Educational Leadership / RESEARCH SUPPLEMENT

EDITED BY ASCD RESEARCH COUNCIL • FREDERICK A. RODGERS, CHAIRMAN

Volume 4 **February 1971** Number 4

CONTENTS

Achievement and Retention as Variables in Comparing Programmed and Conventional Instructions in Geography

***P. O. OKUNROTIFA* ***

THE present universal demand for educational reform has often led to criticisms of the educational profession. Traditional approaches to learning and instruction, which have long been immune to criticisms and secure in their acceptance by the profession, have now been thrown open to question. As one of the most promising innovations in education today, programmed instruction may make a major contribution to the educational problems of developing countries.

However useful programmed instruction may appear to be, the fact still remains that pupils also learn from conventional or traditional methods of instruction. It will therefore be worthwhile, especially in developing countries where financial resources are limited, to ask not only whether programmed instruction is superior or inferior in effectiveness to conventional instruction methods, but also to examine the advantages, or otherwise, of the new instructional device over conventional methods. Some empirical evidence would certainly be needed in this area because, if the effects of programmed instruction are short-lived, it would not seem logical to substitute programmed instruction for other methods.

It is the primary objective of this study

* *P. O. Okunrotifa, Lecturer in Curriculum Development, Faculty of Education, Ahmadu Bello University, Zaria, Nigeria*

to compare, on the bases of achievement and retention, the effectiveness of programmed and conventional instruction methods in the teaching of geography. The study will, therefore, provide a basis for the acceptance or rejection of the following null hypotheses:

1. The use of programmed instructional materials will not result in a significantly better performance on the criterion test than will the conventional instruction method.

2. There will be no significant difference in retention between the experimental and control groups in the study.

The hypotheses were tested at the .05 level of significance.

Method

The Subjects. The subjects for this investigation consisted of three streams of a second form (eighth grade) at Federal Government College, Sokoto, Nigeria. (In Nigeria, "college" includes any post-primary, teacher-training, or higher institution of learning which may or may not award degrees. The Federal Government College, Sokoto, is a high school.) Eighteen of the subjects were girls and 64 were boys. All of them had completed at least six years of elementary school education before gaining admission to the high school. Their mean chronological age was 14 years. As the institution was a secondary grammar school serving the whole of Nigeria, subjects represented not only the present 12 states in the country but also a wide range of socioeconomic levels. The school subjects taken by the pupils included English language and literature, mathematics, and geography.

The second form was selected simply because it was one of the junior classes in our grammar schools where most of the basic concepts in geography were taught.

Development of Criterion Measure. The selected learning material in geography for this study was based on elementary contour map reading. The overall objective of the topic was that learners should understand the concept of contour as a technique for illustrating elevation. The specific levels of understanding which guided the development of test items included recall, explanation, inference, application, and interpretation.

The test consisted of 60 items based on 10 simple diagrams and one large fictitious contour map. All items were constructed by the author. They were multiple choice items with four options for each item. Each item had only one correct answer. The expected perfect score for the test was 60, but the usual correction for guessing was made for a score less than 60. In order not to favour the experimental group, none of the criterion questions were duplicates of instructional items, though some of the same words were used as correct responses. The following are examples of test items:

Recall (5 items). Q. 29: Which of the following gives the correct definition of contours? (a) Lines drawn on a map to divide it into squares; (b) Lines drawn on a map to join places with equal temperature; (c) The difference in height between lines on a map next to each other; (d) Lines drawn on a map joining points which are at the same height above, or depth below, sea level.

Explanation (11 items). *Q. 54: Which of the following reasons best explains why the road stops at Joz and does not continue to Riz? (a) Riz is surrounded by steep slopes; (b) The land between Joz and Riz is too steep for roads; (c) There is a thick forest between Joz and Riz; (d) There is no market at Riz.

Inference (10 items). *Q. 57: Which of the following activities is *not* likely to be practised in the map area? (a) Farming; (b) Hill climbing; (c) Fishing; (d) Coal mining.

