The Mathematics Teacher Shortage, Class Size and Programmed Learning

William Maltby
College of William & Mary - School of Education

Follow this and additional works at: https://scholarworks.wm.edu/etd

Part of the Educational Administration and Supervision Commons, and the Science and Mathematics Education Commons

Recommended Citation
https://dx.doi.org/doi:10.25774/w4-8qfa-fr29

This Thesis is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
THE MATHEMATICS TEACHER SHORTAGE, CLASS SIZE AND PROGRAMMED LEARNING

A Thesis
Presented to
the Faculty of the School of Education
The College of William and Mary

In Partial Fulfillment
of the Requirements for the Degree
Master of Education

by
William Maltby
May 1964
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I THE PROBLEM AND PURPOSE</td>
<td>1</td>
</tr>
<tr>
<td>The Problem</td>
<td>1</td>
</tr>
<tr>
<td>Limitations</td>
<td>1</td>
</tr>
<tr>
<td>How the Study Was Made</td>
<td>2</td>
</tr>
<tr>
<td>II THE SHORTAGE OF MATHEMATICS TEACHERS</td>
<td>3</td>
</tr>
<tr>
<td>III CLASS SIZE</td>
<td>6</td>
</tr>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>Research from the 1930's</td>
<td>7</td>
</tr>
<tr>
<td>Camelback High School and Class Size</td>
<td>10</td>
</tr>
<tr>
<td>Haertter and McGuire Minnesota Experiment</td>
<td>12</td>
</tr>
<tr>
<td>Additional Experiments</td>
<td>14</td>
</tr>
<tr>
<td>Conclusions</td>
<td>16</td>
</tr>
<tr>
<td>IV PROGRAMMED LEARNING AND TEACHING MACHINES</td>
<td>18</td>
</tr>
<tr>
<td>Introduction</td>
<td>18</td>
</tr>
<tr>
<td>History and Development of Automated Teaching Devices</td>
<td>18</td>
</tr>
<tr>
<td>Comparison of Programmed Techniques</td>
<td>20</td>
</tr>
<tr>
<td>Traditional Classroom versus Programmed Techniques</td>
<td>23</td>
</tr>
<tr>
<td>Summary</td>
<td>27</td>
</tr>
<tr>
<td>V THE ROANOKE EXPERIMENT</td>
<td>29</td>
</tr>
<tr>
<td>VI CONCLUSIONS AND IMPLICATIONS</td>
<td>31</td>
</tr>
<tr>
<td>FIGURES AND TABLES</td>
<td>33</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>40</td>
</tr>
</tbody>
</table>
CHAPTER I
THE PROBLEM AND PURPOSE

It was the purpose of this study to evaluate the suitability of programmed learning and the factors of class size for an immediate correction of the deficiency of mathematics teachers in the American High School.

I. THE PROBLEM

The rapid advance of the whole world in the fields of science and technology make it imperative that all students should be adequately instructed in mathematics. Now that the mathematics curriculum is undergoing successful revision, this part of mathematics teaching is prepared for this day and age. However will these new mathematics programmes ever reach all the students. There is, and it will be shown that there is a shortage of fully qualified and available mathematics teachers.

There are two possible solutions, which the writer wishes to investigate, which may lead to an easing of the mathematics teacher shortage. Firstly, to see whether the traditional concept of class size is outmoded. Secondly, there is a possibility that the use of programmed learning may be applied to the classroom situation.

II. LIMITATIONS

The study is almost entirely confined to articles, research journals, books and pamphlets found in the library of the College of William and Mary. Obtaining specific material from outside sources brought little response in the specific areas required. Since this is a study of studies, much of the outside material obtained was irrelevant opinion not based on scientific research.

All the research articles and information refer to mathematics as
the writer, though a mathematics teacher by admission, feels that both in the areas of class size and programmed learning, there are obvious dangers in transferring or drawing conclusions from experiments and results from other subject fields.

This study is concerned with content, method or curriculum problems in mathematics, and it is confined to the one specific problem of the solution of teacher shortage. In addition programmed learning is primarily considered as a teacher or teaching device.

III. HOW THE STUDY WAS MADE

The study was started by consulting the Encyclopedia of Educational Research and the Educational Index and some standard books on programmed learning. These generated a whole range of articles and research reports on which this whole study is based from most educational and psychological journals that are published in the United States.
CHAPTER II

THE SHORTAGE OF MATHEMATICS TEACHERS

Since the whole study is based on the assumption of a shortage of mathematics teachers, it will be demonstrated in this chapter that there is a definite quantitative shortage of teachers, in addition to a more qualitative shortage as demonstrated by an interesting study made in Virginia in 1961-62.

"Teacher Supply and Demand" by R. C. Gibson was a study made to evaluate not only teacher shortage but also all school personnel in Virginia. Gibson made a survey of all the county and city school systems in Virginia via the offices of the school superintendent of each system. The superintendents were each asked to indicate on the basis of their recent experience if the number of properly certified teachers or other personnel were in excess of, about equal to, or short of the required number for each subject and employment area in their schools.

Responses were obtained from all but four school systems in Virginia, though three of those answering indicated lack of sufficient experience to respond intelligently. Thirteen responses represented two school systems each as the superintendents served jointly in adjoining school systems.

As an index of teacher shortage each reply was assigned a value; 2 if there was an excess supply; 1 if there was an adequate supply; 0 if there was a definite shortage. The sum of all values for each response in each teaching and employment field was made and the result divided by the total number of responses. The resulting number is a measure of supply

teachers in Virginia in each employment and subject area, ranging from a value of 2.0 for an abundant supply; to 1.0 for a just adequate supply; to 0.0 for an absolute shortage of teachers in that subject area.

Reproduced in diagram I is a partial copy of the graph given by Gibson. Some areas of employment, for example, elementary, administration, and guidance have been left out so as not to make the graph so large. The form of the graph is essentially the same and the areas left out were classed in the "just adequate" supply.

The most severe shortage indicated by this survey were in the areas of physics, mathematics, and chemistry with index value of .12, .13, and .16 respectively, contrasting sharply with social studies and history at the other end of the scale with values 1.41 and 1.30 respectively. Gibson states that the imbalance in physics and chemistry should be corrected within the next few years but in the area of mathematics the shortage is likely to increase. This is a fairly obvious conclusion as with regards to the science subjects and the grand total of teachers required to correct the imbalance is a great deal smaller than in mathematics.

