World Bank Reprint Series: Number 391

Dean T. Jamison, Barbara Searle, Klaus Galda, and Stephen P. Heyneman

Improving Elementary Mathematics Education in Nicaragua

An Experimental Study of the Impact of Textbooks and Radio on Achievement

Journal of Educational Psychology
1981, Vol. 73, No. 4, 556–567

0022-0663/81/7304-0556$00.75

Improving Elementary Mathematics Education in Nicaragua: An Experimental Study of the Impact of Textbooks and Radio on Achievement

Dean T. Jamison
The World Bank, Washington, D.C.

Barbara Searle
Stanford University

Klaus Galda
Stanford University

Stephen P. Heyneman
The World Bank, Washington, D.C.

Because widespread availability of textbooks in the United States preceded research on the effectiveness of instructional materials, there has been little systematic study of their impact on student achievement. The developing world provides an appropriate setting for such studies. This article reports an experimental study of the impact of textbook availability on mathematics achievement of students in Nicaraguan first-grade classes. This intervention is compared with control classes in which textbooks are relatively rare and with a radio-based instructional program that uses student worksheets but no other textual material. Classes were assigned at random to the three conditions. The control and two treatment groups scored similarly on a pretest of mathematical readiness. Both the textbook and the radio treatments had significant positive effects on achievement. Availability of textbooks increased student posttest scores by about 3.5 items correct, approximately .33 of a standard deviation. Availability of the radio instructional program increased student posttest scores 14.9 items, about 1.5 standard deviations. Both interventions reduced the achievement gap between urban and rural students. However, the question remains whether either radio or textbook use is sufficiently powerful to close the substantial achievement gap that exists between the schools of high- and low-income societies.

Currently many children in the developing world never enter school. However, enrollments have been expanding rapidly, and some projections suggest that most children will have a place at the first-grade level by the end of the century (Meyer & Hannan, 1979). It seems likely that the central concern of educational systems and international agencies that contribute to their support will shift from the difficulties of getting pupils into classrooms to the problem of what pupils find in the classroom after they have entered. Thus, it becomes crucial to identify factors that can increase the quality of primary education at an affordable cost.

Primary schools in the developing world are typically of poor quality in both inputs and outcomes. On the average, teachers, furniture, equipment, and materials are below standards considered minimum in the industrialized societies. In 1977 there were 10 pupils for each available primary-school textbook in the Philippines (Jamison & Montenegro, in press); in 1978 the value of furniture and materials in the average fourth-grade classroom in Bolivia was 1% of that of Maryland classrooms (Heyneman, Note 1). A 1979 survey conducted in Malawi found one chair for each 8 pupils and one

The research reported here was supported principally by World Bank Research Project RP0671-60, entitled "Textbook Availability and School Quality." It could not have been undertaken, however, if it had not been imbedded in the joint United States Agency for International Development (U.S. AID)–Nicaragua Ministry of Education Radio Mathematics Project, supported in part by U.S. AID Contract AID/CM/ta-C-73-40 to Stanford University.

The authors are grateful to those who were part of the project, and to the agencies supporting it, for their collaboration. The views and conclusions expressed in this article are, however, those of the authors and do not necessarily reflect those of any of the supporting agencies. The authors also wish to thank Jamesine Friend and Juan Jose Montenegro for their contributions to the work reported here.

Requests for reprints should be sent to Dean T. Jamison, The World Bank, 1818 H Street, N.W., Washington, D.C. 20433.

desk for each 88 pupils (Heyneman, 1980). Furthermore, the gap in classroom quality between high- and low-income countries is widening; as more pupils enter school, there is less available for each of them. In 1960 the average country in the Organization for Economic Co-operation and Development invested 16 times more per pupil than did any of the 36 countries with per capita incomes below $265 (U.S.) per year. By 1970 the difference had grown to 22 : 1, and by 1975 to 31 : 1 (World Bank, 1980a).[1]

The existing evidence indicates that students in the developing world do not learn as much in primary school as their peers in the industrialized world (World Bank, 1980b, pp. 52–53). A typical finding emerging from the International Evaluation of Educational Achievement, for example, suggests that the mean score on science tests for students in low-income countries is in the bottom 5% to 10% of students from high-income countries (Inkeles, 1977).

Evidence from the United States (Coleman et al., 1966; Jencks et al., 1972; Jencks et al., 1979) has suggested that variation in the level of school inputs is not strongly associated with levels of student attainment. However, a growing body of work from the developing world suggests that the situation there is different. Differences in classroom quality, as measured by physical facilities, availability of materials, and levels of teacher education, appear to be surprisingly robust as predictors of student achievement. The correlation between the proportion of variance in science performance explained by classroom quality and national per capita income is approximately −.37 ($p < .05$) (Heyneman & Loxley, in press), suggesting that the poorer the country, the higher the impact of classroom quality on achievement.[2] This may result from a much higher variance in the quantity and quality of school inputs in low-income countries than in high-income ones.