Application (18 items). *Q. 52: Which of the following places cannot be seen from Joz? (a) Sallah; (b) Femi; (c) Garuba; (d) Goma.

Interpretation (16 items). *Q. 47: Which of the following places will be most suitable for growing sugarcane? (a) Femi; (b) Abdu; (c) Joz; (d) Riz.

* Test items marked by an asterisk were based on a fictitious contour map.

"Face" validity of the test items was checked by two geography teachers. Following their approval, the test was given a trial

run on 60 fourth formers in a secondary grammar school in Zaria.

Item Analysis of Test: Analysis was made of the responses to each item in order to determine the difficulty and discriminatory power of the item used in the trial. A sampling, consisting of the highest 27 percent and the lowest 27 percent, was used for the item analysis. A count was then taken for each item of the number of pupils in the high-achieving group and in the low-achieving group who chose each alternative.

The percentage of the high-achieving pupils who got the item correct and the percentage of the low-achieving pupils who got the item correct were ascertained. The indices of item difficulty and discrimination were then read from an item analysis table (1). If no more than 80 percent or less than 20 percent of the total group answered an item correctly, it was considered to be within an acceptable difficulty range. An item was considered to be highly discriminating if the index of discrimination was .50 or greater, and poor if the index of discrimination was .30 or less. Items considered to be poor were revised.

Reliability of the Test: By means of Spearman-Brown formula, the reliability of the test was estimated, and the coefficient of correlation between the scores on the odd and even items was .87 and the standard error of measurement was .49.

Development of Instructional Materials: The Programme. The topic of elementary contour map reading for this study was selected for two main reasons:

1. Contour map reading has been one of the areas in geography which many new and inexperienced Nigerian teachers often find difficult to teach. It is, therefore, not surprising that map reading had been particularly poor and apparently widely disliked in Nigerian schools. The situation, therefore, demands that more attention be given beyond the conventional approach to this aspect of geography.

2. Subjects had not received any instruction on the topic, and so there was a more uniform entering behaviour against which to assess achievement and retention.

The programme was constructed by the author, and its construction was guided by the overall objective and the specific levels of understanding stated above. The following areas of learning were considered relevant to the selected topic for instruction:

1. Appreciation of compass direction

2. Ability to translate some common symbols used on maps

3. Picture study

4. Understanding of horizontal scale

5. Recognition, in the field, of simple landforms such as hills, spurs, plateaus, valleys, ridges, and mountains

6. Shading method as a technique of representing elevation

7. Representation of simple landforms (see 5 above) by means of contours

8. Determination of vertical intervals and types of slopes—even slope, steep slope, concave and convex slopes—on contour maps.

The areas of learning above were classified into two: (a) prerequisite learning, that is, learning which the author expected learners to have mastered prior to taking instruction, and (b) learning to be programmed or taught. Learning areas under 1 to 6 above were considered to be prerequisite, while those under 7 and 8 were regarded, according to information collected from schools, as new learning areas to be programmed or taught. However, as "the teacher must begin where the student is," learning area under 6 was included under those to be programmed or taught.

The following characteristics of Nigerian pupils were given consideration in the construction of the programme:

1. English is, at best, a second tongue to Nigerian pupils. Therefore, the language used in the programme was kept as simple as possible, ensuring that the vocabulary used could be understood by any second former in Nigeria.

2. In many Nigerian schools, low reading speeds are the rule. For this reason, un-

duly long frames were avoided in the programme. Short frames were used as far as the material would allow without depriving learners of motivation.

3. Nigerian second formers can learn any subject matter if it is made simple and familiar to them. Bearing this in mind, an adequate number of frames was provided to give learners practice and to ensure gradual transition from one idea to another. Nigerian names and maps were used and ideas in the frames were related as much as possible to everyday life in Nigeria.

