary schools and who have experienced the double shift system. Whether today's graduates have less facility in French than earlier students is the subject of much controversy. Our study indicates that most students have poor French language skills, particularly in composition. Whether double shifts have adversely affected the quality of entrants into secondary schools cannot be determined from our data. That many students are unable to benefit fully from instruction in French has been well established.

Elsewhere (Eisemon and Schwille in press), we have proposed improvement and expansion of the teaching of French in the lower stage of the primary cycle while at the same time the use of Kirundi is expanded in grades five and six. This would imply more of a balance between the two languages throughout primary school. The government took an important step in this direction in 1989 by deciding to introduce French as a subject in grade one instead of postponing teaching it until grade three. Nevertheless, additional measures may be needed to prepare students for the teaching of academic subjects in French from grade five, perhaps using French for some academic instruction, such as mathematics, in grades three and four. These measures should also be seen as prerequisites for strengthening the role of Kirundi which is important for the majority of primary school students who will not continue their studies.

Improving assessment of student learning. At present, the Ministry of Primary and Secondary Education has no adequate mechanism for monitoring the learning of all students. The concours national does not serve this purpose. The concours national is intended for selection and is constructed to identify students who will be successful in academic studies offered in French in secondary schools. Teachers also develop and administer end-of-primary-school leaving examinations to students finishing sixth grade. The latter generally cover instruction in academic subjects, but they are not standardized across schools. And many students do not take the tests if they have no intention of returning to school to repeat the sixth year and retake the concours national.

To improve the quality of primary schooling in Burundi will require changes in the concours national. It must be acknowledged that, as long as secondary schooling remains highly selective, the secondary school entrance examination will greatly influence the instruction of all primary school students whether or not they go on to secondary school. If those who are not going on to secondary school are to benefit from primary school, the design of the concours national must take them into account, too.

Therefore, in our view, the examination could be designed for three purposes: to select students for secondary school as is presently done; to certify satisfactory completion of primary school; and to provide feedback to parents, students, teachers, principals and policymakers on what all students are learning. The present end-of-school leaving examination, which differs from school to school and whose quality is therefore hard to ensure, could be abolished.

To implement the different purposes of the examination, different pass rates could be used. The highest cutoff score might be used to identify

students going on to secondary school; it could continue to be based on the number of places available in secondary school. Another level might be used to signify successful completion of primary school. This cutoff should be domain-referenced in the sense of representing as nearly as possible that minimum level of knowledge that students need in order to benefit from primary schooling. The examination should cover all those domains judged important for primary school students to learn, including those elements of health, nutrition and agriculture-related science that are particularly needed by students who will not be going on to secondary school. In all cases, students and parents should be informed of the scores and given help to understand what they mean.

In addition, the present concours national tests of comprehension and knowledge of language production rules in French and Kirundi might be redesigned. Comprehension tasks should be referenced to various text genres and measure different kinds of information processing skills (Eisemon, Patel and Abagi 1988). Moreover, the multiple choice format, while suitable for assessing comprehension is less adequate for measuring knowledge of production rules. Production tasks are better though more difficult to mark. Our findings indicate that it is in the application of production rules for written communication that students experience the most difficulty, especially in a second language. Teachers are not likely to emphasize the development of skills that are not examined.

Strengthening teacher supervision and improving inservice training. We have drawn attention to the importance of teacher supervision. Now that universal primary education is close to being achieved, it may be useful to consider whether present school clusters permit effective supervision. The

number of satellite schools, their size, and the distance from core schools all influence the amount of attention they receive from school directors. Transportation to satellite schools, for instance, is a serious problem for many school directors who either do not have or have only limited access to vehicles, motorcycles or even bicycles to enable them to carry out their responsibilities. Some improvement in school management is likely to occur if transportation is provided to them.

However, strengthening school management may also require more school directors. In comparison to school headmasters in several other East African countries, school directors in Burundi have greater administrative responsibilities, both in scope and substance. In Kenya, for example, school inspectors and the staff of teacher advisory centers are mainly responsible for teacher supervision and inservice training while in Burundi, the school directorships are the locus of these activities.

Improving the quality of teaching is fundamental to increasing school effectiveness and reducing variability in student achievement. Conventional approaches to inservice training focus on improving teacher knowledge of subject matter and/or demonstrating "effective" methods of lesson presentation. Nevertheless, good teaching involves more than understanding the fichiers and being able to present them. It also involves an understanding of what is expected of students, why students make errors, and how to improve their performance.

Involving teachers in the examination process in suggesting items and marking papers can be important in fostering a better understanding of assessment tasks and the difficulties students have in performing them. The knowledge gained from these activities is excellent inservice education and

will help them improve their instruction. That is the experience of many countries where this is practiced (Eisemon and Schwille 1988).

Achieving appropriate rates of repetition. The high repetition rates at the grade six level in primary schools are an outcome of policies intended to increase educational participation and attainment in primary schools but without greater opportunity for secondary education. The rigor of the concours national and until 1989, the lack of feedback to students on their results, encourages the belief that the likelihood of success is increased with subsequent trials. We have no way of confirming or refuting this belief with the data collected in the present study. Nevertheless, for students who repeat, it is clear that repetition raised their levels of performance in the domains which we examined and which are also tested in the concours national.