Which inputs can be expected to raise student achievement in the developing world? The evidence is relatively scarce. A careful literature search located fewer than 30 studies on the cognitive impact of teacher training in low-income countries, in comparison with almost 400 published in the United States in 1 year,[3] and similar imbalances exist with respect to the evidence on textbooks, duplicating machines, audio-visual aids, radio, television—in fact, with respect to all school resources. Furthermore, results from available studies are not unambiguous. Smaller class size and longer teacher-training programs are not always associated with higher achievement levels.

One of the more consistent indicators of higher achievement is the availability of textbooks and other printed materials (Heyneman, Farrell, & Sepulveda-Stuarto, 1978; Heyneman & Jamison, 1980). Table 1 lists 14 studies that report 19 assessments of the relationship between the availability of printed material and student outcomes. Although different methodologies were used and the quality of the data is far from uniform, the positive relationship between input and outcome in 16 of the 19 cases strongly suggests that textbooks are a potentially important and consistent contributor to improved quality in schools. However, the nonexperimental design of these studies often leaves open the possibility, indeed plausibility, of alternative interpretation of the cause for observed correlations between textbook availability and achievement. For this reason we felt it would be desirable to undertake a carefully controlled experiment

[1] If one attempts, however, to adjust for international variations in purchasing power that are not captured in computation of exchange rates, the ratio of spending in industrialized to nonindustrialized societies would decline. For example, at the official exchange rate, the 1975 per capita income of the United States was approximately 53 times that of India; after adjusting for differences in purchasing power, however, this ratio declines to 14 (World Bank, 1979, p. 177).

[2] There is, moreover, some evidence that in developing countries the quality of a primary or secondary school is a robust predictor of a person's success in the labor market, substantially more powerful, for example, than is an individual's socioeconomic status (Heyneman, 1980; Heyneman & Currie, 1979; Fry, 1980; Schiefelbein & Farrell, Note 2). These studies are too recent and too few to generalize these results with confidence.

[3] The Educational Resources Information Center system lists 388 titles published on this subject in the United States in 1977. A recent review of the evidence from low-income countries located 23 studies published between 1963 and 1977 (Husen, Saha, & Noonan, Note 3). A subsequent review, which made a specific effort to locate studies published in non-European languages, found a slightly higher number (Avalos & Haddad, 1981).

Table 1
Availability of Reading Material and Student Outcomes in Low-Income Countries: Synopsis of the Literature

Author	Country	School level	Outcome measures	Relationship[a]	Comments
Beebout (1973)	Malaysia	Grades 10, 11	Achievement gains in secondary school; used national examinations	+	Random sample of 89 schools (N = 7,674)
Comber & Keeves (1973)	Chile	Primary 10-year-olds	Science achievement	+	International Evaluation of Educational Achievement
	India			+	
	Iran			−	
Farrell & Schiefelbein (1974)	Chile	Primary Grades 6–8	Math, language achievement	+[b]	National sample of primary schools
			Educational aspirations	+	
		Same cohort, 4 years later	Type of secondary school entered	+[b]	
			Survival to end of secondary school	+	
Fuller & Chantavanish (Note 7)	Thailand	Primary Grade 3	Language & math achievement	+[b]	
Haron (1978)	Malaysia	12- to 13-year-olds	Bahasa Malaysia, English, math, science, history, geography	+[b]	National probability census sample (N = 6,000)
Heyneman & Jamison (1980)	Uganda	Grade 7	National tests in English, math, general knowledge; Ravens Progressive Matrices used as ability control	+	Results stronger for English than math or general knowledge; random sample of primary schools in 8 districts
Jamison & Montenegro (in press); Heyneman & Montenegro (in press)	Philippines	Grades 1, 2	Filipino, math, science achievement	+[b]	National reduction of pupils to textbooks ratio from 10:1 to 2:1; gains over 1 year ranged from .18 to .51 SDs; national sample of primary schools
Lynch (Note 8)	Ecuador	Grade 1	Reading, math, science achievement	0	88 purposively sampled classrooms; assessed new texts
ODEPOR (Note 9)	El Salvador	Grades 2, 3, 5, 6, 8, & 9	Spanish, math, social & natural science	+	National probability sample of schools; household data
Simmons & Askoy (Note 10)	Tunisia	Individual students, Grades 4–8	Arabic, French, arithmetic achievement	+	44 village students; 80 students from urban suburb
Smart (1978)	Ghana	Grades 8, 9	Reading comprehension	+	40 rural experimental and control schools; used pre- and posttests over 2 years; treatment was access to a school newspaper

Table 1 (*continued*)

Author	Country	School level	Outcome measures	Relationship[a]	Comments
Thorndike (1973)	Chile Iran	Primary 10-year-olds	Reading comprehension	+ –	International Evaluation of Educational Achievement
Wolff (Note 11)	Brazil Rio Grande do Sul	Primary Grade 1	Promotion to Grade 2 based on teacher assessment	+[b]	Random selection of 5% of schools (N = 20,000)

Note. ODEPOR = Oficina de Planeamiento y Organizacion.
[a] This column refers to relationship between availability of printed material and student outcomes (+ = positive relationship; 0 = no relationship; – = negative relationship). [b] Stronger with low socioeconomic status and/or rural students.

assessing the impact of increased textbook availability on student learning. The opportunity to do this arose during the course of an experiment designed to investigate the effectiveness of a radio-based instructional program. This experiment, known as the *Radio Mathematics Project*, was conducted in Nicaragua from 1974 to 1978 (Friend, Searle, & Suppes, 1980).