The resulting programme was a linear constructed response type printed on thirty 13 by 8 inch sheets with answers on the left-hand side of the next page. The programme consisted of 47 diagrams and 111 frames calling for 299 responses. The first five frames were written to teach subjects how to use programmed material. The material contained in the familiarization programme was totally unrelated to the main programme. Checks were made to ensure that each frame called for the appropriate response and that every answer was the correct response called for.

Programme Validation. The first validation group consisted of seven second formers at St. Enda's College, Zaria. They were of average ability and their average age was 14 years. They were first given the pretest, and then the programme, followed by the post-test. The mean scores on the pretest and post-test were 23 and 34 respectively. The notes taken during the working session on the programme helped the author in the revision, addition, and omission of many frames.

The second validation group consisted of six second formers at the Nigerian Military School, Zaria. The trial was conducted under similar environmental conditions to the first validation group. The subjects' "feedback" and error frequency provided bases for further programme revision. The mean pretest and post-test scores were 20 and 42 respectively.

The third validation group consisted of six boys at St. Paul's College, Zaria. The mean pretest and post-test scores were 24 and 48 respectively. Only minor revisions were made at this stage because the author was satisfied with the pupils' favourable comments and the learning gain achieved. The final programme was printed on twenty-eight 13 by 8 inch sheets; the sheets were then bound into a booklet. The validation of the programme resulted in a 102-frame linear programme consisting of 50 diagrams. Two hundred ninety-two responses were required.

Conventional Instruction. Two months before the experiment was administered, the author obtained the assistance of an undergraduate student at the Ahmadu Bello University, Zaria, Nigeria. The student indicated considerable willingness to participate in the study. He was a second-year student reading geography and education. His work in both subjects was considered to be well above average. The student had attended a teachers college and had taught for three years before gaining admission to the university. At the university, he had once been exposed to a four-week school practice in one of the local secondary grammar schools.

The author had detailed discussions with the student on the synopsis of this study. The student was given a copy of the synopsis including a list of the overall and specific objectives for the selected topic and a list of the prerequisite learning areas and those on which instruction was needed. He agreed to develop detailed lesson notes from the materials provided.

Having completed the lesson notes, the student was given opportunity to present his lessons to 15 Nigerian second formers during weekly micro-teaching sessions organized by the University Department of Education. The student had four meetings with the pupils, and at the end of each meeting his classmates made many valuable comments which were used to improve his lesson plans.

Only 12 of the 15 pupils attended all the four micro-teaching sessions. It was not possible to give the pupils any pretest; they were, however, given the post-test and their median score was 65 percent. The final materials prepared by the student-teacher for his les-

sons included four detailed lesson notes, a wall map of Nigeria, and duplicated contour maps.

Administration of the Experiment. The author was assisted during the first three days of the experiment by four undergraduate students reading history, geography, and education at the University in Zaria. Eighty-five second formers of the Federal Government College, Sokoto, were available for the experiment. They were assured during the first meeting that the ulterior motive of the experiment was to teach them a new topic in geography rather than to test them.

Having completed a short questionnaire about their age, sex, and home, subjects were given the pretest. Forty-one pairs were made by the author from the 85 subjects who took the pretest; three pupils were excluded from the study. The pupils were paired on the bases of their pretest scores, age, sex, and socioeconomic level. One pupil in each pair was randomly assigned to an experimental group and the other to a control group. The experimental group used the programmed materials and the control group was taught by conventional methods.

On the second day, subjects left for their respective groups.

The Experimental Group. The first five frames of the programme were used to explain to subjects how to learn with programmed materials and how to use response sheets. The group had two study sessions—morning and evening—each of 40 minutes. Subjects were told to work at their own pace and that everyone would be given enough time to finish. At the end of each session programmed materials were collected.

Subjects continued with their programmes on the third day (Saturday). There was no time limit for the subjects that day, but a record of the starting and finishing times for each of them was kept. During the programme administration, supervisors did not instruct but answered subjects' questions on words and numbers which were considered not clear.