Repetition probably increases students' ability to learn in French and in doing so, to perform better on examinations. Apparently, the shift from Kirundi to French as the medium of instruction in grade five is too abrupt for many students for whom repetition is a way of catching up. The retention of French as the medium of instruction in grades five and six and for examination for admission to secondary school favors the persistence of high repetition rates. Because students are not given unique national inscription numbers in the first grade, repetition cannot be easily monitored. The introduction of national inscription numbers would facilitate further research on the effects of repetition, which in turn could suggest appropriate measures of control.

In 1989, more than 73,000 students sat for the concours national and only about 7,000 received places in secondary schools. Between thirty and forty thousand students would be expected to repeat the sixth grade. It will require considerable expenditure of scarce resources to educate them.

Therefore, a limit on the number of times a student is allowed to repeat might be established, based on a minimum score on the secondary school entrance examination. Or the proportion of secondary school admissions awarded to students who retake the entrance examination might be limited to further discourage repetition. This, in effect, would make the threshold score higher for repeaters than for nonrepeaters. But reductions in repetition make sense only if considered in conjunction with other changes designed to increase learning or at least to ensure the same level of learning outcomes at less cost.

Needed Experimentation

In large part, these findings are a vindication of conventional thinking about what makes for effective primary schooling in Burundi, such as the value of principals making frequent classroom visits and the positive effects of repetition. However, given the shortage of resources in Burundi and the difficult conditions under which many teachers and principals work, just doing more of the same is not likely to be effective in strengthening primary schooling. The variables found to have a positive impact may be nonlinear in their effects, resulting in diminishing returns. Moreover, in some respects our findings deviate from conventional thinking. For example, conventional wisdom might suggest that weak students suffer most from assessment in French. But our results indicate that testing in French may have a still more negative effect on evaluating the ability of the most able students.

It is important to proceed with careful experimentation in pilot projects to see how mean levels of student performance might be increased and variation in performance reduced. Teachers and principals can play an important role in making plans and developing projects to answer questions such as the

following: Which students should be allowed to repeat, at what grade level and how often? What modifications of the double shift system should be encouraged? How can one get principals who do not make enough classroom visits to make more of these visits and to monitor teaching more effectively? What changes in incentives are needed and feasible to increase teacher engagement, reducing absenteeism and lateness? What combinations of French and Kirundi across subject-matters and grades would bring about the best balance of the two languages so that the foundations of literacy are established as much as possible in both languages?

To take the double shift policy as an example, it should be possible to suggest variations in double shifts that would make the system both more effective and less onerous. Different patterns might be developed for different situations, depending on such factors as number of students, number and size of classrooms available, and so on. For example, shifts might be combined on certain days of the week and kept separate on others. If there were 5 hours of combined shifts on Monday, Wednesday and Friday (instead of 3.5 hours for each of two shifts); 4 hours of combined shifts on Saturday; and 2 shifts of 2.5 hours each on Tuesday and Thursday, this schedule would result in an increase in instruction for each student (24 hours instead of the present 19.5) and a reduction in the hours of teaching for teachers (29 hours instead of 39). But this is only an example. The best and most feasible arrangements remain to be determined through planned experimentation based on input from teachers, principals and ministry staff.

Likewise, in language policy much experimentation is indicated to follow up the results of this study and to arrive at a better balance between the two languages. The results of the study point to both advantages and

disadvantages of using Kirundi more as a language of instruction. These results are evidenced not only by interviews but by the findings obtained from the analyses of tests and classroom observations. Pilot studies could serve to test variations in the use of Kirundi and French for instruction and assess the consequences for student learning. Our results suggest that mathematics might be taught earlier in French while integrated science might be taught throughout primary school in Kirundi.

A common element to these experiments or pilot projects would be a more adequate means of assessing the learning of all students. This includes not only the students who are going on to secondary school, but those who are not. Our study is a first step in that direction.

References

Alexander, L., & Simon, J. (1975). The determinants of school achievement in developing countries: The educational production function. Washington, DC: World Bank.

Almond, G.A., & Coleman, J.S. (1960). The politics of developing areas. Princeton, NJ: Princeton University.

Berry, J.W. (1985). Learning mathematics in a second language: Some cross-cultural issues. Learning of Mathematics, 5, 18-23.

Bracewell, R.J., Frederiksen, C.H., & Donin-Frederiksen, J. (1982). Cognitive processes and comprehending discourse. Educational Psychologist, 17, 146-164.

Cole, M., & Means, B. (1981). Comparative studies of how people think: An introduction. Cambridge, MA: Harvard University Press.

Eisemon, T.O. (1988). Benefiting from basic education: School quality and functional literacy in Kenya. Oxford: Pergamon.

Eisemon, T.O. (1989). Becoming a modern farmer: The impact of primary schooling in Kenya and Burundi. In D.M. Warren, L.J. Slikkerveer, & S.O. Titilola (Eds.), Indigenous knowledge systems: Implications for agriculture and international development (pp. 41-67). Ames, IA: Iowa State University.

Eisemon, T.O., Patel, V., & Abagi, J. (1988). Read these instructions carefully: Examination reform and improving health education in Kenya, International Journal of Educational Development, 8, 55-66.

Eisemon, T.O., & Schwille, J. (1988, November). Examinations policies to strengthen primary schooling in African countries. Paper prepared for World Bank Seminar on Using Examinations and Standardized Testing to Improve Educational Quality, Lusaka, Zambia.

Eisemon, T.O., & Schwille, J. (in press). Should schools prepare students for secondary education or for self-employment? Addressing a dilemma of primary schooling in Burundi and Kenya. Elementary School Journal.

Eisemon, T.O., Schwille, J., & Prouty, R. (1989, January). Does schooling make a better farmer?: Schooling and agricultural productivity in Burundi. Paper presented at the annual meeting of the BRIDGES Project, Bangkok.