There is, of course, no quantitative figure given in the survey to indicate the severity of the mathematics teacher shortage, and in addition one must question whether the three categories that Gibson set up were distinctive enough to make his results significant. However, this survey is useful in indicating that if there is a shortage at all in the educational world it is located in mathematics teaching. In addition Gibson points out that the acute shortage is even more obviously found in those counties in Virginia which pay low salaries as compared with many of the city school systems and the northern counties around and near Washington, D. C.

To obtain some quantitative figures of mathematics teacher shortage
one has to turn to the NEA Research Bulletin.² In 1962, there was a 14.3% increase of mathematics teachers, whilst in the whole range of teaching there was an increase of 13.9%. This meant that if anything there was a slight increase or small improvement over the previous year but there is no indication that there was anything but a superficial relieving of the acute shortage. It appears ludicrous that then in 1962, there were 7,700 physical education majors as compared with 7,000 mathematics majors ready to enter the schools. Since only one third of these physical education majors, many without an adequate minor, can be employed in their own subject area, they are often employed in other subjects (and one presumes this is often mathematics).

In 1963, the National Education Association tell us, that the percentage of school staff teaching mathematics was 11.4% of the total. Of all new prospective teachers for 1963, those who are qualified to teach mathematics represent 8.4% of the whole, indicating that the supply does not even keep up with the demand. Some 8,100 mathematics teachers entered the field of education in 1963, but at least 11,200 were required to preserve the status quo. Thus the position over mathematics teacher shortage deteriorated in 1963 rather than improved.

Though the writer has found it impossible to find any exact data for the total overall shortage of mathematics teachers, the Virginia Survey of 1961-1962 and the NEA Research Bulletin figures for 1962 and 1963 indicate a measure of the shortage of mathematics teachers, and in addition the position is not improving—if anything, deteriorating.

CHAPTER III
CLASS SIZE

I. INTRODUCTION

Ask any teacher the question, "How many pupils in a class do you think is best for the teacher and the student?"; invariably the answer lies between the magic numbers twenty to thirty and more often in the low twenties.

The writer sought to find out where this magic number of students per teacher came from. By mere chance the writer was fortunate enough to stumble across, whilst casually reading unrelated topics, on two independent sources which give the answer.

"We traced the origin of the twenty-five-pupils-to-one-teacher myth (said Mr. Coombs) back to the Talmud. In the Talmud it is stated that there shall be one teacher for twenty pupils and if there are more than twenty-five the teacher shall have an aid. If there are more than thirty the class shall be divided into two. That was in the year 400. But not all teachers kept that ratio; Jesus did not, for instance. Why do teachers today stick to it?"3

"Two millennia ago the Rabbi Raba established the rule that twenty-five students are to be enrolled in one class. If there are twenty-five to forty an assistant must be obtained. Above forty, two teachers are to be engaged."4

Apart from disagreeing slightly in their historical positioning by some 500 years, both Coombs and Sir Eric Ashby agree that the pupil-teacher ratio has its origin as command of law rather than based on a more critical approach.

In this chapter pertinent experiments on class size will be examined to see whether research can give guidance over this whole problem, for if


the pupil-teacher ratio is a myth as has been suggested, then the class sizes can quickly be enlarged. This in turn would solve the mathematics teacher shortage. On the other hand if this figure is not a myth it may be better to maintain the present pupil-teacher ratio, even if it means the recruitment of some mediocre teachers, so that the teacher can communicate his mediocrity in an intimate environment.

Before one can examine the details of the experimental evidence of the pupil-teacher ratio, it is necessary to consider the basic terminology used. By a small or traditional class it usually means that the class has some twenty to forty students, though generally the classes are in the twenties. Large classes are designated to be those with more than forty students and may be as large as one hundred or more students.

Out of some 218 references found pertaining to class size, by far the great majority were mere opinion, no better than the utterances of the Rabbi Raba. By the time research reports in other subject fields were rejected: only a few experiments in mathematics resulted. After rejecting some eight reports as they were unaccompanied by any kind of results but only conclusions, the writer was left with some five original sources that have any accuracy and relevance. The five studies are considered here and for each there is an individual criticism. Following the studies there is a summary and general conclusion of the research material, whilst overall conclusions and implications pertaining to the teacher shortage are left till Chapter IV.

II. RESEARCH FROM THE 1930'S

It may appear bad educational policy to return to an experiment that was conducted in 1930, but since it is one of the very few experiments to come through as a worthwhile and accurate study it is certainly justifiable

---

to include this work.

Cunningham wished to consider the outcomes of live parallel large and small classes under particular controlled conditions. Approximately one hundred students were assigned to a B-9 (beginning 9th grade) algebra class. The two classes were assigned in a two-to-one ratio, though at the end of the half-year, when the final results were made up, complete records were only available for fifty-five and twenty-eight pupils respectively. An IQ test, taken by the students when they entered the 8th grade was used to form the two groups and a prognosis test in algebra was administered in order to make an initial comparison of the large and small classes. Quarterly tests were given during the semester and finally both groups were given the Columbia Research Bureau Algebra Test with a local City Semester Test, compiled by the City (Los Angeles) Division of Psychology and Educational Research.

The teaching method was reported to be the same for both classes including the same teacher for both groups. Books, assignments, tests were identical for the large and small classes.

Table I has been constructed from the detailed results published. No figures have been altered or modified, but some detail has been removed in order to clarify the main results. From these results Cunningham suggests that there was no statistical evidence of a significant difference in achievement between the two groups in one semester of algebra.

There are several features of this experiment that need to be examined before the conclusions that were drawn are seriously accepted. The two groups were divided by means of the IQ test administered one year prior to entering the 9th grade. The results of the prognosis test in algebra show, that for mathematics, these two groups were not equal. The author ran a t-test from the results shown in Table I. The resulting t-score was very close to the 5% level of confidence, though just within the restrictions usually allowed.
It is well within the realms of possibility that the large group was significantly better than the small group before the experiment was started. It would therefore not be surprising to find the larger group still superior at the end of the semester.

From the report the following statements appear:

"They were told that the group had been formed for experimental purposes." 6

"In the large group at the beginning of the term stress was laid upon the greater need of individual responsibility in matters of conduct, concentration and regular attendance than perhaps had been previously felt in small groups." 7

"The novelty of working in a large group seemed to be a factor which tended to arouse members to an unusual interest and desire for success in the class." 8

With these admissions, it would therefore appear that the controlled conditions implied, were not as they appear to be. Granted that it was necessary to have the large group well-ordered, but it is now uncertain how much effect this had on the controlled nature of the experiment and thus on the results obtained. Transfer to an everyday classroom situation of this kind of control by appealing to the experimental nature of the class would no longer have the same effect. One can only speculate that the total performance under the normal school conditions would now be different, and the writer suggests that it would be significantly different.

