Inductive versus deductive methods of teaching : an experimental research / by W. H. Winch.

Contributors

Winch, W. H.

Publication/Creation

Baltimore : Warwick & York, 1913.

Persistent URL

https://wellcomecollection.org/works/d9repu5k

License and attribution

Conditions of use: it is possible this item is protected by copyright and/or related rights. You are free to use this item in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s).

Wellcome Collection
183 Euston Road
London NW1 2BE UK
T +44 (0)20 7611 8722
E library@wellcomecollection.org
https://wellcomecollection.org

Educational Psychology Monographs

3370.

THE
CHARLES MYERS
LIBRARY

Spearman
Collection

NATIONAL INSTITUTE
OF
INDUSTRIAL
PSYCHOLOGY

22500416485

Med
K41773

NATIONAL INSTITUTE OF
INDUSTRIAL PSYCHOLOGY
LIBRARY

7/6 net

Digitized by the Internet Archive
in 2016

https://archive.org/details/b2806298x

INDUCTIVE vs. DEDUCTIVE METHODS OF TEACHING

Educational Psychology Monographs

Edited by Guy Montrose Whipple

No. 11

Inductive versus Deductive Methods of Teaching: An Experimental Research

By

W. H. WINCH

External Member of the Psychological Board of Studies for the University of London
Chairman of the Committee of the Teachers' Guild of Great Britain and Ireland
on Psychological Research in Schools; Lecturer for the London County
Council on Pedagogical Methods in Schools.

Author of "Problems in Education," "German Schools," "When Should a Child Begin School," etc.

Baltimore, U. S. A.
WARWICK & YORK, Inc.
1913

Copyright, 1913
WARWICK & YORK, Inc.

WELLCOME INSTITUTE
LIBRARY
Coll. WelMGmec
Coll.
No. WM

UEB

EDITOR'S PREFACE.

It affords me great pleasure to call editorial attention to this interesting and instructive contribution to experimental pedagogy. Mr. Winch writes with the authority of long experience born of his professional duties as one of the official inspectors of English schools. He is, indeed, well known as the first Englishman to bring the technique of experimental and statistical methods to bear upon the actual practical problems of the school.

Those who have followed with any care the modern developments of educational theory know how significant is that trend of investigation which seeks to study the concrete problems of education at first hand in the classroom and with all the exactness of experimental control. The movement for experimental pedagogy is yet in its infancy, but it has already shown the possibilities that lie before it. In the *Journal of Educational Psychology,* with which this series of *Educational Psychology Monographs* is affiliated, there has appeared of late an important series of articles which show for various school subjects what important problems offer hope of solution by experimental investigation. This monograph presents what is at the very least a first approxima-

tion toward the solution of one of these vexed questions of educational practice: Is it better to follow deductive or inductive methods in the teaching of various types of subject-matter? The presentation has the special merit of being sufficiently detailed that any teacher who desires to do so may of himself repeat the experiments and verify the conclusions.

G. M. W.

AUTHOR'S PREFACE.

This is, I believe, the first attempt to decide between the conflicting claims of 'inductive' and 'deductive' methods by experimental procedure. In the 'world of science' it is not usual for results to be accepted unless the methods by which they have been obtained are described in such detail as enables other workers to repeat, corroborate, or modify them. Nor are they regarded as valid unless they are obtainable under widely differing external circumstances. To produce similar evidence for educational science will be the aim of all serious workers in education during the next two or three decades, and I am therefore offering this research as a contribution to the scientific knowledge of the results of inductive and deductive methods in actual application under school conditions.

I am quite well aware that much valuable knowledge is collected by school administrators and school inspectors during the ordinary course of their work. They know much about the results of the application of different methods in different schools. But to disentangle all the contributory factors—even to realize them—is very difficult, and inspectors are likely to be misled; for the teacher is, naturally, mainly desirous of showing that his school is a good one, and not of

settling, by experimental tests, the value of a particular method. The work reported in this monograph is not subject to this source of error, since the teachers, in this case, were working *with* the experimenter, and not *against* him. It is my firm and ever-growing conviction that without that kind of co-operation on the part of teachers there can never be, in an applicable sense, a 'Science' of Education.

W. H. W.

London, September, 1912.

CONTENTS.