Eisemon, T.O., Schwille, J., & Prouty, R. (in press). What language should be used for teaching?: Language policy and school reform in Burundi. Journal of Multilingual and Multicultural Development.

Frederiksen, C.H., Bracewell, R.J, Breuleux, A., & Renaud, A. (1988). The cognitive representation and processing of discourse: Function and dysfunction. In Y. Joanette & H. Brownell (Eds.), Discourse ability and brain damage: Theoretical and empirical perspectives. New York: Springer-Verlag.

Gay, J., & Cole, M. (1967). The new mathematics and an old culture. New York: Holt, Rinehart and Winston.

de Glopper, K. (1988). The results of the international scoring sessions. In T.P. Gorman, A.C. Purves, & R.E. Degenhart (Eds.), The IEA study of written composition 1: The international writing tasks and scoring scales (pp. 59-75). Oxford: Pergamon.

Greenland, J. (1974). The reform of education in Burundi: Enlightened theory faced with political reality. Comparative Education, 10, 57-63.

Heyneman, S.P., & White, D.S. (1986). The quality of education and economic development. Washington, DC: World Bank.

Joreskog, K.G. (1978). Structural analysis of covariance and correlation matrices, Psychometrika, 42, 443-477.

Joreskog, K.G., & Sorbom, D. (1986). LISREL VI: Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least squares methods [computer program]. Mooresville, Indiana: Scientific Software, Inc.

Lemon, N. (1981). Language and learning: Some observations on the linguistic determination of cognitive processes. In J. Gay & B. Lloyd (Eds.), Universals of human thought (pp.201-214). Cambridge, MA: Cambridge University Press.

Lockheed, M.E., & Komenan, A. (in press). Teaching quality and student achievement in Africa: The case of Nigeria and Swaziland. Teacher and Teacher Education.

Mamari, G.R.V. (1975). Tanzania's experience in and efforts to resolve the problem of teaching mathematics through a foreign language. In Accra workshop proceedings: The problems of developing a workable language policy as the medium of instruction changes, mimeo.

Ministry of National Education, Republic of Burundi. (1988a). Politique et stratégies pour améliorer la qualité de l'éducation de base [Policies and strategies for improving the quality of basic education]. Bujumbura: Author, mimeo.

Ministry of National Education, Republic of Burundi. (1988b). Rapport final: Séminaire national de réflection sur l'évaluation systématique de l'enseignement primaire, 16-21 mai 1988. [Final report: National seminar of deliberation on the systematic evaluation of primary education, May 16-21, 1988]. Bujumbura: Author.

Mwamwenda, T.S., & Mwamwenda, B.B. (1989). Teacher characteristics and pupils' achievement in Botswana primary education. International Journal of Educational Development, 9, 31-42.

Proceedings of the Seminar of Deliberation on Improving the Internal and External Effectiveness of Primary Education. (1989). East Lansing, MI: Michigan State University, USAID BRIDGES Project, in collaboration with the Ministry of Primary and Secondary Education, Bujumbura, Burundi and the Centre de Perfectionnement et de Formation en Cours d'Emploi (CPF), Bujumbura, Burundi.

Purves, A.C., Gorman, T.P, & Takala, S. (1988). The development of the scoring scheme and scales. In T.P. Gorman, A.C. Purves, & R.E. Degenhart (Eds.), The IEA study of written composition 1: The international writing task and scoring scales (pp. 41-58). Oxford: Pergamon.

Ramsay, J.O., & Abrahamowicz, M. (1989). Binomial regression with monotone splines: A psychometric application. Unpublished paper, McGill University, Montreal.

Reed, H.J., & Lave, J. (1979). Arithmetic as a tool for investigating relations between culture and cognition. American Ethnologist, 6, 568-582.

Sapir, E. (1963). Language. London: Rupert Hart-Davies.

Schwille, J., Eisemon, T.O., & Prouty, R. (1989, April). Between policy and students: The reach of implementation in Burundian primary schools. Paper presented at the annual meeting of the Comparative and International Education Society, Boston.

Schwille, J., Eisemon, T.O., & Prouty, R. (1990, March). How wasteful is grade repetition? Unanswered questions posed by test scores of rural sixth grade repeaters in Burundi. Paper presented at the annual meeting of the Comparative and International Education Society, Los Angeles.

Whorf, B.L. (1956). Language and Reality. New York: MIT Press and Wiley.

World Bank. (1988). Education in Sub-Saharan Africa: Policies for adjustment, revitalization, and expansion. Washington, DC: World Bank.

Zaslavsky, C. (1973). Africa Counts. Boston: Prindle, Weber and Schmidt.

Zepp, R.A. (1982). Bilinguals' understanding of logical connectives in English and Sesotho. Educational Studies in Mathematics, 13, 205-221.