To help the teacher of the large group, a substitute teacher-clerk checked the papers and tests of the large group and recorded the results, thus the work load of the regular teacher was reduced. No longer was the large group teacher working under normal everyday school situation.

Of the three criticisms made, the most serious is the actual division

6 Cunningham, op cit, p. 19.
7 ibid, p. 20.
8 ibid, p. 23.
of the two groups into supposedly two equal ability groups. They were cer­
tainly equal by the IQ test, which is largely verbal, but in aritmetic or
algebraic ability, groups were certainly not equal to a level of absolute
confidence. It would thus appear that the rest of the results must be treated
with caution. The prognosis test (the Orleans Test of Ability in Algebra)
would have been a sounder basis on which the two groups could have been divided.

The other major criticisms imply that the groups were not controlled
and that because of special attention the large group were unduly favoured.
Given artificial incentive the "novelty" effect could well have been operat­
ing. The teacher of the large class was favoured by the special assistant
received. The large group achievement may partially have been caused by
these effects.

III. CAMELBACK HIGH SCHOOL AND CLASS SIZE

Turning from a pre-war study to a more recent study\(^9\) conducted at
Camelback High School, Phoenix, Arizona, in 1960-61, we find another which
dealt with ninth grade students in algebra.

One hundred and twenty students were picked from eighth and ninth grades
of the Differential Aptitude Numerical Test. It was decided to split
the group into two classes, one large of eighty students and one small of
forty students. The students were assigned randomly and then were compared
by administering the Iowa Test of Educational Development. This test showed
that the larger class was slightly but not significantly lower in achievement.
The two classes were taught by the same teacher, who was given an extra hour
of preparation time for the large class. Two twelfth grade assistants were
utilized for clerical work and routine checking and grading. The large class
was held in a choral room, which did not have any apparent disadvantages as

compared to the normal classroom.

After one semester of instruction in February, 1961, the two groups took an alternate form of the Iowa Test, which is normally taken after two semesters' work. The overall results of the pre-test and post-test are shown in Table II. From these results it was concluded that there was no significant difference between the two groups after one semester work of algebra. The standard error of the difference between classes is 1.668 for the pre-test and 1.115 for the achievement test, well within the allowable statistical limits.

This well-conducted study has also two major drawbacks. In the group of one hundred, all but six were at the nintieth percentile or above. This means that the students were of an extremely select group of the school population, certainly atypical. This would indicate that to transfer these results into the everyday school situation would not be justified in view of such a select group. It is not very surprising to find no differences showing up; with such a select group, any type of instruction could possibly lead to the same kind of improvement. However it would be a fair conclusion to say that students of extremely high ability do not seem to have their learning impaired by large classes.

As in Cunningham's experiment, the teacher received assistance with the large class. To quote Anderson, "many of these problems (administrative) were eliminated or relieved by arranging to have two twelfth grade assistants." ¹⁰

"The one hundred twenty students, divided into two classes were to be considered the equivalent of three regular classes. This meant the teacher was to be in the classroom two periods and have one preparation and believe me, one needs the 'extra' preparation period." ¹¹

¹¹ ibid, p. 156.
"I found that doubling the student load more than doubled the time needed for preparation, grading and reporting, and working with the students. I wouldn't recommend a large group for regular mathematics classes."\(^{12}\)

According to Anderson, he does not feel that the large class could be immediately transferred into the normal school situation, but sees large group teaching possible with high ability students provided the teacher has clerical assistance and compensatory time.

IV. HAEFTER AND McGUIRE MINNESOTA EXPERIMENT

Earl Hudelson reported on various class-size experiments conducted prior to 1930.\(^ {13}\) Within the article Hudelson reports an experiment concerning the teaching of plane geometry to large and small classes, undertaken by Haertter and McGuire at the Minnesota University High School in 1926-27. Though this report appears to be from a secondary source it in fact includes all the results published by Haertter and also unpublished material by McGuire.

The experiment was due to run for two years, but Haertter, who started the experiment, left after one year to take up another appointment. The experiment was taken up by Miss McGuire, who had not previously been associated with the experiment. Here then is an experiment taken over by an inexperienced experimenter, which could be likened to a school situation, where increase in class size might be suddenly implemented.

Haertter had a large class of fifty-five students and a small class of twenty, where two pupils in the large class were paired with each member of the small class, according to sex, mental age, average IQ in six standard intelligence tests, average freshman marks in mathematics, and three objective mathematics tests (the Reeve Minimum Essentials Test, Minneapolis Minimum

\(^{12}\) ibid, p. 158.

Essentials Test in Plane Geometry, and Haertter Test in Plane Geometry).

Haertter's small sections met as under normal good teaching situations, whilst the salient feature of the large class was the breaking down of the group into subsections which worked separately. Each unit taught to the large group was mimeographed, as were drill questions, tests and assignments.

In Table III the results of Haertter's experiment are given. It will be seen that according to average IQ, average freshman marks and the pairing tests used, the groups were very well matched. In the achievement test given after one year's instruction the difference, on a test with a maximum of 780 points, between the large and small groups was only ten points in favour of the small group, which is not statistically significant. It would thus appear that the large class did as well in the experimental situation as did the small class.

When Miss McGuire took over the experiment in the fall of 1926, she formed two groups. Her small class was twenty-three compared with a large class of forty-four, in which nineteen in each were closely paired as Haertter had done the previous year.

It is reported that McGuire had immediate difficulty with the large class; she had discipline problems, she tried cooperative development of theorems with little success, and she tried class supervised study, all to no avail. After awhile she organized the large class into homogeneous squads and these were instructed separately. The accelerated squad were allowed to proceed at their own pace, whilst the mediocre and retarded groups were supervised by McGuire herself.

At the end of the year accomplishment for both large and small classes was measured by nine objective tests. The results of McGuire experiment are

\[^{14}\text{Hudelson, op.cit., p. 200.}\]
\[^{15}\text{ibid., p. 200.}\]
\[^{ibid., p. 202.}\]
given essentially in Table IV. It can be seen that not only were the two groups essentially the same statistically before the years' work but at the end of the various standardized and other tests, the level of achievement still was about the same. It would appear that the large class did not suffer through being in a large group.