The Radio Mathematics Project provided an ideal setting for undertaking such an experiment. Not only were random assignment and testing procedures already in place, but inclusion of the textbook experiment in the radio project offered the possibility of comparing the two interventions with one another as well as with the status quo. Previous evaluations had shown that the radio instructional program has a strong positive influence on children's mathematics achievement in Grades 1–3, and particularly in Grade 1 (Searle & Galda, 1980). Thus, first grade was chosen as the major focus of the textbook intervention. This article compares mathematics achievement levels of control classes with classes provided with radio lessons and adequate textbooks.

Educators have become increasingly aware that placing new or additional instructional materials in classrooms does not guarantee their use. Their research reported here attempts to assess the impact of materials under naturalistic conditions. That is, the level of effort expended in encouraging teachers to use the new materials was limited to that considered feasible given the administrative capabilities and the resources available in Nicaragua. Thus, the work does not assess the potential that textbooks or radio lessons have for improving student achievement under optimal circumstances. Rather, it attempts to assess their impact as they might be adopted in the typical developing country.

Method

Subjects

In earlier work conducted by the Radio Mathematics Project, all primary schools in three provinces of Nicaragua were classified as *radio* or *control*, using a random sampling process stratified by urbanization (Searle & Galda, 1980). For the present experiment 88 first-grade classrooms were selected at random, 48 from radio schools and 40 from control schools. The experiment was conducted in the 1978 school year (February to November). Half of the classrooms from control schools were assigned to the textbook treatment, and the remaining 20 served as controls for both the textbook and radio treatments. Close to 3,000 students participated in the two experiments, but because some children registered late and others dropped out, not all students were enrolled for the entire school year. The fluctuation in numbers of students is discussed below.

Materials

Textbooks. In 1977, 13 rural and 7 urban fourth-grade classrooms were surveyed to determine the number of mathematics texts present and the ways they were used by the teachers. The number of textbooks available varied widely, ranging from 2 to 28 per classroom. However, the high number (28) was found only in one urban and one rural classroom. Four classrooms had between 10 and 14 texts; 12 of the 20 classrooms had

3 or fewer texts. Most teachers reported that they used the textbooks as resources for themselves and did not make them available to the children. In contrast, the experimental treatment provided a text for each child.

In choosing a textbook, all those approved by the Ministry of Education were examined. Only one text series conformed reasonably well to the official curriculum and was available locally in sufficient quantities. This was *Matematica Para la Educacion Primaria*, an adaptation of *Elementary School Mathematics* by Eicholz et al. (1968). The translation and adaptation were designed specifically for use in Central America; the first-grade text is in workbook form.

Teachers were given a textbook for each child as well as the teacher's editions of the texts for Grades 1 and 2, the latter because some topics in the official Nicaraguan curriculum appeared in the second-grade text. A chart cross-referencing topics in the text and in the curriculum guide was prepared to help teachers accustomed to following the guide. In addition, a handout was prepared that explained the differences between the text and the official curriculum and provided suggestions for teaching the topics that did not appear in the text.

Radio lessons. The first-grade radio lessons are described in detail in Searle, Friend, and Suppes (1976). The full course consists of 150 daily mathematics lessons, each in two parts (i.e., a 20- to 30-minute broadcast and a postbroadcast lesson taught by the classroom teacher). Most Grade 1 lessons require individual student worksheets; these were distributed to teachers in three installments during the year. At the beginning of the school year teachers were given a guide describing activities for teachers before, during, and after the radio lessons. The instructional program was designed to cover the topics in the official Grade 1 mathematics curriculum.

Instruments

Mathematics achievement was assessed using a posttest, which is described in detail in Searle and Galda (1980). The posttest uses a matrix-sampling design (Shoemaker, 1973) and has 84 items, one fourth of which are taken by each student. The test was constructed by selecting items from pools defined from the topics in the official curriculum. The number of items for a topic was chosen to be proportional to the time allotted to that topic by the curriculum guide. Since all three instructional programs in the present experiment are based on the same set of objectives, the method for developing the posttest assures a fair assessment of achievement across programs. (Although students taking radio lessons may learn test-taking skills that give them an advantage, this does not seem to explain the score differences reported below. For discussion of this potential difficulty, see Searle & Galda, 1980, p. 112.)

A test of mathematical readiness was administered as a pretest (see Searle & Galda, 1980). The pretest was used to obtain scores that could serve as covariates in the prediction of the posttest score. The pretest scores were also used to check on the comparability of the groups of classrooms randomly assigned to treatments. The test used for the pretest is an adaptation of the mathematics section of the Spanish version of the kindergarten-level *Test of Basic Experiences* (Moss, 1970). The test booklet consists of 28 pages, one item per page; the questions are read aloud. All items are multiple choice, and students respond by drawing a line through the illustration of their choice. In adapting the test, several questions were modified because either the illustration or the text was unsuitable for use in Nicaragua.