Table 1
Student Performance in Reading Comprehension, Written Composition, Mathematics and Science-Agriculture: French vs. Kirundi as Language of Assessment

Test:	Mean Scores			F value
	Standard French	Simplified French	Kirundi	
Reading comprehension total test (15 items)	6.04	6.53	8.05	F=44.6***
Narrative text subtest (8 items)	3.02	3.36	4.27	F=47.86***
Procedural text subtest (7 items)	3.01	3.18	3.78	F=21.48***

	French	Kirundi	F value
Composition Test (ratings 0-10)			
Overall quality	2.27	4.11	F=137.32***
Coherency of narrative	1.96	3.06	F=60.28***
Use of language	2.14	4.01	F=147.11***
Mathematics (19 items)	8.65	8.19	F=2.56
Science-agriculture (19 items)	7.16	9.31	F=168.67***

* $p<.05$
** $p<.01$
*** $p<.001$

Table 2
Mean Test Score Differences by Test: Repeaters Versus Nonrepeaters at Sixth Grade

Test	Mean test scores		F value
	Repeaters	Nonrepeaters	
French mathematics (19 items, n=469)	10.46	6.06	136.80***
Kirundi mathematics (19 items, n=486)	9.69	6.40	81.03***
French reading, comprehension of standard text (15 items, n=269)	6.94	4.99	34.49***
French reading, comprehension of simplified text (15 items, n=241)	7.78	5.19	57.51***
Kirundi reading comprehension (15 items, n=481)	9.09	6.79	74.01***
French composition: general quality (ratings 0-10, n=474)	3.33	1.10	154.64***
Kirundi composition: general quality (ratings 0-10, n=451)	4.83	3.24	49.71***
French composition: coherence (ratings 0-10, n=474)	2.80	1.02	116.17***
Kirundi composition: coherence (ratings 0-10, n=451)	3.64	2.36	37.20***
French composition: use of language (ratings 0-10, n=474)	3.14	1.03	144.45***
Kirundi composition: use of language (ratings 1-10, n=451)	4.71	3.17	47.35***
French science-agriculture (19 items, n=469)	7.82	6.20	51.94***
Kirundi science-agriculture (19 items, n=486)	9.74	8.79	16.57**

* $p < .05$
** $p < .01$
*** $p < .001$

Table 3
Grade 6 Repetition by Type of School

	Core w/o Satellites N=374	Core w/ Satellites N=841	Satellites N=727
Grade 6 repeaters	62%	58%	48%
Grade 6 nonrepeaters	38%	42%	52%

$X^2 = 23.61$, $p < .001$

Table 4
Supervisory and Instructional Correlates
of Mean School Test Results (N=47)

Characteristics	Language arts French r=	Language arts Kirundi r=	Math & sci French r=	Math & sci Kirundi r=
No. of class visits	.35**	.43***	.41**	.47***
Teacher punctuality	.03	.32*	.28*	.46***
Application of language policy	.27*	.28*	-.08	-.05
No. of times shifts combined past two weeks	.23	.25*	.13	.06
No. of extra hrs. of teaching past two weeks	.21	.17	.18	.26*
No. of lessons skipped past two weeks	-.14	-.10	.06	.12
No. of days required to finish syllabus	.00	-.34**	.04	-.03
No. yrs. of experience in grade six	.06	.19	.22	.30*
No. of demon-stration lessons	.21	-.15	.17	.32*
Proportion of students last year who passed concours	.23	.30*	.34*	.28
Teacher assessment language arts	.12	.11	-	-
Teacher assessment math & science	-	-	.21	.39**