The experiment started by Haertter was carried out in a controlled scientific manner. The large class was specially treated to mimeographed material and a well-organized routine developed by the teacher. However, when McGuire took over the experiment there was admitted chaos. She admits she spent much time in solving problems that confronted her. The experimental control on the large class was largely lost, additional time was given to the students above that allotted to the small class. In effect McGuire reduced the large class into a number of small groups which she could handle. The high-ability went on at their own rate whilst McGuire taught the other reduced number students. It can therefore be speculated that the large class did as well as the small due to the special attention given to it and the large class was reduced to small groups to which additional time was given.

IV. ADDITIONAL EXPERIMENTS

At this point the amount of good experimental research in mathematics and class size comes virtually to a halt. The following experiments have not the scientific rigour of the previous ones, but are worth including for the information they have to offer.

In 1961, Johnson and Hobb, reported on the Jefferson County three-year study, which included a comprehensive study on class size in many subject fields as well as mathematics. Unfortunately they do not provide any


experimental results but only the conclusion they came to at the end of the study.

It was found that class size did not essentially matter, classes of twenty, thirty-five, sixty achieved the same level. However, looking further into the results the study claimed to show that seventy students and a two-teacher team is better than two 35-students-and-1-teacher classes and also better than seventy students with one teacher and one assistant. Thus an additional factor seems to be in operation. Previous experiments have tried to show that increasing or doubling the class size has no effect on the achievement of the pupils. It has been pointed out that each of the three previous experiments had limitations and one of these was that teacher had assistance. Johnson and Hobb now come up with evidence to show that by employing a two-teacher team with seventy students better results can be obtained than by the conventional one teacher and thirty-five students approach; in addition better than the large class of one teacher and seventy students, with or without a professional helper.

From this Jefferson County report it would appear that class size may not after all be so important, whether classes are considered now as either one-teacher or multi-teacher units. However, without further experimental evidence definite conclusions cannot be drawn from the Colorado study.

The final experiment worth any consideration, is one, however, that is not concerned with mathematics but with science. Immediately it therefore has limitations as stated in Chapter I, but the writer feels that the experiment is worth noting since it relates to the overall conclusions that are drawn later.

Anderson\textsuperscript{18} compared students' final examination marks, from seventy-

\textsuperscript{18} Anderson, K. E., "The Relationship between Teacher Load and Student Achievement." \textit{School Science and Mathematics}, vol. 50, 1950, pp. 468-70.
three high schools in nine states, with teacher load measured as a function of the number of pupils taught per day by the teacher. The report tells that the various classes were equated as the basis of intelligence and previous knowledge of the subject. The abbreviated results stated by Anderson indicate that the students were found to make significantly higher achievement scores when their teachers were assigned few students. This experiment has no detailed results and high school marks are notoriously variable from class to class, so equating the various scores must be questioned. However since Anderson claims that there were significant results where one would expect to end up with an inconclusive result due to the limitations explained, the overall conclusion has some validity.

V. CONCLUSIONS

The problem, as stated in Chapter I, refers to the shortage of mathematics teachers and whether a solution to this can be quickly found by increasing class size. The research evidence offered in class-size experiments needs careful assessment before any definite conclusions can be drawn. The number of actual scientific experiments carried out in mathematics is small and from the data gathered regardless of the tentative conclusions that may be drawn, it is clear that more experimentation is needed.

The results do not give any positive answer. Cunningham,\(^{19}\) Anderson\(^{20}\) and Haertter\(^{21}\) all say that class size appears to have no effect on student achievement. But this is not really the topic being considered. In all cases either the teachers had special clerical assistants and/or they were given extra time to prepare the lessons. Therefore there is little or no saving in teacher time and hence it would appear unjustified to immediately

---

19 Cunningham, op.cit.
20 Anderson, F. H., op.cit.
21 Haertter, op.cit.
increase the number of students unless the teacher was allowed help or extra free time.

There has been a movement to bring in clerical assistants to help the teacher and thus relieve the teacher for more teaching duty. Johnson and Hobb\(^\text{22}\) say that achievement does depend on total teacher pupil-load, and therefore increase in the number of students a teacher has to handle will lower the achievement. In addition Anderson\(^\text{23}\) says that two teachers and a large class is better than one teacher and a large class with or without an assistant. It may be fairer to say that class size in itself may not be so important but the teacher/pupil ratio seems to have a limit for effective teaching.

The research evidence is therefore even very inconclusive, within the controlled scientific experimental situation of the single variable, in solving the shortage of mathematics teachers. The single variable does operate in the school room, so in addition there are many other factors. Probably total teacher load, including teaching, administration, and other sundry duties play an important part in the limiting factor on teacher/pupil ratio than does actual class size alone. Remember, too, in the experimental condition the teacher was only teaching one large class, whilst if this kind of approach was used to solve the shortage, a teacher might have to teach more large classes a day. Miss McGuire\(^\text{24}\) admitted that she had a strain with only one large class.

It would thus appear that an across-the-board increase in class size to solve the mathematics teacher shortage is not viable without built-in precautions of assistance and special techniques.

\(^{22}\) Johnson and Hobb, \textit{op.cit.}

\(^{23}\) Anderson, K. E., \textit{op.cit.}

\(^{24}\) McGuire, \textit{op.cit.}
CHAPTER IV
PROGRAMMED LEARNING AND TEACHING MACHINES

I. INTRODUCTION

In recent years industry and commerce have undergone a revolution. From the nineteenth century organization of the mechanization of man's abilities, we have now realized the labour resources are not unlimited and that machines must be employed to exploit to its greatest potential the limited man-power we have. So far, however, education seems to have been neglected in this revolution. Maybe the resources of human material for teaching mathematics are limited, too. Therefore the problem is two-fold: to increase the supply, and to make certain that we utilize our resources to the full.

In this paper the writer is not concerned with the methods that might be used to increase the supply of teachers. However, programmed learning does come into the second category mentioned above, as does television, calculators and other teaching aids which will not be discussed here.

II. HISTORY AND DEVELOPMENT OF AUTOMATED TEACHING DEVICES

A teaching machine is a device for presenting a programme of self-instruction to a student in such a way that he must take an active part in the proceedings and receives immediate knowledge of progress at every stage of the learning process.

One of the earliest of such devices was produced by S. Pressey in the 1920's. This consisted originally of a simple multiple choice testing device. Information and questions were presented to the student, who was required to select an answer from a group given to him. The machine recorded this answer and thus acted as a simple testing device. This was modified later to become an instructional device by arranging that, if a student answered a question correctly then the next item appeared, but if he answered incorrectly he was required to try again, as the machine did not move forward until the correct
answer had been selected. None of our current social, economic and educa-
tional pressures were present at the time, and research did not receive any
impetus.