Procedure

The experiment was conducted during 1978 in the context of a project that had been in operation in Nicaragua since 1974. Most procedures were those that had been established for implementing and evaluating the Radio Mathematics instructional programs (Searle et al., 1976; Suppes, Searle, & Friend, 1978); in particular, the assignment of classrooms to treatment conditions (see previous *Subjects* section) was based on groups established for the Radio Mathematics Project evaluation.

Teacher orientation. Teachers using radio lessons attended one 3-hour training session before school opened in February 1978. The session described the instructional program and gave teachers guidance in using the radio, the teacher's guides, and the student worksheets. A session of similar duration was held for teachers of classes in the textbook group at which the handouts were distributed and discussed and teachers were given suggestions about using the texts effectively. No meeting was held for teachers of control classes.

Table 2
Variables Used in Regression Analyses

Variable	*n*	*M*	*SD*
	Classroom level		
Pretest score[a]	87	21.9	1.9
Posttest score[b]	87	46.1	8.3
Urban school[c]	88	.477	.502
Radio group[d]	88	.545	.501
Text group[e]	88	.227	.421
Control group[f]	88	.227	.421
	Individual level		
Pretest score	2,070	21.9	4.62
Posttest score	1,699	46.6	11.85

[a] Maximum possible score = 28. [b] Maximum possible score = 84. [c] 1 if school is urban, 0 otherwise. [d] 1 if class assigned to radio group, 0 otherwise. [e] 1 if class assigned to text group, 0 otherwise. [f] 1 if class assigned to control group, 0 otherwise.

Table 3
Numbers of Classes and Students Tested

	Pretest			Posttest		
		No. of students tested/class			No. of students tested/class	
Group	No. of classes	*M*	*SD*	No. of classes	*M*	*SD*
Breakdown by treatment						
Control	20	24.1	6.1	20	18.1	7.3
Textbook	20	24.3	5.3	20	20.7	6.0
Radio	47	23.6	5.3	47	19.6	7.6
Breakdown by urbanization						
Rural	46	20.0	3.0	45	17.4	5.6
Urban	41	28.2	4.1	42	21.8	8.0

Note. One urban radio class was not pretested; one rural radio class was not posttested.

Supervision. Classes assigned to experimental treatments by the Radio Mathematics Project were not supervised or visited by project staff members. This policy had been adopted in evaluating the radio lessons to assess their effectiveness under normal operating conditions. The decision not to supervise treatment groups requires some explanation. An important trade-off exists: It is necessary to monitor program use to assess the degree of implementation of an innovation. However, classrooms in Nicaragua are not normally visited by supervisory personnel during the school year. Visits to check on implementation would have to be made by project personnel, and disguising their purpose would be difficult. Thus, visits made to check on implementation could be expected to influence strongly the degree to which teachers used the radio lessons or the textbooks. As a compromise, the project planned to make one surprise visit to each class assigned to the textbook treatment to determine whether the books were being used. However, the visits were not made because of political turmoil at the time they were scheduled. Instead, the visits were conducted early in the following school year.

Testing. Pretests were administered during the first 3 weeks of school. Posttests were scheduled to be administered in the last 3 weeks of school but were given 2 weeks early because there were rumors that school would close early. All tests were administered by project staff members, not by classroom teachers. To minimize variation in procedures, taped instructions were used for the posttests. Thus, all students heard the same instructions, and timing for the tests was precisely controlled.

Not all students were tested. To reduce copying when a classroom was overcrowded, test administrators were instructed to excuse students by using a systematic procedure that gave each child an equal probability of being excused. For logistical reasons it was not possible to identify for posttesting those students who were pretested. Thus, fluctuations in the student population on pretest and posttest can be attributed to any of several factors: late registration, dropping out, absence from school on the day of the test, and failure to be tested even though present because of overcrowding.

Data Analysis Methods

Major analyses were conducted using the class as the unit of analysis. This is appropriate, since both the assignment to treatment and the implementation were at the level of the classroom (Cronbach, Deken, & Webb, Note 4, chap. 1). Multiple linear regression was used to estimate treatment effects on the posttest score. Regressions were estimated using ordinary least squares. Variations in classroom size, which might have suggested the use of generalized least squares, were regarded as unlikely to have an important effect.

Results

Data

Table 2 presents the variables used in the analyses as well as their means and standard deviations. The numbers of classes tested and the mean number of students in classes for different treatments are shown in Table 3. The number of students pretested (per class) ranged from 9 to 32. For the posttest the range was 6 to 37. (The reader is reminded that in larger classes students may have been excused from testing sessions.)

Means and standard deviations for pretest and posttest are presented in Table 4, at both the classroom and individual level. Pretest scores for the three treatments are not significantly different.[4] This result is expected, since classrooms were randomly assigned to treatments. Pretest scores were obtained primarily to serve as covariates in the regression analyses reported below; that they are comparable across treatments confirms the effectiveness of the randomization. All of the differences in posttest scores between treatment groups at the classroom level are statistically significant. The textbook group scored significantly higher than the control group, t (38) = 2.74, $p < .01$; the radio group scored significantly higher than the control, t (65) = 10.63, $p < .001$, and significantly higher than the textbook group, t (65) = 8.54, $p < .001$. At the individual level, the textbook group scored about .33 of a standard deviation higher than the control group; the radio group scored 1.5 standard deviations above the control group.