Statistical Note 7

I. Introduction 11

II. The Problem of the Experiments 17

III. The General Plan of the Experiments 20

IV. First Series of Experiments:

1. General Characteristics of the Children Who Worked the Exercises 23
2. The Preliminary Tests 23
3. The Method of Marking the Preliminary Tests . 25
4. The Teaching of the Two Groups 31
5. The Immediate Testing of the Two Groups . . . 35
6. The Marking of the Tests 36
7. The Subsequent Testing of the Two Groups on the Same Subject-matter 37
8. The Testing of the Two Groups on New Material . 38
9. The Marking of the New Material 40
10. Chronology of the Experiment 44
11. Results:
    - (a) The Marks for the Preliminary Tests . . . 45
    - (b) The Marks for the Test Immediately After the Definitions Had Been Taught and Learnt 46
    - (c) The Marks for the Tests of Deferred Reproduction 47
    - (d) Correlation Between Immediate and Deferred Reproduction 49
    - (e) Results When the Two Groups Are Tested on New Material 50
12. Pedagogical Conclusions 53

V. Second Series of Experiments:

1. General Plan 55
2. The Preliminary Tests and the Method of Marking. 56
3. Chronology of the Experiment 59
4. The Final Tests and the Method of Marking . . 60
5. Results of the Experiment:
    - (a) Results of the Preliminary Tests 61
    - (b) Results of the Tests in Immediate and Deferred Reproduction 63

(c) Correlation Between Immediate and Deferred Reproduction 65
(d) Results of the Test on New Material . . . 67
VI. Third Series of Experiments:
1. General Plan 69
2. The Preliminary Tests and the Method of Marking. 70
3. Chronology of the Experiment 75
4. The Final Tests and the Method of Marking . . 76
5. Results of the Experiments:
(a) Results of the Preliminary Tests 90
(b) Results of the Tests in Immediate and Deferred Reproduction 92
(c) Correlation Between the Results of Immediate and Deferred Reproduction 95
(d) Results of the Test on New Material . . . 96
VII. Fourth Series of Experiments:
1. General Plan 100
2. The Preliminary Tests and the Method of Marking. 101
3. Chronology of the Experiment 104
4. The Tests of Immediate and Deferred Reproduction 107
5. The Test of Application to New Material 107
6. Results 114
VIII. Fifth Series of Experiments:
1. General Plan 119
2. The Preliminary Tests and the Method of Marking. 120
3. Chronology of the Experiment 122
4. The Tests of Reproduction 124
5. The Test of Application to New Material 129
6. Results:
(a) Of the Preliminary Tests 133
(b) Of Immediate Reproduction 134
(c) Correspondence Between Immediate and Deferred Reproduction 135
(d) Results of the Test on New Material . . . 138
IX. General Summary 140

STATISTICAL NOTE.

Suppose we have two measurements of any mental function for a number of children, that the second measurement gives higher results than the first in most cases, and that the average mark for the second measurement is a little higher than for the first. May we, therefore, conclude that some general tendency is at work, or must we regard the higher average of the second measurement as the result of 'chance' or mere variability? To answer this question I propose to illustrate the usual statistical check on results of this kind by means of one or two examples. Suppose the children are measured for their power of spontaneous definition; that, a week later, they are measured again, and that the marks are as shown in the following table:

Name.	First measurement.	Second measurement.
A. B.	9	10
C. D.	8	9
E. F.	7	8
G. H.	6	7
I. J.	5	6
K. L.	4	5
M. N.	3	4
O. P.	2	3
R. S.	1	2
	Average, 5	Average, 6

Common sense has no difficulty in deciding that there *is* a 'general tendency' to improvement from one exercise to another. Let us now calculate the

'probable errors.' The 'probable error' of an average is $\frac{.67449\sigma}{\sqrt{n}}$, where '$\sigma$' is the standard deviation, and 'n' is the number of cases measured.* Worked out on this formula, the 'probable error' of the average 5 is approximately .6, and of the average 6 is also approximately .6. The 'probable error' of the *difference between two averages* is $.67449 \sqrt{\frac{\sigma^2_1 + \sigma^2_2}{n}}$, where '$\sigma_1$' is the standard deviation of the first average, 'σ_2' is the standard deviation of the second average, and 'n' is the number of cases measured. Applying this formula to the present example, we have the 'probable error' of the difference between the two averages $= .67449 \sqrt{\frac{(2.6)^2 + (2.6)^2}{n}}$, which is approximately .8.

It is required statistically that the difference between two means shall be twice (or more) the 'probable error' of that difference before the difference is supposed to be 'significant,' that is, indicative of a general tendency. But the difference between the means in this case is only 1 and its 'probable error' is .8, so that, apparently, we have no 'significant' difference at all.