* $p < .05$
** $p < .01$
*** $p < .001$

FIGURE 1

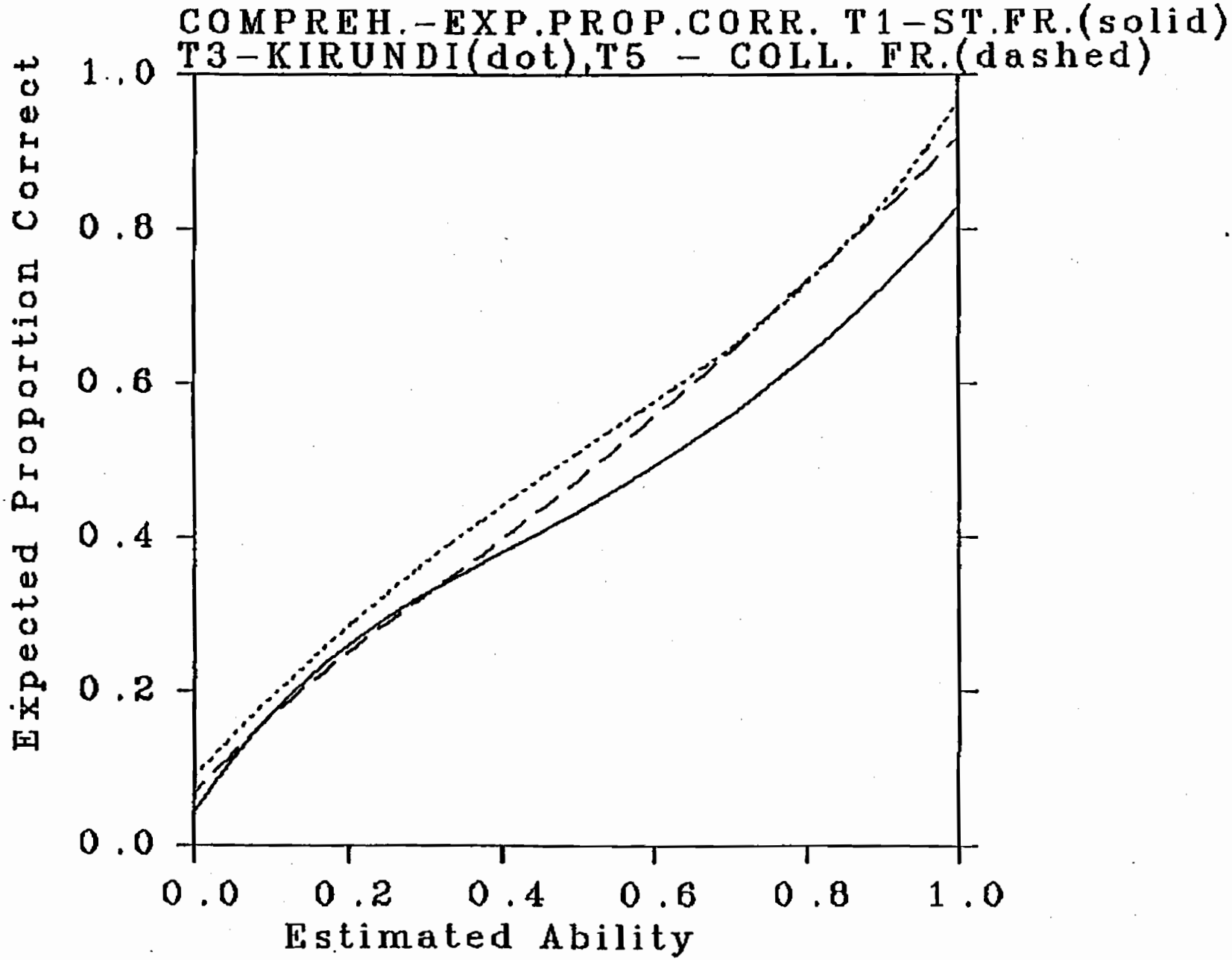

COMPREH.-EXP.PROP.CORR. T1-ST.FR.(solid)
T3-KIRUNDI(dot),T5 - COLL. FR.(dashed)
Expected Proportion Correct
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Estimated Ability

FIGURE 2

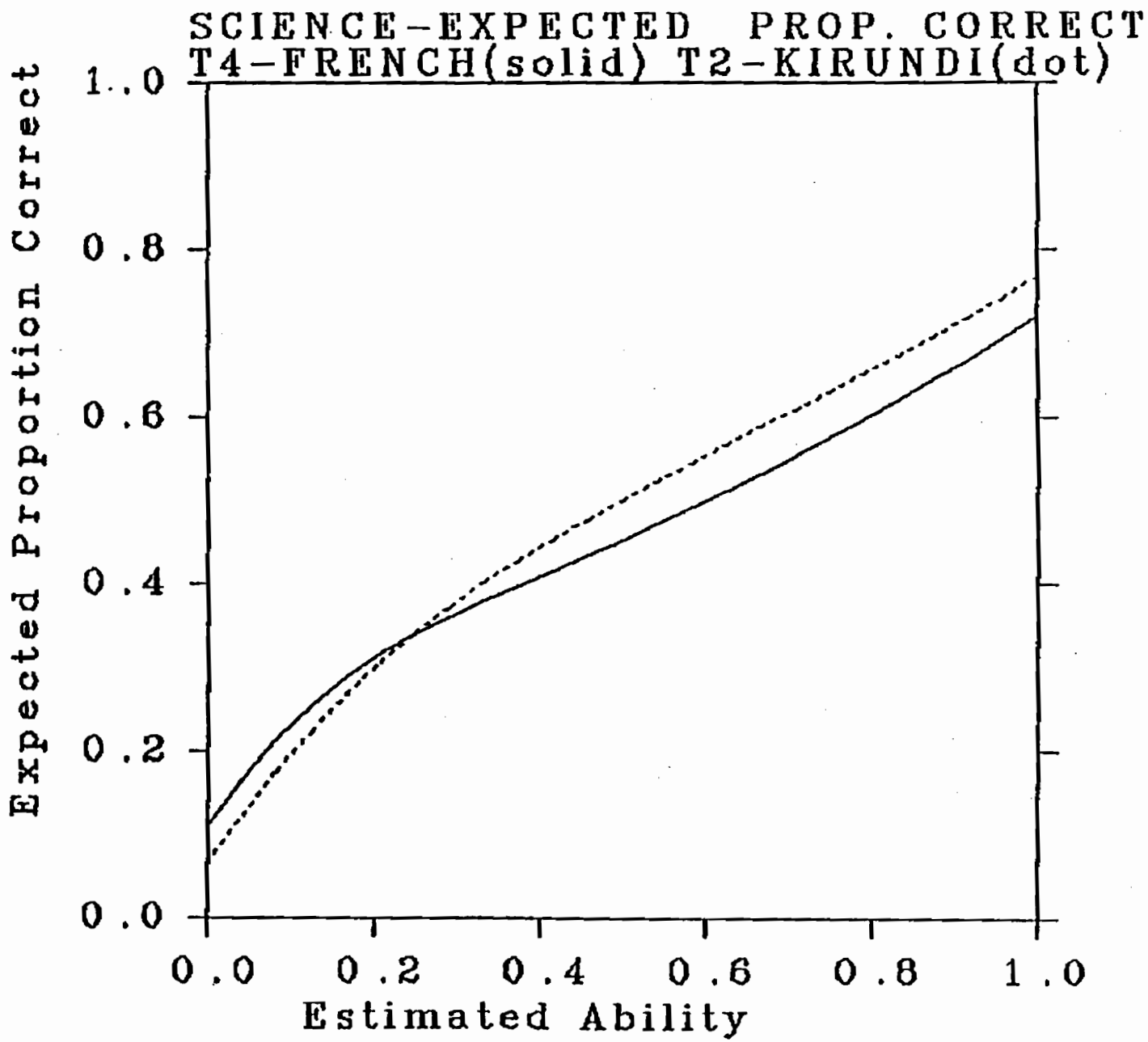

SCIENCE-EXPECTED PROP. CORRECT
T4-FRENCH(solid) T2-KIRUNDI(dot)
Expected Proportion Correct
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Estimated Ability