Group	N	Mean	S.D.	t	p
Experimental	41	33.77	8.07	.41	ns
Control	41	32.53	9.50		

Table 1. Comparison of the Pretest Scores of the Treatment Groups

The Control Group. This group also had two meetings—morning and evening—each of 40 minutes during the second day. Under the guidance of the student-teacher, subjects spent the first period of their time on a slope in the school compound.

One of the major activities of the subjects outside the classroom was the alignment of the subjects themselves along some contours, thus forming some "human contours." They were questioned as to heights in between the contours and were required to record on paper the shape of the lines they themselves represented. The record of the contours made by the subjects themselves, numbered in feet above sea level, was next linked to the contours shown on a large-scale map of the school area. The activities ended in the classroom, where the experiences collected outside were translated into blackboard drawings.

	Experimental Group	Control Group
Shortest Time	98 minutes	
Mean Time	129 minutes	140 minutes
Longest Time	150 minutes	

Table 2. Comparison of Learning Times Spent by Treatment Groups

The group had a long period of 60 minutes on the third day (Saturday). All work was done in class. Instructional techniques included lecture, questioning, demonstration, and discussion. Subjects were given work to do in class and individual help was given to some of the subjects.

Post-Test. On the third day, at the end of instruction, each subject in both treatment groups was given 15 minutes to review the material learned. After the brief review, the criterion test was administered. The control group had a uniform starting time because lessons ended for the group at the same time, while subjects in the experimental group

started at varying times when individuals completed their programmes.

Subjects were allowed to spend as much time as they needed to complete the test. They were also warned at the beginning of the test that they would be penalized for guessing.

Retests. The post-test was given in the middle of July 1969. As the school was due to have a terminal examination before it broke up during the last week in August, it was not possible to administer retention tests later than three weeks after the completion of instruction. It was therefore decided to retest at intervals of one week, two weeks, and three weeks.

To avoid practice effect of repeated testing with the same test, each treatment group was divided into three for retest. Thus, for both treatment groups, 15 were retested after one week, 14 after two weeks, and 12 after three weeks. The author was, however, aware of the fact that the resulting values of test population were small, but it was considered important to control the effect of repeated testing. Subjects in each retest group were matched on post-test scores, age, and sex.

Method of Scoring. All answers to tests were scored by the author. Correction for guessing was made by deducting one-third of the number of incorrect answers from the number of correct answers.

Results and Discussion

Differences between the pretest and post-test scores for individual subjects were obtained. The difference score for each subject represented the amount of gain in learning, which, in turn, could be regarded as the achievement made by the subject from instruction. Achievement, however, is a variable closely related to retention. Table 3 indicates a comparison of the achievement scores (difference scores) in both treatment groups. A t-test for correlated means was computed and none of the differences reached statistical significance at the .05 level. The study, therefore, indicated that there was no significant difference between the achievement of the experimental and control groups. The first null hypothesis of this study was thus accepted. However, a comparison of the means of the pretest (Table 1) and post-test (Table 5) scores, in each treatment group, indicates that both methods of instruction contributed significantly to the achievement of subjects. It should also be pointed out that there was a mean difference of 1.30 points in favour of the experimental group.

Group	*N*	*Mean*	*S.D.*	*t*	*p*
Experimental	41	10.67	5.77	.83	ns
Control	41	9.37	5.01		

Table 3. Comparison of the Means and Standard Deviation of Achievement Scores

The retention scores for subjects were derived from the total number of items answered correctly on *both* the post-test and retest.

Table 4 indicates that none of the differences reached statistical significance and, therefore, the second null hypothesis of this study was again not rejected.

Table 5 shows the differences in the post-test scores; t-tests for correlated means were computed for the comparison of:

(a) The means of post-test and retest scores in each treatment group. The obtained t-values of 7.98 (experimental group) and 10.30 (control group) were significant at the .05 level.

(b) The means of pretest and retest scores in each treatment group. The obtained t-values of 2.87 (experimental group) and 2.64 (control group) were also significant at the .05 level.