But let us consider one more illustration in which the averages are the same, but in which common sense would *not* find a general tendency to improvement:

---

*Simple illustrations, in which 'σ' is found from easy examples, are given in the statistical note attached to my monograph, *When Should a Child Begin School?*

Name.	First measurement.	Second measurement.
A. B.	9	2
C. D.	8	3
E. F.	7	4
G. H.	6	5
I. J.	5	6
K. L.	4	7
M. N.	3	8
O. P.	2	9
R. S.	1	10
	Average, 5	Average, 6

The difference between the means is again equal to 1, and the 'probable error' of the difference, calculated just as before, is .8. Statistically, therefore, we are precisely in the same position as in the previous example, and there is no 'general tendency' to improvement. But quite obviously the two cases are by no means similar and their 'probable errors' are *not* the same, for we have overlooked the positive correspondence between the first and second measurements of A. B., C. D., and the rest in the first illustrative case and the negative correspondence in the second illustrative case. The theory of statistics takes account of this correspondence, or lack of correspondence, in the following formula for the 'probable error' of the difference between two averages:

$$.67449 \sqrt{\frac{\sigma^2{}_1 + \sigma^2{}_2 - 2r\sigma_1\sigma_2}{n}},$$

where 'σ_1' is the standard deviation of the first average, 'σ_2' is the standard deviation of the second average, 'r' is the coefficient of correlation between the two series of measurements, and 'n' is the number of cases measured. I suggest now that the correlation

coefficient be worked out for both cases. In the first case 'r' will be found to be + 1 and in the second case — 1.* Let us then apply the formula, corrected for correlation, to the two illustrative cases. The 'probable error' in the first case is

$$.67449 \sqrt{\frac{(2.6)^2 + (2.6)^2 - 2\,(+1)\,2.6 \times 2.6}{n}}.$$

It will be seen at once that the expression disappears, for $2\,(2.6)\,(2.6) = (2.6)^2 + (2.6)^2$: that is to say, the difference between the averages is *perfectly* representative of the two series of measurements, as common sense would suppose. In the second case the 'probable error' is

$$.67449 \sqrt{\frac{(2.6)^2 + (2.6)^2 - 2\,(-1)\,2.6 \times 2.6}{n}},$$

that is, double what it was when the negative correlation was neglected. It now reaches 1.6, and is greater than the difference between the averages, which is only 1. Hence the conclusion is against any general tendency, again in accordance with common sense.

These illustrations will probably be sufficient to show that the use of probable error formulae without regard to correlation may be very misleading, and also that mere averages, without some indication of the nature and extent of the variability of the measurements, may be even more so.

---

*Easy illustrations will be found in the statistical note previously referred to.

INDUCTIVE VERSUS DEDUCTIVE METHODS OF TEACHING: AN EXPERIMENTAL RESEARCH.

I. INTRODUCTION.

In England—it is for Americans to speak for their own country—there is a widely-spread opinion that the theory and the practice of teaching are two very different things. The young student leaves the normal school or training college, and, doubtless crudely enough, begins to put, or to try to put, into practice some of the pedagogical methods which he has been taught as theoretically sound.

Not infrequently—I had almost said invariably—his confidence in his theoretical instruction receives a violent shock. His superiors in the school assure him that he will do no good if he goes on like that. What is worse and much more disconcerting (for, after all, principals and head masters must find fault; it is their métier), his confrères look on with amused tolerance and 'chaff' him about his 'new-fangled' ways. Then, dropping into friendly confidence, they explain to him that it was all very well for him to 'get up' and describe methods of that sort in examination papers; it was expected of him, and, naturally enough, he wished to get his certificate of

proficiency and to do credit to his college. Examiners required these things; they were unpractical persons like professors, but, of course, a wise student humored them; besides, how else could he pass his examinations? Let these fellows take off their coats and come and do a day's teaching in the schools, and they would very soon change their opinions. Their stuff is all theory, and in actual school life is simply no good. Now you have become a man, you must put away from you childish things; and so on. Thus the experienced and disillusioned confrères to the neophyte.

It is not clear that taking off their coats would assist much in such professional conversions as are here anticipated, but the suggestion is a protest against what the teachers regard as a rather visionary and unpractical existence.

If this rude shock resulted in complete divorce, there would be some hopes of other and happier marriages between theory and practice later on; but, in England at least, what happens is rather of the nature of a judicial separation.