Determinants of Posttest Score

Main effects. Table 5 presents the results of a linear regression with posttest score as the dependent variable. The model fits the data quite well, accounting for almost 70% of the variability (R^2 = .67).[5] The variable with strongest effect is radio, the indicator for assignment to the radio group. The regression coefficient for text is also significantly different from zero, indicating a significant treatment effect, although it is weaker than that of the radio treatment.

The coefficient on pretest score in Table 5 is small and not statistically significant. The failure of pretest to contribute significantly to predicting posttest score can be attributed to the low variability across classrooms in pretest scores. The standard deviation is 1.9, and all but six scores fall between 19 and 24.9. (The variability of pretest scores across individual students is substantial [see Table 2], and in regressions run at the student level the significance of the coefficient of pretest on posttest is extremely high, $t > 13$.) As expected, urban classes perform better on the posttest than rural classes.

Table 4
Pretest and Posttest Means and Standard Deviations for Three Treatments

	Pretest			Posttest		
Treatment	*n*	*M*	*SD*	*n*	*M*	*SD*
Classroom level						
Control	20	21.8	2.0	20	37.2	5.0
Textbook	20	22.1	1.8	20	40.9	3.5
Radio	47	21.8	2.0	47	52.2	5.4
Individual level						
Control	486	21.9	4.6	362	37.6	9.7
Textbook	486	22.3	4.8	414	41.1	9.4
Radio	1,098	21.8	4.5	923	52.5	10.1

Table 5
Regression of Treatment Variables on Class Mean Posttest Score

Determining variable	Regression results *b*	*t*
Radio	14.88	11.36
Text	3.93	2.52
Urban	1.81	1.67
Pretest	.24	.86
Constant	31.02	
R^2	.67	

Note. n = 86. The Radio and Text regression coefficients refer to the control group; the Control variable was omitted from the regression.

Interaction effects. The presentation of findings in Table 5 fails to allow for the possibility that either the radio or the textbook intervention could work differentially well for different groups of students.

[4] Because of attrition between the pretest and posttest, it is useful to examine the pretest scores of those who did and did not take the posttest. There is a small but statistically significant difference between the mean pretest scores for those who were present for the posttest and those who were not present (22.4 vs. 21.7, t = 3.10). The difference is in the direction expected if less able students drop out, but in this case, it is very small. Similar results are discussed in Searle, Sheehan, Gonzalez, and George (1978). However, there is no difference in pretest scores across treatment groups for those not posttested. The scores are 22.1, 22.0, and 21.5 for control, textbook, and radio groups, respectively. Thus, there is no differential attrition across treatments, as measured by the pretest.

[5] We used a double-log (or Cobb-Douglas) specification of the relation between posttest score and its determining variables to explore for the possibility of a stronger nonlinear than linear relationship between the two main variables; both the R^2 and the estimated effects of the determining variables were little changed.

Table 6
Interaction of Urban and Treatment Variables, With Posttest as the Dependent Variable

			Regression runs			
			Against rural control		Against urban control	
Variable	Treatment	No. of classes	*b*	*t*	*b*	*t*
Location						
Urban	Text	7	5.67	2.33	1.13	.47
Urban	Radio	24	17.93	9.66	13.38	7.29
Urban	Control	10	4.54	2.06	—	—
Rural	Text	13	6.08	2.96	1.54	.75
Rural	Radio	22	16.39	8.81	11.84	6.30
Rural	Control	10	—	—	−4.54	2.06
Other						
Pretest			.22	.78	.22	.78
Constant			30.18		34.72	
R^2			.678		.678	

Note. $n = 86$. The unit of analysis in these regressions is the classroom.

Plausibly, the effects could differ among students with different pretest scores, for students in rural and urban schools, as well as for other possible student characteristics.

To test for a possible interaction between treatment condition and students' competence at entrance, as measured by the pretest, we ran a series of regressions (with individual students as the unit of analysis rather than classrooms) that included interaction terms between text and pretest and radio and pretest.[6] All estimated interaction coefficients were very small and did not approach statistical significance. We concluded that the effectiveness of both interventions is independent of the initial competence of the student.

Differences in effectiveness between rural and urban schools did emerge. Since both the attribute (rural–urban) and the treatment (control–text–radio) variables are discrete, we formed six (0, 1) indicator variables to denote the group each classroom is in. For example, if a classroom is in a rural school that received the textbook treatment, the indicator variable for that class denoting rural and text would take on the value of 1, and the other five indicators would take on the value of 0. These indicator variables are then entered, along with the average pretest score for the class, into regressions predicting posttest score. Table 6 shows the results. (The two sets of results columns in Table 6 contain the same information, with different bases formed by deleting different indicator variables from the regression.)