From the results (a) above, a comparison of the means of post-test and retest scores in each treatment group indicates that, for both groups, there were significant losses in retention during the three weeks, but the loss in retention (the amount of forgetting) was higher in the control group than in the other group. However, the significance of the differences between the means of the pretest and retest scores, in each treatment group (see b above), shows that, despite significant losses in retention, a substantial amount of the material taught was retained until the end of the three weeks.

	Experimental					Control				
	N	*M*	*S.D.*	*t*	*p*	*N*	*M*	*S.D.*	*t*	*p*
First Week Retest	15	44.11	7.66	1.59	ns	15	37.78	9.64		
Second Week Retest	14	38.55	9.51	.37	ns	14	36.55	6.21		
Third Week Retest	12	39.40	7.82	.91	ns	12	35.50	9.50		
Totals (3 groups)	41	40.50	8.79	1.70	ns	41	36.57	8.56		

Table 4. Comparison of Retention Scores for Treatment Groups

The major result of this study was obtained by analyses which indicated that the two treatments did not differ significantly in affecting the pupils' achievement and retention. The result is important because the initial favourable reaction to programmed instruction had led, during the past decade, to some claims including statements that programmed instruction was superior to conventional instruction and that pupils were able to learn several times faster than the normal rate through the use of programmed material (2). The findings of this study have not given any support to these claims.

Programmed instruction might not have proved significantly superior to conventional instruction methods in this study because of three factors, each of which has some educational implications.

First, programmed instruction is essentially based on reinforcement theory. The instructional device certainly allows the use of other motivational devices such as self-pacing, maintaining interest, and confidence, all of which could prove helpful to learners. Yet there is definitely a limit to the amount of motivational devices which can be included in programmes, especially in Nigeria, where low reading speeds are not uncommon. The teacher used a variety of incentives for the control group in this study.

A teacher should consider the incentives he is offering his pupils, remembering that specific incentives will not serve equally well in motivating all pupils, as it is often assumed in programmed instruction. Pupils who vary individually with respect to intelligence, physical makeup, and emotional constitution will also vary with respect to potential for motivation. It is important that effective means of instruction should involve different incentives in a variety of ways and will elicit different results.

Apart from the motivational devices included in the programmed materials in this study, there were 50 diagrams in the programme to lend meaning to the material. Despite these learning opportunities, the experimental group might not have been able to achieve and retain significantly more than the control group because of the variety of instructional techniques to which the control group was exposed. The teacher's techniques in the control group ranged from the use of lecture, field study, questions, and demonstrations to the use of assignments and maps. Subjects were given opportunity to practice, and some of them received individual attention.

In contrast to programmed instruction, questions to which original or different answers were equally correct were answered by the teacher. Furthermore, while the experimental group had the benefit of limited visual aids, the control group had the benefit of a combination of auditory aid, such as the teacher himself, with visual materials such as maps, blackboard, and others which are all capable of increasing the depth of meaning of important concepts essential for pupils' achievement and retention.

A third factor which, perhaps, influenced the result of this study was the role of an interesting personality in the control group. Although, in this scientific age, the concept of teaching as a personality function is being justifiably assailed from all sides, no one would deny that an interesting personality is an effective lubricant in the instructional process. The student-teacher used in this study possessed certain personality characteristics. He was helpful, approachable,

	Experimental					Control				
	N	*M*	*S.D.*	*t*	*p*	*N*	*M*	*S.D.*	*t*	*p*
Group 1	15	47.22	6.17	1.44	ns	15	42.44	9.01		
Group 2	14	44.55	7.49	.83	ns	14	42.36	5.28		
Group 3	12	45.70	5.84	.97	ns	12	42.50	6.80		
Totals (3 groups)	41	45.73	6.71	1.92	ns	41	42.43	7.09		

Table 5. Differences in the Post-Test Scores Between Treatment Groups

friendly, and intelligent. He had a good command of his subject, planned effectively, and knew how to communicate his subject to pupils. It should be noted that the vital factor in teaching and learning geography, as in teaching and learning every other subject, is the teacher. To promote its effectiveness, the programme used in this study was adequately validated, and it also included motivational devices. It would, therefore, be considered a good programme, but the degree of inspiration which it could give learners might not be comparable with that from a good classroom teacher.