The theoretical methods are not absolutely discarded; they are laid by and put in evidence only on special occasions; the practical methods do duty day by day. For it is dangerous, from the standpoint of professional advancement, for the teacher boldly to renounce the methods he has been taught; he is pursued all his life by unpractical, theoretical persons, to wit, inspectors, and he will often deem it to his advantage to profess a method he does not believe in.

Head masters, too, mindful of the repute of their

schools, will say, "Yes, that's all right, but don't do that when Mr. I——r comes; he does not like it; he thinks you ought to teach that this way."

Well, yes, no doubt, but what is there in all this but the usual difficulty which besets the young instructed person when he takes an actual place in the working world: it is the old difficulty of science versus practice. In a few years the teacher, like other people, will have allowed his theory and his practice to interpenetrate, and both will have been improved. In such wise may we suppose an experienced administrator pooh-poohing my criticism.

If I could admit this, my complaint would indeed lose much of its poigancy. But I do not admit it. On the contrary, I believe that with the great majority of teachers there remains permanently an irreconcilable breach between dominant theory and current practice. It is true that experienced teachers—some of them—will attend lectures about educational topics. Two men speaking somewhat loudly after leaving a lecture hall—modesty forbids me to name the lecturer—said one to the other: "I didn't get much from him; he's like all the rest of 'em." "Oh! I don't know," said the other, with a give-the-devil-his-due air, "one gets ideas." "Yes," was the prompt reply, "but they don't work." And this, let us quite clearly understand that, is not merely an expression of a private grumble; it is a strongly held and often a clearly reasoned view.

There are always "new methods" in education, of course, and I hope that those of us who hold the very newest of them are more or less prepared to see our darlings cold-shouldered for a newer birth. Still, behind all temporary fluctuations, there is a line of

steady meaning in such phrases as 'new method,' 'inductive procedure,' 'developmental method,' 'psychological method,' *et id omne genus.* And behind all temporary oscillations there is a steady trend of opinion amongst experienced teachers that these methods have certain serious disadvantages; that though they may be valuable for show purposes in teaching, they are too slow, and the information thus acquired is not really available when it is wanted.

An experienced head master in London wrote to a lecturer on pedagogy in the following terms:

"I (recently) asked a question on the difficulty of covering a present average course (by using the new methods) in the time given to it on the school timetable, and I should like to press the point and illustrate its importance from my knowledge of the views held by others, and especially by the class teachers in my own school, for at almost every conference with my staff this question arises strongly.

"It seems impossible to train children by much individual work in class by inductive methods, much questioning and the consequent necessary waiting for the child's expression to be formulated in a sufficiently acceptable form, and at the same time *to get through the course set in a given time,* and especially properly to prepare the children also *for examination purposes.*

"For instance, in an illustration given of obtaining from the class the definition of the diameter of a circle, the time taken, if similarly applied to other parts of the course, would not permit of a present average syllabus being more than about half completed, nor would the information got be available

throughout the class for reproduction at a more or less distant examination.

"Another illustration: Two years ago an inspector (*fons et origo malorum* W. H. W.), examining Standard V (approximately American Grade VI), asked for a definition of a proper noun, and, not getting a satisfactory answer, tried to obtain it from the boys with the aid of many questions and illustrations. He took up twenty minutes of the lesson, and failed in the end to get what was wanted.

"Of course, all the time the children were being educated on the best lines; they showed eagerness, interest and active thought. (This, I fear, is a concession to the lecturer as a theoretical person. W. H. W.) *But,* the time taken, in view of the rest of the syllabus, was excessive, and the result at the end was not satisfactory."

Then follows a paragraph in which the writer expresses his willingness to carry out the new methods provided the educational authority will dispense with tangible results.

This is a strong letter. It expresses views which are very common, and which, moreover, are held by some of the most successful schoolmasters I know. And they will have to be met by educational science. I, myself, believe that, until these questions are dealt with in such a way as to be acceptable to teachers, the breach between theory and practice will remain. Professors, teachers of method and inspectors may continue as now to receive lip-homage for the methods they advocate, but their directions will be honored rather in the breach than in the observance. In actual practice there will be little, if any, change.

What, then, do I suggest? I propose the experimental determination of these disputed questions in the schools themselves. There is an increasing number of teachers willing—nay, anxious—to carry out such experiments if adequate guidance be given to them. To the description of one attempt at an experimental solution of some of these disputed points I now proceed.