Table 6 shows, for example, that in comparison with the rural control group, the urban control group is expected to average 4.54 additional items correct on the posttest, or the rural text group would average an additional 6.08 items. This indicates that the availability of mathematics textbooks more than compensates for the disadvantage of being in a rural school. More than this, Table 6 indicates that both interventions *reduce* the difference in performance between urban and rural groups. Whereas the difference between urban and rural control groups is 4.54 items correct, the difference between urban and rural radio groups is only about 1.5 items, and the urban text group actually performed slightly less well than did the rural text group. Providing instructional media, whether print or audio, seems to reduce or even eliminate urban–rural differences in school quality.

Degree of implementation of treatment. As indicated earlier, the implementation of the textbook treatment was not monitored during 1978. Early in 1979 all schools in the

[6] See Delaney and Maxwell (1980) for a discussion and justification of examining Attribute × Treatment interactions through multiple regression analysis including discrete covariates (as our urban variable is in this case).

textbook group were visited and information about textbook usage was collected from all but one of them. Eleven teachers reported using the texts all year, and 4 reported 6–8 months usage. In the four remaining classes, the texts were used little, if at all.

One would expect that the degree to which teachers claim to use the texts would be reflected in posttest scores. This is not the case. The mean posttest score for the 11 classes where teachers said they used the texts faithfully is 41.3, compared with 40.6 for the remaining classes. However, from another point of view this finding is not surprising. The score difference between control and textbook groups is only 3.7 points (see Table 4). With small numbers of classrooms it is difficult to detect differential effects within this relatively limited range.[7]

Discussion

The experiment reported here took place during the 1978 school year in the first grades of 88 schools.[8] Students in 48 of these schools received radio lessons; students in 20 of the schools received individual workbooks; and students in the remaining 20 schools served as controls. Assignment of schools to treatment and control conditions was random. All students were pre- and posttested, and the resulting data were analyzed at the classroom level, using linear regression methods to predict the posttest score from indicator variables for treatment condition and location (rural–urban), with pretest controlled.

Both the textbook and the radio treatments had significant positive effects on achievement; availability of textbooks increased student posttest scores by about 3.5 items correct, and availability of radio lessons increased the scores by about 14.9 items correct. These differences are substantial: The standard deviation of the classroom posttest means is 8.3, and the standard deviation of the individual posttest scores is 11.8.

Both interventions, but particularly radio, have the capacity to eliminate much of the gap in achievement across classrooms and a noticeable amount of the more substantial gaps among individuals. Moreover, an analysis of the extent to which treatment efficacy differed between rural and urban schools showed both radio and textbooks to be relatively more effective in rural schools; both interventions, even if uniformly applied, would close a substantial fraction of the existing differences in quality between rural and urban schools. It should be borne in mind that the achievement differences we have reported resulted from interventions

---

[7] Other project results bear on the relationship between intensity of treatment and outcome. The results reported here for first-grade radio and control classes in 1978, the year of the revolution, replicate findings for 1976, a more normal year. All the communities covered by the project were affected by revolutionary activities, and battles were fought in several of them, resulting in school closures and other disruptions of the educational process. Yet regression results on data pooled from both years showed that revolution caused less than a 2-point decrement in posttest scores, and the effect was the same for radio and control classes (Searle & Galda, 1980, pp. 126–128).

In other (unpublished) work we investigated several measures of radio usage and failed to find one that provides more predictive information than the dichotomous radio variable. That is, any level of exposure to the radio-based instructional program results in increased posttest score and variability in outcome associated with different levels of exposure is too small to be detected.

These results suggest that reducing instructional time during the school year does not substantially diminish achievement. (The pattern of missed lessons may, however, be quite important. That is, missing every other day has a smaller impact than missing the first or second half of the year.) These results are inconsistent with the emerging learning-time literature (Wiley, 1976; Wiley & Harnischfeger, 1974; Berliner, Note 5) and with our own previous analysis of the effect of attendance on learning in Nicaragua (Jamison, 1978, pp. 209–210). However, Cooley and Leinhardt (1980) reported similar results in a careful study of the impact of instructional programs on first- and third-grade reading and mathematics achievement. They examined a large number of process variables (more than 75) and found that for first grade, the best predictors of mathematics achievement gain were pretest and "opportunity to learn," which was defined operationally as the inclusion in the material taught by the teacher of the topics appearing on the posttest. The amount of time scheduled for a subject bore little relation either to gain or to raw achievement.

[8] A parallel experiment was run at the fourth-grade level. However, differences between treatment groups and control were small and are difficult to interpret because the control classes were tested in 1977, the year prior to the experiment. Some evidence exists that fourth-grade classes were more substantially disrupted by the civil disorder than other grades. The fourth-grade results are discussed in Searle and Galda (1980) and are not reported here.

lasting for 1 school year; ultimately it will be desirable for research to examine consequences of radio or increased textbook availability over the entire elementary school experience of a child.

Why did the radio group improve at a rate significantly greater than that of the textbook group? It seems likely to us that at least part of the difference can be attributed to the more inconsistent application of the textbook treatment (in the hands of teachers with relatively low levels of education) in comparison with the uniformly administered, high-quality radio lessons that embodied learning principles such as active responding, immediate feedback about answers, and careful timing and systematic review of past lessons. The tradeoffs implied by this situation are interesting: The use of radio as a delivery medium makes possible standardized delivery of lessons, but the pedagogical value of the lessons can be no better than the design and content of what is transmitted. On the other hand, teachers can use textbooks poorly or well and have the potential for making highly effective use of even poorly designed learning materials.