FIGURE 3

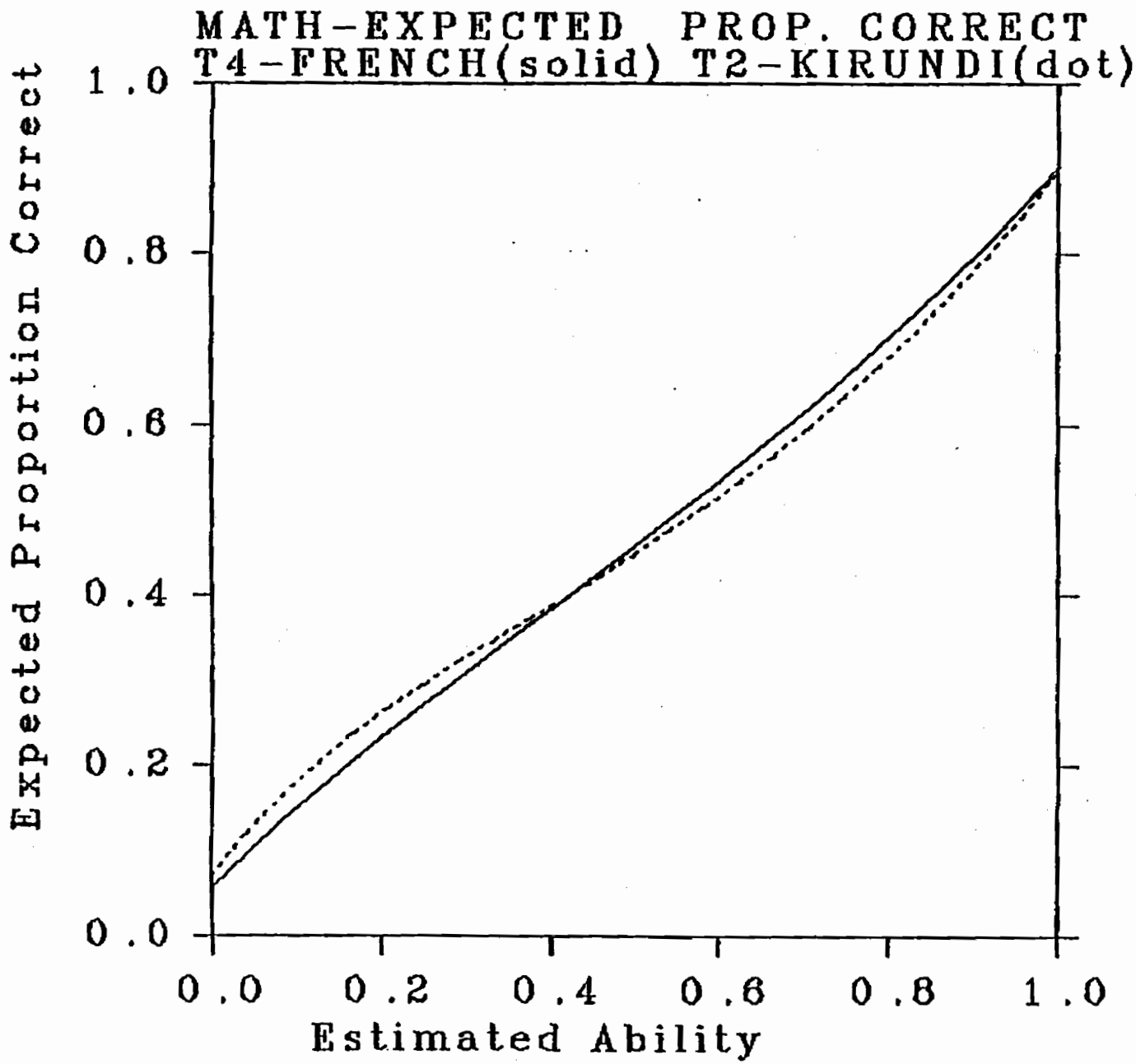

MATH-EXPECTED PROP. CORRECT
T4-FRENCH(solid) T2-KIRUNDI(dot)
Expected Proportion Correct
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Estimated Ability