II. THE PROBLEM OF THE EXPERIMENT.

No one can hope to solve all the questions raised in the never-ending controversy about inductive and deductive methods by means of a single experiment or by means of a single series of experiments. Yet, if one attempts to deal with the matter experimentally, one must deal with some definite subject-matter. There is danger in this, since we may find out afterwards that our conclusions are true only for subject-matter of that particular kind; but it is a danger which must be faced.

A good plan, if one is conscious of bias, is to select subject-matter which favors the chances of the method in which personally one does not believe. So I, believing in inductive rather than in deductive methods, chose geometrical definition as the subject-matter for my experiment.

There is much good opinion in favor of deductive treatment of definitions, especially when, as in demonstrative geometry, a sort of system of reasoned conclusions is to be built up upon those definitions. It is argued that a pupil ought to know exactly what the definition means, that the exact wording of it is very important for that purpose, and that at some stage in the procedure the definitions should be memorized. There is no question here about the *introduction* to demonstrative geometry. It is supposed by both parties to the dispute

that manual work and geometrical construction are necessary propaedeutics to any rational system of geometry. But if we are ever to have a system of *demonstrative* geometry we must, it is said, have exact meanings for our terms or we shall never be able to reason in words at all.

This is by no means a weak case, and to it additional importance is given at the present juncture, when so much dissatisfaction is being expressed as to the 'chaos' into which geometrical teaching is falling through what is alleged to be an excess of inductive method.

Those who advocate purely inductive methods urge that the memorizing of the definition and the study of its application to examples is not, in the truest sense, knowing the definition; it is urged that it can be known better if it is built up from the examples. It is asserted that the memorizing of definitions leads to bad errors, 'howlers,' of which the following, once given to me, is a choice example: "A circle is a figure bounded by a straight line, which is such that every point within it is at equal distance from every other point." There is a tendency to concede that inductive methods are slow; there is some tendency also to concede the point that inductive methods will not prepare a pupil so well for examination purposes. But it is argued that what he does learn he will remember longer, and that he will be made more intelligent.

The use of the word 'intelligent' in educational disputes amounts almost to a public scandal, and I do not propose to use it without giving an opponent some way of checking the assertion.

By intelligence, in this case, I am going to mean

the power gained to deal with new material in consequence of the mental processes which the pupil has passed through in acquiring the old. I have found this interpretation of the word acceptable to both parties in the dispute.

Let me now endeavor to disentangle the threads and see how far the assertions made, first on one side and then on the other, are susceptible of experimental determination.

First of all, we can quite easily find out whether pupils taught inductively or deductively know the required definitions better *immediately* after they have acquired them. We shall demand exact meanings, but not a stereotyped form of words. A child taught inductively would not, of course, know a particular form of words for a definition, but no deviation from accuracy of meaning must be allowed.

Secondly, we can find out, by repeating the exercises later on, whether the pupil forgets more or less after one method than after the other. We are thus testing the *durability* of his knowledge.

Thirdly, we can find out how many '*bad errors*' are made by pupils who are taught by inductive and deductive methods, respectively.

Fourthly, we can find out which of the two methods gives the pupil the greater *power of attacking new material successfully*.

In so far as we can determine these issues, we are in possession of the clues which will lead us to reasonable conclusions on most, if not all, of the questions raised in this section and in the section entitled Introduction.

III. THE GENERAL PLAN OF THE EXPERIMENT.

One difficulty in all work of this kind is to find some unsophisticated material with which to experiment. I wanted to work, if I could, with some school children who had never learnt or even heard of a geometrical definition throughout the whole of their school lives. I think I succeeded in getting this condition fulfilled with some Standard V girls in the southwest of London and some Standard III boys in the southeast.

In the course of the experiment one of the boys' fathers told him it was Euclid, which it wasn't, and gave him one or two 'tips' which spoiled some of his papers; but, with that exception, nothing transpired to indicate that we were not working on virgin soil. I worked also with a Standard VI and VII girls' class in a poor school. These girls, though not *au courant* with geometrical definition, had nevertheless done much constructive work and were accustomed to express themselves freely and exactly. I worked also with a Standard VII boys' class in a poor school. There was a little difficulty here with one or two of the definitions, owing to the boys' attendance at the Manual Training Center, where they had learnt something about them. And, finally, the work was done with an ex-VII boys' class, the highest class of a higher grade school. The teacher of this class was

a man who had for years attended lectures on psychology, and was accustomed to teach very largely by inductive methods.