After the Second World War, new work was developed by Professor Skinner
at Harvard. Believing in the importance of immediate knowledge of correct
results as motivation, avoidance of any wrong answers as likely to produce
misconceptions difficult to eradicate, and the need for active student parti-
cipation in the learning process (an application of his stimulus-response-
reinforcement psychology of learning). He designed small step programmes in
psychology for college students and obtained impressive performances from the
students in learning time and test scores. The steps in the programmes must
be carefully designed and ordered, and students should make few if any errors
in order to insure the motivation to continue. Machines to contain these
programmes can be simple in design—all that is necessary is a method of pre-
senting the item and question to which the student responds before seeing the
correct answer. This may be done either by a mechanical machine or in a modi-
fied book form.

Programmes following the Skinner pattern are called "constructed response"
programmes since the student is required to make his own answer and then seek
confirmation from the programme. On the other hand, the earlier method of
Pressey offered the student a selection or "multiple choice" of answers from
which to select one. A development of the "multiple choice" method has been
made by Crowder and others. Instead of taking a particular theory of learning
as their starting point they have attempted to simulate the tutor-student situ-
ation, maintaining that wrong answers are likely from time to time in the nor-
mal tuition process and it is the tutor's job to correct errors and misconcep-
tions. Their programmes, therefore, provide corrective material for wrong
answers. To achieve this the programme offers a multiple choice of answers to each question—one correct and the others "likely" wrong ones. With each wrong answer is additional information intended to clear the misunderstanding and the student is required to try again. These types are generally known as branched programmes.

Experiments in programmed learning fall into two classes: comparison of various programmed methods and, secondly, a comparison of the traditional teaching method with some programmed method or technique. In relation to the problem, the experimental evidence will be examined, firstly, to compare the various modes of programming and then to examine the programmed instruction experiments versus the traditional teacher method.

Consider first briefly a few experiments comparing the various techniques of programmes, so that, when these methods are used to compare with the traditional teaching method, there will be some standard by which the methods may be measured.

III. COMPARISON OF PROGRAMMED TECHNIQUES

Roe compared linear (Skinner-type) with forward and backward branched programmes. He constructed a ninety-three-item programme in the three methods and each programme was set up in identical boxes or machines. The programme was part of an introductory course of statistics, and was given to ninety-odd students randomly divided into three groups. The groups were pre-tested and post-tested after the period of time working on the material. For statistical comparison, analysis of variance was employed.

The programmes were effective, first, in showing a significant change in the performance level in all three groups. The forward branching group and the linear programme group showed to have achieved the same learning

---

and took the same time in doing it, but the backward branching programme proved to be inferior in time and learning.

The programme was indeed short, and to draw any definite conclusions would be rash. Perhaps a significant difference between the forward branching and the linear programmes might show up if the total programme had been longer. Otherwise the experiment appears to be well conducted.

Eigen took seventy-seven 8th grade New Hampshire boys, split them randomly into three groups. He took the same programme—developed from Chapter I of Modern Math by Eigen, Kaplan and Emersen—and made it into three different forms. He constructed a linear programme for a machine, in a vertical book form and a horizontal book form. The vertical book was such that the pupils went down the page from item to item or frame to frame, whilst in the horizontal form, the students had to turn over the page for the answer to a question.

There were only eighteen machines available, so the machine group was small, the other two were equally divided. The programme was of sixty-five items and the time taken by all students was about the same. The groups were pre-tested for equality and twice post-tested, the first shortly after the programme had been administered, the second given after a two-week interval.

Eigen found that there was no significant difference between the three groups in performance, but the machine group did take a much longer time. Eigen suggests that in the regular school situation the vertical book programme would have the advantages of reduced cost of printing and general ease of operating, even though it is not cheat-proof like the machine.

This experiment has the same failing that Roe's had before—it is so short, taking only two hours as an average, that it is not surprising that a

---

significant difference was not found. One wonders how this may be applied to the school situation.

Stolurow and Walker compared overt (written) and covert (normally thought) responses in a programmed learning experiment in statistics. They randomly distributed fifty-six subjects into two groups, and gave them a programme in statistics of unknown length. The group were pre-tested and three paper and pencil tests were used to measure the performance after the programme had been taken.

Stolurow and Walker found that there was no significant difference in learning or retention for the two groups, though the time taken was significant in favour of the "covert" group.

Unfortunately the validity of the results of this experiment is severely reduced as the authors did not mention the length of the programmed material. If the programme was short the problem arises again whether it possible to expect a significant difference after a short time. As there appear to be no other experiments in this particular format, there is no evidence to show whether the written response is better than the mere thought response, though educational psychologists state that the written response is better.

To summarize the experimental data of the three experiments, it would appear that there has been no significant difference consistently found that makes any one type of machine or book, one technique of presentation linear or branching, superior to another; though there seems to be an indication that the linear programme might be better if a careful controlled experiment was carried out. The main advantages appear to lie outside the experimental situation, in that the linear programme is probably easier to construct, whilst the book form is cheaper, more versatile, easier to manufacture and lends

itself to either covert or overt response. However it is very clear that the experimental evidence is weak, long term programmes need to be conducted in the experimental situation before definite conclusions can be drawn.

IV. TRADITIONAL CLASSROOM versus PROGRAMMED TECHNIQUES

Virtually carrying on from the previous described experiments, Dessert has carried out a very good experiment comparing not only various programmed techniques but also comparing them with the traditional teacher-taught class.27

The programme used came from 8th grade material based on the University of Maryland Mathematics Project.

Taking eighty students with a mean IQ of 121, as measured by the California Test of Mental Ability, they were split into seven random groups and pre-tested according to reading ability, IQ, and mathematical competence. Dessert had one teacher-taught class designated "T.T." The six other classes formed were each identified by a serial of three letters, each letter standing for a certain type of group they were in. The groups either had a linear (L) or a branching (B) programme. The groups were aided by the teacher (W) or had no aid (N). The groups either finished the programme in their own time and therefore studied the programme just once (S) or they used up the time to equalize with the T.T. group denoted (F).

The six groups were BSW, BSN, BFN, LSW, LSN, LFN, with the letters standing for the abbreviations indicated above.