Figure 4

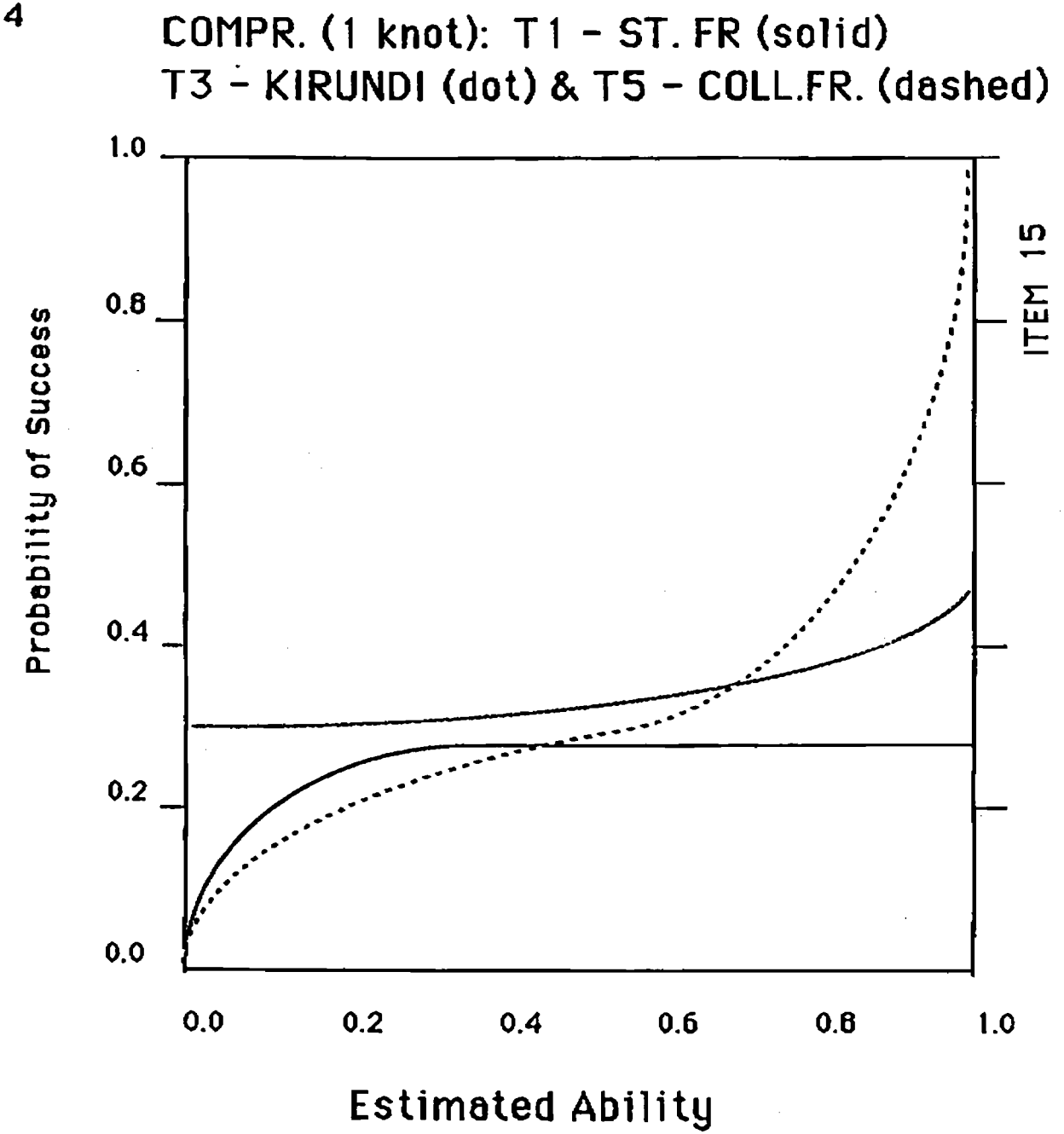

COMPR. (1 knot): T1 - ST. FR (solid)
T3 - KIRUNDI (dot) & T5 - COLL.FR. (dashed)
ITEM 15
Probability of Success
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.2
0.4
0.6
0.8
1.0
Estimated Ability