I think it will be conceded that we have here a good variety of material; and, since in every case we worked with all the pupils of the class, and not merely with selected pupils, it will also be conceded that if any tendencies show themselves throughout the whole range of our material, the probability that they are chance results is very small indeed. So much for a general survey of the material; it remains to be seen how it was utilized. Briefly, though there were local variations in procedure, the same general plan was followed throughout.

A whole class, under one teacher and working under the same syllabus of instruction, was divided into two equal groups. The children were required to attempt the definition of some geometrical shapes which were placed before them in the form of large drawings, and on the results of these attempts the class was divided. One of the two groups subsequently acquired the definitions inductively. The other group learnt them, but the children were instructed that the exact words given them in the definition were not required as long as they gave all the meaning. The two groups were tested immediately, and after shorter and longer intervals. And, after some interval of time, new material of a somewhat analogous kind was given to the children to define with a view to discovering which of the two groups could better apply their old knowledge, as we say, though it is, in some aspects, rather an application of a method than an application of knowledge.

Two of the classes were taken by me in the induc-

tive acquisition of the definitions; in the three other cases they were taken by their own teachers.

In one of the schools in which I took the exercises myself one of the other teachers inquired of the teacher of the class whether I wished the inductive method to succeed. "I think he does," was the reply. "Then," came the prompt rejoinder, "he ought to have let you take it; you would have got it into them much better than he would." With this unsolicited testimonial to my handling of the method, I proceed to a detailed description of the five series of experiments actually carried out in the schools above cited.

IV. FIRST SERIES OF EXPERIMENTS.

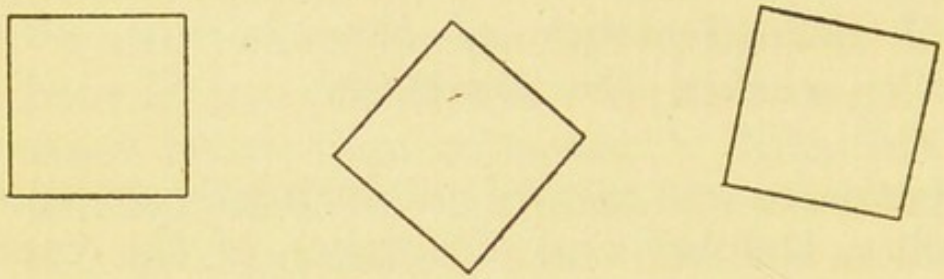
1. *General characteristics of the children who worked the exercises.*

This experiment was carried out during the months of September, October and November of the year 1911. The work was done with the whole of a Standard V class of girls of an average age of 11 years 8 months. The children knew nothing about geometrical definitions and were not biased by practice or novelty towards either of the methods employed. It is, of course, necessary to know the customary lines of teaching in a school in order to prevent one from drawing misleading conclusions from the results of experiments of this kind. The infant school work which these girls had done some four, five or six years previously was very little affected by tendencies to the kindergarten or sensory type of instruction. The school was situated in a neighborhood slightly above the average for municipal elementary schools, and the girls of the class were accustomed to give full attention to their school work.

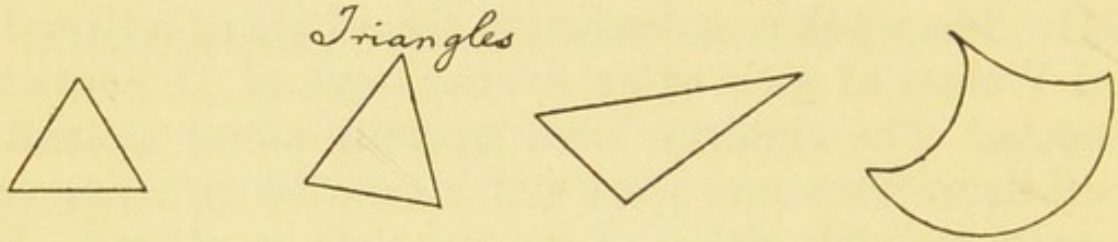
2. *The Preliminary Tests.*

Drawings of squares, triangles, oblongs and diameters of circles were placed upon large blackboards, with their names written above them, thus:

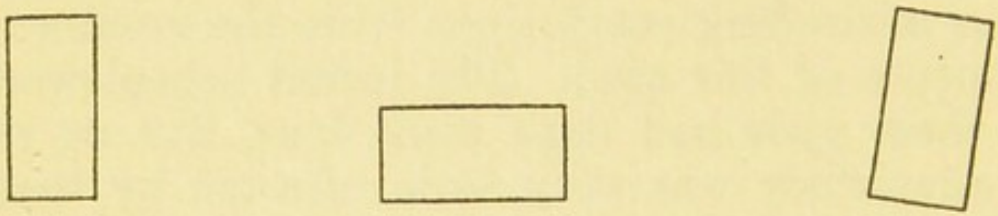
Squares



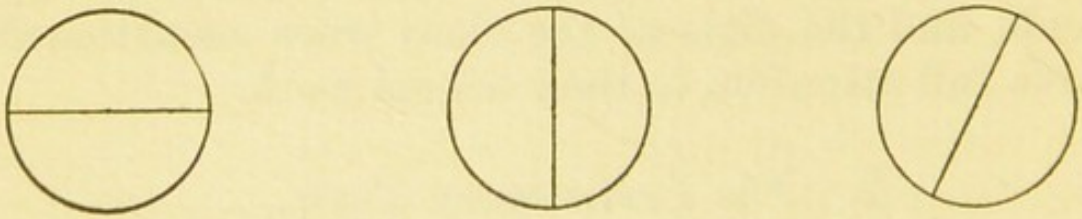
Triangles



Oblongs



Diameters of Circles



The children had the drawings pointed out to them, with the accompanying words, "These are squares," etc. In the case of the last set of figures the straight line was pointed to with the words, "This is a diameter of a circle, and this is a diameter of a circle," etc. Then upon a blackboard the following questions were written:

1. What is a square?
2. What is a triangle?
3. What is an oblong?
4. What is a diameter of a circle?

All the children in the class were required to answer the questions. They were told to do so thoughtfully and without hurry. No time limit was imposed in this or in any other of the subsequent exercises of this experiment.

3. *The Method of Marking the Preliminary Tests.*

The importance of this question merits a short discussion of the principles involved. I suppose one's first notion is something like this: Let us just take any currently accepted definition of a square, etc., Euclidean or other, and mark the children's exercises in accordance with their correspondence or non-correspondence with that. It will not matter very much which definition we take, provided that we keep to the same one all through the marking.

A perusal of the answers shows us immediately, if we had not known it before, that this method will not do. The children will hardly be likely, for example, to write down that a square is a rectilinear quadrilateral with a right angle, as one good definition gives it. They know the meaning of a straight

line; they know the meaning of four sides; they do not know the meaning of right angle, and if they did they would tell you quite candidly, if they had been properly taught, that the definition just given was wrong. A square has *four* right angles, not *one,* they would urge. The teacher might quite authoritatively inform them that every four-equal-sided straight-lined figure, if it have one right angle, *must* have four; therefore, why say four in the definition? and that a definition should not say more than it need. I hope the teacher won't, because the redundancy is no redundancy to the child at this stage; indeed, is no redundancy at all until the child is in a mental condition to deduce some of the properties from the others. Then, and only then, can one strike out the derivable properties and be satisfied with the others as sufficiently defining the term. This is an exceedingly interesting exercise, but its time is not yet. "Well," I can hear an impatient mathematician say, "this isn't mathematics; this is psychology; the children are not going to be marked on real definitions at all." On the contrary, I assert that they are going to be marked on the very realest of definitions; they will be marked according to the number of qualities and properties which they can themselves see to be common to all the specimens called by the same name. And I go further, and assert that the mathematician's definition, suitable and right for those who know the system of knowledge within which the definition finds a place, is just mere arbitrariness to those who do not. And, moreover, to tell the child that he *may* mention some of the common properties that he finds, and that he may *not* mention others, unless he is in a position to see for himself that some

are derivatives or derivable, is to shut him up sharp before a mystery. He won't do much spontaneous defining after that. Perhaps, some day, we may have a system of demonstrative geometry built up by psychological research. If so, the definitions will change as knowledge accumulates and reasoning becomes more penetrating, *just as our definitions do of the common objects of daily life.* Imagine a system of geometry for school children with evolving definitions. It is doubtless too horrible, and I do not, at present, ask my reader to accept such a thing, but only to grant that if I want to mark fairly the efforts of untaught children in spontaneous definition I must be guided by what they do, and not mark on an *à priori* scheme, settled beforehand by Euclid or by another geometer, or by myself, who am no geometer.