The seven groups studied the elementary topics of convergence and divergence of infinite series under the seven procedures discussed above for about one week using this 200-item programme, developed for the purpose by professionals. At the conclusion the students were administered a test, which had been constructed at the design phase of the programme and had been tested

on eleven pre-experiment students. The post-test scores were studied statistically by an analysis of variance with the individual post-test means compared by Kramer's extension of Duncan's Multiple Range Test. The post-test means were further analysed by an analysis of co-variance in which IQ was used as the control variable. The set of adjusted means (for slight differences in the groups) were analysed as above, but the results did not change significantly.

Table V shows that the T. T. post-test mean was significantly greater than the post-test means of LSW, BSN, and BSW groups. The TT group was significantly better on the post-test than the combined branched groups, but did not exceed significantly the linear groups. If the time factor was reduced the differences remain but no longer become significant.

The fact that the TT group did better than the programmed groups over-all leads one to speculate that if this programmed instruction had been lengthened to more than the 2 1/2 hours that these differences would have become significant. However it is apparent that time is an important consideration in learning. Dessert suggests that programmed material could be integrated into the classroom in order to free the teacher for more personal pupil contact. He further claims that his experiment has shown the importance of the teacher in the classroom.

Up to now all the programmed learning experiments in mathematics have been very short. In 1961-62 Banghart and others conducted a one-year study in which programmed material was compared with the traditional teaching approach for a modern mathematics course in the 4th grade. A linear programme was utilized for the experimental group, whilst the control group utilized a student text and supplementary material.

---

28 Dessert, op.cit., p. 520.

A total of one hundred ninety-five pupils from Norfolk, Virginia, were divided into six classes, three control and three experimental. The classes were spread over an unstated number of schools but one suspects it was three schools. It is stated by the author that the groups were a cross-section of 4th graders as shown in terms of intelligence, achievement, and socio-economic status. There is however no mention as to how the groups were formed and no indication given as to whether a pretest was used.

The experimental and control classes were taught for one year for thirty to forty minutes per day. At the end of the experiment, the Metropolitan Battery Arithmetic Section was given to measure the achievement of all groups. The post-test results given in Table VI indicate that the experimental programmed group was superior to the control group at a significant level of \( p < .05 \). This must immediately be questioned as there is indication that the groups were truly equated in any recognized scientific manner. Of more importance to the problem in hand is that the experimental groups always had a teacher in the classroom and the teacher was used in the instructional process. This means, therefore, that in this experiment there was no saving of teacher resources. It would thus appear, allowing for the absence of the pre-testing and proper allocation of the groups, that programmed learning is successful with the aid of a teacher to help the students individually.

It would be quite possible to relate several small experiments that have been done with programmed materials compared with the "teacher-taught" class. Generally the results have been very similar in that no significant difference has been found between the experimental and control groups or that programmed materials are shown to be superior. Generally the various experiments have had the same fault in that they have been conducted over a short time span, lessening the chance that a significant difference might appear and the experimental rigour has left much to be desired.
To show the weakness of many of these so-called experiments in programmed learning, Pressey has come up with some interesting results. For some years now the experiments have taken small parts of a normal textbook material and rewritten these parts in programmed form. They have then run an experiment and have attempted to show that the programmed method is better or at least as good as the conventional method. This they often achieved and by expert presentation have made their results seem very impressive.

Now Pressey has done the reverse process. He took a part of Holland and Skinner's *Analysis of Behavior*, some fifty-four frames and eight pages of it in fact, filled in all the blanks, reprinted it and ended up with a continuous but repetitive discourse of some 1100 words. He then reconstructed this material into a passage of some three hundred sixty words of normal written English. Colleagues agreed that it taught the same amount of material. To this he added seven auto-elucidative questions (right answers appear when correctly marked).

Four groups were formed both in psychology and education. The groups were tested for equality and were found to be so statistically. The average size of each group was thirty-one students. The four groups were a control group with no instruction, a Skinner programme group, a Pressey statement group, and a Pressey statement plus seven questions group. After the study period all the groups were post-tested.

Table VII shows that all groups made significant gains over the control group. The education group learned as much by reading the Pressey statement as doing the Skinner programme, and in one-fifteenth of the time. With the Pressey statement and the seven auto-elucidative questions, both the psychology and education students did better than those on the Skinner programme. All

---

these results are an exact reversal of the numerous other experiments.

The Pressey experiment was conducted in psychology, but the writer believes that the weaknesses purposely shown up by Pressey apply to many of the programmed learning experiments in mathematics. Though Pressey's experiment is short, this is exactly the kind of fault he is showing up. Not only, he claims, are the majority of experiments too short, but often the "Hawthorne" or novelty effect is operating in the experimental situation, just as it is in his own. He states:

In short, students do learn and very rapidly from silent reading without overt responding. Challenging questions, asked by a teacher or an auto-instructional device, which differentiate right from wrong ideas, appear to help in such differentiation. It would seem evident that the meaningful learning here under consideration is not a series of bit learnings struck on by "reinforcements" but rather a progressive process of cognitive clarification and integration.31

V. SUMMARY

Pressey's experiment, though not in the field of mathematics, clearly shows that many programmed learning experiments have failed to show significant results to the inadequacies of them, and the over-optimistic conclusions drawn from the results.

There have been, however, one or two experiments that are worthy of further consideration. Banghart's one-year experiment32 at least overcomes the main objection that Pressey puts forward. It would be indeed difficult to maintain for one year the novelty effect that is said to operate in the experimental situation. In Banghart's experiment, he show that programmed learning can be as successful as the traditional teaching method. However, in his experiment the teacher did not disappear from the experimental classroom, neither does the teacher have an increased class population. It would thus appear by the extremely limited experimental evidence that none of the

31 Pressey, op.cit., p. 415.
32 Banghart, op.cit.
various programmed methods have indicated a measure of improvement that would allow a teacher to take on an additional load and thereby reduces the shortage of mathematics teachers.
CHAPTER V

THE ROANOKE EXPERIMENT

It was very clear at the outset of this study that an experiment of a long-term nature was needed to evaluate the use of programmed materials in schools.

In 1960, Roanoke City Schools embarked on a long-term appraisal of programmed materials. The writer has searched diligently into every reference pertaining to this experiment, but was unable to find any worthwhile report on the whole project. Just recently and too recently to include in the previous chapter, the writer received some information from Roanoke and Encyclopedia Britannica Films (EBF), who are supplying the materials for the project.

There is planned to be a full-scale report of the project, which is due to be published in the summer of 1964, though this report has already been delayed once.