FIGURE 5 A GENERAL PATH MODEL OF SCHOOL EFFECTIVENESS

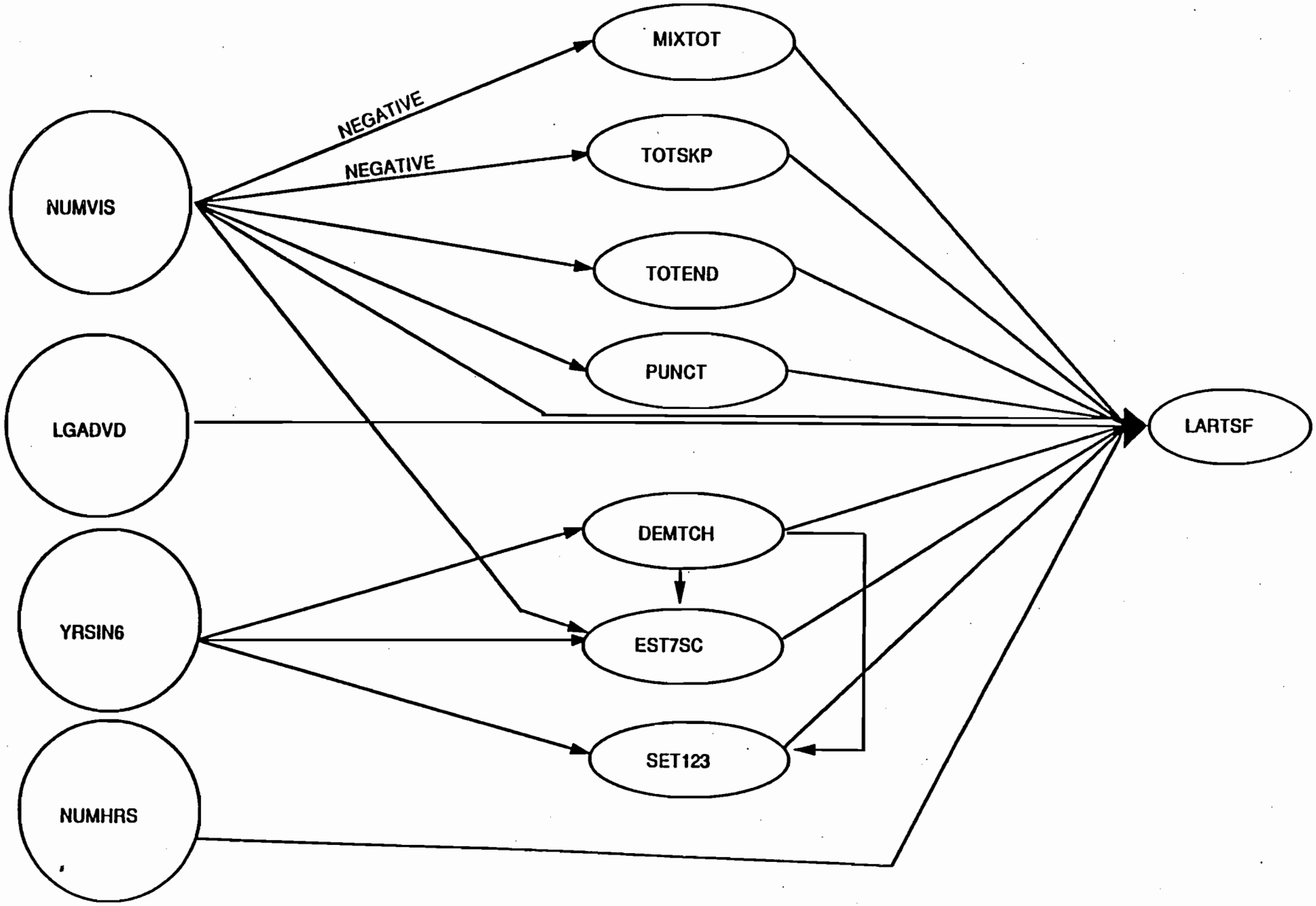
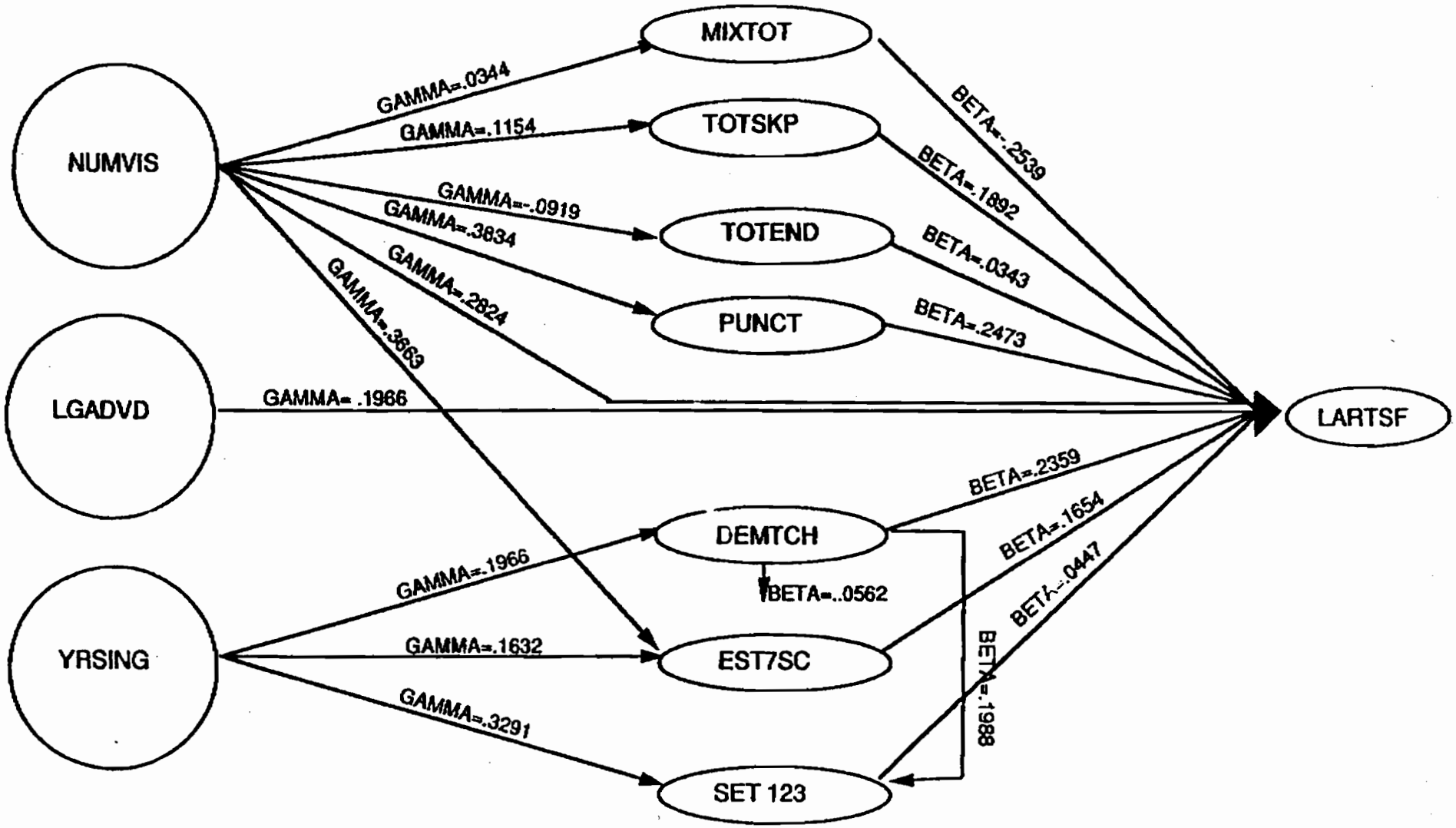


FIGURE 6 PATH MODEL FOR ACHIEVEMENT IN FRENCH LANGUAGE ARTS



x^2=35.57, df=33, p=.348