As we have noted, textbooks are far from universal in low-income countries, and the decision to increase their availability has cost and logistical implications that can be substantial for financially pressed ministries of education.[9] Thus, research into the extent of cognitive gains that could be expected from increased investment in textbooks has immediate policy relevance for low-income countries.

Our principal purpose in this article has been to examine experimentally whether provision of textbooks can result in a pedagogically significant improvement in school quality in an environment where an adequate supply of texts is lacking. We conclude that it can. We further conclude that instruction by radio, if the lessons are properly designed, can do even more for in-school quality improvement than can the provision of textbooks. But the provision of textbooks is a more straightforward operation, both logistically and in terms of the demands made on skilled project management; for this reason, in many circumstances expanding the supply of textbooks will be a safer and more sure way of obtaining the desired results. However, the question remains whether either radio or textbooks contain sufficient potential for closing the very substantial achievement gap that exists between the schools of high- and low-income societies.

---

[9] The potential of textbooks for improving school quality has led such agencies as the United Nations Educational, Scientific and Cultural Organization and the World Bank to examine carefully the production, logistical, and cost implications of national textbook programs; see Bolibaugh et al. (Note 6) and Neumann (1980).

Reference Notes

1. Heyneman, S. P. *Primary education in Bolivia: What's wrong?* Unpublished manuscript, The World Bank, Education Department, 1979.
2. Schiefelbein, E., & Farrell, J. P. *Education and occupational attainment in Chile: The effects of educational quality, attainment, and achievement.* Unpublished manuscript, The World Bank, Education Department, 1980.
3. Husen, T., Saha, L. J., & Noonan, R. *Teacher training and student achievement in less well developed countries* (Staff Working Paper No. 310). Washington, D.C.: The World Bank, 1978.
4. Cronbach, L. J., Deken, J. E., & Webb, N. *Research on classrooms and schools: Formulation of questions, design, and analysis* (Occasional Papers of the Stanford Evaluation Consortium). Palo Alto, Calif.: Stanford University, 1976.
5. Berliner, D. C. *The beginning teacher evaluation study: Overview and selected findings, 1974–75.* San Francisco: Far West Regional Laboratory for Educational Research and Development, 1975.
6. Bolibaugh, J., et al. *Guidelines for the preparation, production, and distribution of textbooks* (Educational Financing Occasional Paper No. 5). Paris: United Nations Educational, Scientific and Cultural Organization, 1980.
7. Fuller, W. P., & Chantavanish, A. A study of primary schooling in Thailand: The final report. In, *Factors affecting scholastic achievement of the primary school pupils.* Bangkok, Thailand: Office of the National Education Commission, 1977.
8. Lynch, P. D. *A national textbook program: Promise and frustration.* Paper presented to the annual meeting of the American Educational Research Association, Chicago, April 1974.
9. Oficina de Planeamiento y Organizacion. *The relations of social, economic, and cultural factors and national achievement exam results.* El Salvador, San Salvador: Ministry of Education, 1977.
10. Simmons, J., & Askoy, S. *Schooling for development? Students and workers in Tunisia.* Unpublished manuscript, Harvard University, Department of Economics, 1972.

11. Wolff, L. *Why children fail first grade in Rio Grande do Sol: Implications for policy and research.* Rio de Janiero, Brazil: United States Agency for International Development, 1970.

References

Avalos, B., & Haddad, W. *A review of teacher effectiveness research in Africa, India, Latin America, Middle East, Malaysia, Philippines, and Thailand: Synthesis of results.* Ottawa, Ontario, Canada: International Development Research Centre, 1981.

Beebout, H. S. The production surface for academic achievement: An economic study of Malaysian secondary education (Doctoral dissertation, University of Wisconsin—Madison, 1972). *Dissertation Abstracts International*, 1973, *33*, 3099A–3100A. (University Microfilms No. 72-27,310)

Coleman, J. S., et al. *Equality of educational opportunity.* Washington, D.C.: U.S. Government Printing Office, 1966.

Comber, L. C., & Keeves, J. P. *Science education in nineteen countries.* New York: Wiley, 1973.

Cooley, W., & Leinhardt, G. The instructional dimensions study. *Educational Evaluation and Policy Analysis*, 1980, 2, 7–25.

Delaney, H. D., & Maxwell, S. E. The use of analysis of covariance in tests of attribute-by-treatment interactions. *Journal of Educational Statistics*, 1980, *5*, 191–207.

Eicholz, R. E., et al. *Matematica para la educacion primaria* (Libro 1). Reading, Mass.: Addison-Wesley, 1968.

Farrell, J. P., & Schiefelbein, E. Expanding the scope of educational planning: The experience of Chile. *Interchange*, 1974, *5*, 18–30.

Friend, J., Searle, B., & Suppes, P. (Eds.). *Radio mathematics in Nicaragua.* Palo Alto, Calif.: Stanford University, Institute for Mathematical Studies in the Social Sciences, 1980.

Fry, G. W. Education and success: A case study of the Thai public service. *Comparative Education Review*, 1980, *24*, 21–34.