In the material that has been received, and based on early work in 1960-61, the Superintendent of the City School System reports that a mathematics class has successfully completed a year of algebra in one semester at the 8th grade level. The students were without a teacher, textbooks, or homework. However there is no statistical evidence, save that 41% of the 8th graders surpassed the average score made by 9th graders in a national examination. This is not evidence that is reliable as those 41% could well have been the brightest students and the fact that they took a programmed course does not make the slightest difference.

One can only await the final analysis of the experiment and hopefully it will be a well-constructed and analytical report.

Already the writer has heard "gossip" to the effect that the introduction of programmed materials into several systems in Virginia has met with
failure. Talk of locking the machines away in a cupboard after one year's work is common and even the Roanoke experiment has come under the eye of the critical even before the results are fully known.

Naturally big business has a stake in the Roanoke Experiment, and if the reports from EBF are equally as vague, it appears that there will be no reported hard facts and figures from Roanoke. EBF put out glossy sheets, on which so-called case histories of programmed materials are printed. Beautifully presented, just as good experimental reports should be, these reports on closer examination show that the evidence given is not so foolproof as it first appears.

One outstanding fact appears in these reports. In all cases save one, the students were helped by the teacher. Mr. Kettner states, referring to a programmed algebra course, "I feel you need a strong teacher for this programme. You need a teacher with a wide background of mathematics." 33

Though some experts claim that essentially the teacher need not be present in the programmed learning situation it is becoming more and more obvious that the teacher is an essential part of the learning situation.

Before further considerations and conclusions can be drawn about this long-term project at Roanoke, the report promised for mid-1964 must be seen, though the writer does not hope for much of an improvement over the non-scientific reports that have been published up to now and also doubts whether it will shed much more light on the experimental evidence on programmed learning already known.

CHAPTER VI
CONCLUSIONS AND IMPLICATIONS

The conclusion may be summarized as follows:

(i) There is no experimental evidence to show that there is an optimum class size.

(ii) There is no evidence, however, to show that a teacher can successfully handle a large group of students, without the teacher having assistance and/or extra time for preparation.

(iii) There is no evidence to show that by teaching large classes the teacher saves time and thus could take on extra classes.

(iv) Immediate increase in class size to solve the shortage of mathematics teachers has no experimental backing.

(v) Learning in mathematics can take place by using programmed materials.

(vi) No technique of programming whether linear or branching has been shown to be absolutely superior in method to traditional teaching method.

(vii) Under limited conditions programmed learning materials with a teacher have been shown experimentally to be as good as the traditional teacher approach.

(viii) There is no evidence to show that programmed learning can replace the teacher completely and still maintain the same standard of achievement, within the normal school situation.

(ix) Immediate introduction of programmed materials to eliminate the teacher shortage in mathematics has no experimental backing.

In order to prevent this study's becoming a mere overview of many topics, it was first limited to mathematics and to two specific areas of class-size and programmed learning. It is now apparent that there are many other factors that need to be considered. Among these factors are team-teaching, television, total teacher load, the place of non-professional assistants, personality of the individual student and the real purpose of secondary education. Before the shortage of mathematics teachers can be overcome, these factors will need to be considered.

It has been indicated previously that the manpower available for mathe-
Mathematics teaching may indeed be running out and that it will be necessary to exploit this available manpower to near perfect efficiency as possible. It will be therefore necessary to examine all the factors mentioned, before a rearrangement can be made to reduce the shortage of mathematics teachers or even to ensure that if the output of mathematics teachers cannot be increased there is a reserve plan that the present supply is used to maximum capacity and efficiency.
FIGURE I
EXTENT OF TEACHER SUPPLY BY TEACHING FIELD BASED ON OPINIONS OF EMPLOYING AGENTS

[Bar chart showing supply and shortage by subject, including Social Studies, Sociology, Government, English, Economics, Physical Education, Mathematics, Art, Spanish, Primary Graded English, Latin, Chemistry, Mathematics, and Physics.
TABLE I
PROGNOSIS AND ACHIEVEMENT DATA IN CUNNINGHAM'S CLASS-SIZE EXPERIMENT

<table>
<thead>
<tr>
<th>I.Q.</th>
<th>PROGNOSIS</th>
<th>CITY TEST</th>
<th>COLUMBIA TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASES</td>
<td>L. 55</td>
<td>S. 28</td>
<td>L. 55</td>
</tr>
<tr>
<td>TOTAL RANGE</td>
<td>80-150</td>
<td>80-150</td>
<td>60-200</td>
</tr>
<tr>
<td>MODAL STEP</td>
<td>100-110</td>
<td>110-120</td>
<td>110-120</td>
</tr>
<tr>
<td>MEDIAN</td>
<td>113.2</td>
<td>113.3</td>
<td>137.5</td>
</tr>
<tr>
<td>MEAN</td>
<td>114.8</td>
<td>113.6</td>
<td>133.6</td>
</tr>
<tr>
<td>S.D.</td>
<td>12.4</td>
<td>13.3</td>
<td>23.7</td>
</tr>
<tr>
<td>Age exceeding other group</td>
<td>49.8</td>
<td>50.4</td>
<td>61.8</td>
</tr>
</tbody>
</table>
TABLE 11

PRE-TEST AND POST-TEST DATA FOR CAMELBACK HIGH SCHOOL
EXPERIMENT IN ALGEBRA.

<table>
<thead>
<tr>
<th></th>
<th>MEANS</th>
<th>STANDARD ERROR OF THE DIFFERENCE</th>
<th>DIFFERENCE OF MEANS BETWEEN CLASSES</th>
<th>DIFFERENCE OF MEANS BY CLASS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PRE-TEST</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LARGE 66</td>
<td>40.2</td>
<td>39.6</td>
<td>5.4</td>
<td>1.668 for large class</td>
</tr>
<tr>
<td>SMALL 33</td>
<td>41.0</td>
<td>41.2</td>
<td>1.3</td>
<td>1.22</td>
</tr>
<tr>
<td>POST-TEST</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LARGE 66</td>
<td>31.0</td>
<td>31.0</td>
<td>5.8</td>
<td>1.115 for small class</td>
</tr>
<tr>
<td>SMALL 33</td>
<td>32.2</td>
<td>32.5</td>
<td>5.8</td>
<td>.99</td>
</tr>
</tbody>
</table>

For large class, the difference of means is significant.
For small class, the difference of means is not significant.
TABLE III
PAIRING AND ACHIEVEMENT DATA IN HABERTER'S EXPERIMENT IN PLANE GEOMETRY.