Although this study indicated, contrary to the predictions of its enthusiastic adherents, that programmed instruction was not significantly superior to conventional instruction methods on the bases of achievement and retention, a careful examination reveals that the slight differences in almost all instances are in favour of the experimental group.

Despite the suitability of the teacher for the control group, the size of his class made it extremely difficult for him to give *each* pupil individual attention, which is one of the major attributes of programmed instruction. An important implication of this study for instructional purposes, therefore, is the demonstration of the need to lessen the current importance being attached in our schools to the class group of uniform size taught as an assumed homogeneous unit in terms of subject attainment. Programmed instruction, by its nature, emphasizes the fact that our work as teachers is the education of individual persons.

The fact that, despite the nature of the teacher's personality and the variety of methods and aids used in the control group, that group could neither achieve nor retain more than the experimental group clearly shows the considerable potentials of programmed instruction. The results justify the fact that programmed instruction also includes many essential principles of good teaching such as individual differentiation, graduated sequence, appropriate practice, immediate knowledge of results, and active response.

This study, therefore, implies that some good programmes would very likely be superior, on the bases of achievement and retention, to classroom teaching under the control of incompetent and unimaginative teachers. In other words, programmed materials in geography are likely to serve a useful purpose in some of our schools where there is a shortage of suitable teachers and equipment. Even in schools where suitable teachers and equipment are available, programmed materials can profitably be used without any significant difference in the time devoted to the same amount of learning done by conventional methods.

This study strongly indicates that some good programmes in geography can compare favourably with excellent classroom teaching. This means that programmed instruction can provide a new dimension for geography teaching in schools. The teacher can use the new instructional device either in conjunction with or as an alternative to conventional methods as a means of encouraging a certain degree of overlearning. The extra effort often demanded in overlearning is a wise investment in terms of long range retention. As an alternative to conventional methods, programmed instruction can be used successfully to encourage periodic review of learning

Continued on page 517

Continued from page 508

materials in order to secure maximum retention, especially when the learning materials are to be used in climbing to higher levels of learning.

This study has illustrated one particular contribution which programmed instruction could make to conventional methods. The roots of programmed instruction are deep both in the theory of education and in the psychological learning theory. It would be recalled that the specific objectives and the content outline obtained through programming process were those used by the teacher for the control group to develop conventional lessons. The rigorous analysis and the formulation of a topic's objectives in operational terms often require some hard thinking about the necessary content of teaching and can teach one far more than the preparation of many conventional lesson plans. The process of programming calls for the breaking down of learning materials into units which are meaningful in terms of pupils' ability and background, with stress being placed on generalization and application, which tend to withstand forgetting. Geography teachers involved in the process would, no doubt, increase their understanding for the development of more meaningful lessons, which are important to pupils in terms of achievement and retention.

Summary

The results of this study show that, on the bases of achievement and retention, programmed instruction is comparable with conventional instruction methods. The results obtained imply that programmed instruction, being a practical and a valuable instructional method, can be used either in conjunction with or as an alternative to conventional methods to promote retention and achievement in the classroom. Contrary to speculation that programmed instruction may replace conventional methods in Nigeria, the results of this study imply that the new instructional device is only one of an arsenal of teaching devices at the command of the teacher to help him to promote a more permanent learning in the classroom.

Acknowledgments

The cooperation of the Department of Education, Ahmadu Bello University, Zaria, Nigeria, and of the staff and pupils of the secondary schools in Zaria and the Federal Government College, Sokoto, Nigeria, is gratefully acknowledged.

References

1. Chung-Teh Fan. *Item Analysis Table.* Princeton: Educational Testing Service, 1952.

2. J. L. Hughes and W. J. McNamara. "A Comparative Study of Programmed and Conventional Instruction in Industry." *Journal of Applied Psychology* 45 (4): 225-31; 1961. □