Haron, I. Social class and educational achievement in a plural society: Peninsular Malaysia (Doctoral dissertation, University of Chicago, 1977). *Dissertation Abstracts International*, 1978, *38*, 7262A–7263A.

Heyneman, S. P. *The evaluation of human capital in Malawi* (Staff Working Paper No. 420). Washington, D.C.: The World Bank, 1980.

Heyneman, S. P., & Currie, J. K. *Schooling academic performance and occupational attainment in a non-industrialized society.* Washington, D.C.: University Press of America, 1979.

Heyneman, S. P., Farrell, J. P., & Sepulveda-Stuarto, M. *Textbooks and achievement: What we know* (Staff Working Paper No. 298). Washington, D.C.: The World Bank, 1978.

Heyneman, S. P., & Jamison, D. T. Student learning in Uganda: Textbook availability and other factors. *Comparative Education Review*, 1980, *24*, 206–220.

Heyneman, S., & Loxley, W. Influences on academic achievement across high and low income countries: A re-analysis of IEA data. *Sociology of Education*, in press.

Heyneman, S. P., & Montenegro, X. P. Social environmental factors which influence academic achievement in the Philippines. In D. T. Jamison & G. Feliciano (Eds.), *Improving school quality in the Philippines: Evaluation, research, and educational policy.* Manila: University of the Philippines Press, in press.

Inkeles, A. Review of the International Evaluation of Educational Achievement. *Proceedings of the National Academy of Education*, 1977, *4*, 139–200.

Jamison, D. T. Radio education and student repetition in Nicaragua. In P. Suppes, B. Searle, & J. Friend (Eds.), *The Radio Mathematics Project: Nicaragua, 1976–1977.* Palo Alto, Calif.: Stanford University, Institute for Mathematical Studies in the Social Sciences, 1978.

Jamison, D. T., & Montenegro, X. P. Evaluation of the Philippines textbook project: Multivariate analysis of data from grades 1 and 2. In D. T. Jamison & G. Feliciano (Eds.), *Improving school quality in the Philippines: Evaluation, research, and educational policy.* Manila: University of the Philippines Press, in press.

Jencks, C., et al. *Inequality: A reassessment of the effect of family and schooling in America.* New York: Basic Books, 1972.

Jencks, C., et al. *Who gets ahead? The determinants of economic success in America.* New York: Basic Books, 1979.

Meyer, J. W., & Hannan, M. T. (Eds.). *National development and the world system: Educational, economic, and political change, 1950–1970.* Chicago: University of Chicago Press, 1979.

Moss, M. H. *Tests of basic experiences.* Monterey, Calif.: CTB/McGraw-Hill, 1970.

Neumann, P. H. *Publishing for schools: Textbooks and the less developed countries* (Staff Working Paper No. 398). Washington, D.C.: The World Bank, 1980.

Searle, B., Friend, J., & Suppes, P. *The Radio Mathematics Project: Nicaragua, 1974–1975.* Palo Alto, Calif.: Stanford University, Institute for Mathematical Studies in the Social Sciences, 1976.

Searle, B., & Galda, K. Measurement of the effect of radio mathematics lessons on student achievement. In J. Friend, B. Searle, & P. Suppes (Eds.), *Radio mathematics in Nicaragua.* Palo Alto, Calif.: Stanford University, Institute for Mathematical Studies in the Social Sciences, 1980.

Searle, B., Sheehan, J., Gonzalez, J., & George, E. Patterns of promotion and wastage for Nicaraguan first-grade students. In P. Suppes, B. Searle, & J. Friend (Eds.), *The Radio Mathematics Project: Nicaragua, 1975–1976.* Palo Alto, Calif.: Stanford University, Institute for Mathematical Studies in the Social Sciences, 1978.

Shoemaker, D. A note on allocating items to subtests in multiple matrix sampling and approximating standard errors of estimate with the jackknife. *Journal of Educational Measurement*, 1973, *10*, 211–219.

Smart, M. The Densu Times: Self-made literacy. *Development Communication Report*, 1978, *21*, 1–4.

Suppes, P., Searle, B., & Friend, J. *The Radio Mathematics Project: Nicaragua, 1976–1977.* Palo Alto, Calif.: Stanford University, Institute for Mathematical Studies in the Social Sciences, 1978.

Thorndike, R. L. *Reading comprehension in 15 countries.* New York: Halsted Press, 1973.

Wiley, D. E. Another hour, another day: Quantity of schooling, a potent path for policy. In W. J. Sewel, R. M. Hauser, & D. L. Featherman (Eds.), *Schooling and achievement in American society.* New York: Academic Press, 1976.

Wiley, D. E., & Harnischfeger, A. Explosion of a myth: Quantity of schooling and exposure to instruction, major educational vehicles. *Educational Researcher*, 1974, *3*, 7–12.

The World Bank. *World development report, 1979.* Washington, D.C.: Author, 1979.

The World Bank. *Education sector policy paper.* Washington, D.C.: Author, 1980. (a)

The World Bank. *World development report, 1980.* Washington, D.C.: Author, 1980. (b)

Received October 13, 1980 ■