<table>
<thead>
<tr>
<th></th>
<th>AVERAGE I.Q.</th>
<th>MEDIAN I.Q.</th>
<th>AVERAGE FRESHMAN MARKS</th>
<th>AVERAGE MENTAL AGE mths.</th>
<th>PAIRING TESTS</th>
<th>MEDIAN ACHIEVEMENT SCORES max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>SMALL-20</td>
<td>111.1</td>
<td>112.0</td>
<td>63.0</td>
<td>196.2</td>
<td>21.6</td>
<td>13.3</td>
</tr>
<tr>
<td>LARGE-20*</td>
<td>111.1</td>
<td>110.5</td>
<td>63.3</td>
<td>202.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LARGE-40</td>
<td>111.4</td>
<td>109.5</td>
<td>60.3</td>
<td>202.3</td>
<td>21.6</td>
<td>12.3</td>
</tr>
</tbody>
</table>

* The twenty pupils most closely paired man for man with the members of the small class.
### TABLE IV
PAIRING AND ACHIEVEMENT DATA CN MISS. McGUIRE'S EXPERIMENT

#### PAIRING DATA

<table>
<thead>
<tr>
<th>SECTION</th>
<th>SIZE</th>
<th>SEX PAIRS</th>
<th>AGE</th>
<th>I.Q. TESTS</th>
<th>FRESHMAN</th>
<th>CHARACTER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M</td>
<td>F</td>
<td>MONTHS.</td>
<td>TESTS</td>
<td>MEAN</td>
</tr>
<tr>
<td>SMALL</td>
<td>23</td>
<td>9</td>
<td>10</td>
<td>178.5</td>
<td>116.1</td>
<td>67.2</td>
</tr>
<tr>
<td>LARGE</td>
<td>44</td>
<td>9</td>
<td>10</td>
<td>179.9</td>
<td>116.2</td>
<td>69.4</td>
</tr>
</tbody>
</table>

#### ACHIEVEMENT DATA

<table>
<thead>
<tr>
<th>SECTION</th>
<th>SIZE</th>
<th>SEX PAIRS</th>
<th>AGE</th>
<th>LOCAL</th>
<th>MINNEAPOLIS</th>
<th>HAWKESWOOD</th>
<th>SCHARLING</th>
<th>ESSENTIALS</th>
<th>TEST</th>
<th>SANFORD</th>
<th>TEST</th>
<th>MEAN</th>
<th>MEAN</th>
<th>MEAN</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SMALL</td>
<td>23</td>
<td>as above</td>
<td>526.1</td>
<td>569.0</td>
<td>h05</td>
<td>37.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LARGE</td>
<td>44</td>
<td>as above</td>
<td>572.8</td>
<td>577.9</td>
<td>h1.4</td>
<td>39.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There is no significant difference between the two groups tested
### Table V

**Achievement Results in Dessart's Programmed Learning Experiment**

<table>
<thead>
<tr>
<th>GROUP MEANS</th>
<th>BS</th>
<th>BN</th>
<th>BN</th>
<th>LS</th>
<th>LS</th>
<th>LNS</th>
<th>TT</th>
</tr>
</thead>
<tbody>
<tr>
<td>POST-TEST MEANS OF 7 GROUPS</td>
<td>33.2</td>
<td>30.4</td>
<td>37.0</td>
<td>31.8</td>
<td>40.2</td>
<td>37.5</td>
<td>41.8</td>
</tr>
<tr>
<td>ditto ADJUSTED MEANS</td>
<td>33.6</td>
<td>30.3</td>
<td>37.7</td>
<td>32.6</td>
<td>40.2</td>
<td>37.6</td>
<td>39.8</td>
</tr>
<tr>
<td>MEAN STUDENT TIME IN MINUTES</td>
<td>94.6</td>
<td>100.4</td>
<td>150</td>
<td>109.9</td>
<td>109.9</td>
<td>150</td>
<td>150</td>
</tr>
<tr>
<td>POST-TEST GROUP MEANS</td>
<td>33.6</td>
<td>36.9</td>
<td>41.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ditto ADJUSTED</td>
<td>34.0</td>
<td>37.2</td>
<td>39.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*NOT SIGNIFICANT*
### TABLE VI

ACHIEVEMENT RESULTS OF FOURTH GRADE PROGRAMMED MATERIAL EXPERIMENT AT NORFOLK, VIRGINIA.

<table>
<thead>
<tr>
<th></th>
<th>EXPERIMENTAL GROUP</th>
<th>CONTROL GROUP</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEAN SCORE</td>
<td>58.64</td>
<td>55.68</td>
</tr>
<tr>
<td>STANDARD DEVIATION</td>
<td>10.33</td>
<td>11.48</td>
</tr>
<tr>
<td>NUMBER OF PUPILS</td>
<td>107</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>.71</td>
<td>.87</td>
</tr>
<tr>
<td></td>
<td>( t = 2.37 )</td>
<td>( p &lt; .05 )</td>
</tr>
</tbody>
</table>
TABLE VII
EXPERIMENTAL RESULTS FOR PRESSEY'S "CONTRA-EXPERIMENT" IN
PROGRAMMED LEARNING.

<table>
<thead>
<tr>
<th>GROUPS</th>
<th>MEDIAN TIME FOR PSYCHOLOGY IN MINS</th>
<th>MEDIAN TIME FOR EDUCATION IN MINS</th>
<th>POINTS PER PSYCHOLOGY SUBJECT (MEAN)</th>
<th>POINTS PER EDUCATION SUBJECT (MEAN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONTROL</td>
<td>-</td>
<td>-</td>
<td>2.0</td>
<td>3.6</td>
</tr>
<tr>
<td>H/SKINNER PROGRAMME</td>
<td>22</td>
<td>23</td>
<td>8.5</td>
<td>10.5</td>
</tr>
<tr>
<td>PRESSEY STATEMENT</td>
<td>1.1</td>
<td>1.5</td>
<td>6.0</td>
<td>10.5</td>
</tr>
<tr>
<td>PRESSEY STATEMENT</td>
<td>3.5</td>
<td>4.3</td>
<td>9.3</td>
<td>12.4</td>
</tr>
<tr>
<td>PLUS SEVEN ELUCIDATIONS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SELECTED BIBLIOGRAPHY

A. BOOKS

Hudelson, Earl. _Class Size at the College Level_. University of Minnesota Press, 1928.


B. PERIODICALS


Anderson, K.E., "The Relationship between Teacher Load and Student Achievement.", _School Science and Mathematics_, vol. 50, 1950, pp. 468-70.


NEWSPAPERS AND BULLETINS.