What, then, do the children say in answer to these questions, "What is a square?" etc. There are before us between forty and fifty sets of answers, and, though it would be illuminating for anyone to read the whole of them, they cannot be reproduced here. The interest in them lies in the fact that they represent untaught, spontaneous attempts; it is an interest which is at first psychological; the pedagogical interest comes later. Perhaps one of the best papers and one of the poorer ones may be found worthy of attention. I request the reader to look at the drawings whilst considering the children's definitions. The specimen papers follow exactly as they were written:

Nellie W., aged 12 years 3 months, wrote:

1. A square is an object with four corners and four lines, two for the sides and one for the top and bottom.
2. A triangle is an object with three corners, and three lines, two slanted ones and one straight one at the bottom.

3. An oblong is an object with four corners and four lines, the lines at the top and bottom being longer than those of the sides.
4. A diameter of a circle is a line drawn right across the circle from one side to the other each line is called a diameter.

Winnie T., aged 12 years 8 months, wrote:

1. A square is a kind of box with four lines all the same length.
2. A triangle is a thing with three sides not all the same length.
3. An oblong is a thing with two short sides and two long sides.
4. A diameter is a circle with a number of lines going from one side to the other.

Even from these two papers one may get a hint as to the way the children are going to sum the figures up, and a careful search through all the papers reveals that by one or another the following points of accurate definition are mentioned:

Common qualities mentioned in children's definitions.

1. A square is a shape, figure, drawing.
   It has lines or sides.
   It has four lines or sides.
   It has equal lines or sides.
   It has straight lines or sides.
   It has corners.
   It has four corners.
   It has equal corners.
   (A total of eight points.)
2. A triangle is a shape, figure, drawing.
   It has lines or sides.
   It has three lines or sides.
   It has corners.
   It has three corners.
   (A total of five points.)

3. An oblong is a shape, figure, drawing.
   It has lines or sides.
   It has four lines or sides.
   It has straight lines or sides.
   It has two long sides.
   It has two short sides.
   The two long sides are the same length.
   The two short sides are the same length.
   The two long sides are opposite each other.
   The two short sides are opposite each other.
   It has corners.
   It has four corners.
   The corners are all the same size.)
   (A total of thirteen points.)
4. A diameter of a circle is a line.
   It is a straight line.
   It goes through the middle or center of a circle.
   It goes from one side or edge of the circle to the other.
   (A total of four points.)

The papers were marked thus: One mark was allowed for each common attribute correctly given. It was decided not to allow thing or object as equivalent to shape, diagram, etc., for it was thought that 'thing' was so wide a term as to be hardly available for the purpose of these definitions, and that by the word 'object' children usually meant a material thing in three dimensions of space.

It is not contended that all the above units of marking are exactly equal in value; it is contended only that marking on these units is easy, practically serviceable, and yields results from which one can draw reliable conclusions for practical purposes.

On the results of this marking all the children of the class were divided into two equal groups, one of the best children being placed at the head of the first group—Group A—an equivalent child being placed at the head of the second group—Group B—then the children next in order were placed, one in each group, and so on down the list, carefully preserving the balance all the way down, till all the children were divided between the two groups.

There is a weakness here which needs attention. It is not usually satisfactory to divide a class *on the basis of one test only*. It is probably much more satisfactory where the higher mental functions are concerned than it is where simple sensory functions are measured, but it is risky even in the former case. It adds enormously to the probability of a valid result if several tests of the same kind are taken and the results of these correlated. One then feels confidence, if the results of the tests correlate highly with one another, that one is testing some function or group of functions which is operating steadily, and that each child is working at about its 'true form' as compared with the others. In some of the subsequent experiments of this kind I adopted that plan, but as I wished to be present during the whole time, and on each occasion when exercises were done in this case, I reduced the number of Preliminary Tests to one, but I did so with a full consciousness that I should feel less reliability in the equality of my two groups.

The marks obtained in the Preliminary Tests by the two groups, respectively, will be shown in the section headed 'Results.'

4. *The Teaching of the Two Groups.*

About a week later the two groups were separately instructed in the subject-matter of the definitions. As each child had already made some attempts for herself under test conditions to define the terms, she was in a favorable condition for the reception of knowledge by any method.

I had decided that one of the two groups should have the definitions written out for them, with illustrative drawings underneath, and that the children of this group should be instructed to study and learn the definitions. They were told that they were going to be examined afterwards, that they might then write down the exact words if they liked, but that as long as they got down all the meaning they would lose no marks because they had failed to remember the exact words. This resembled the way in which I remember myself to have learnt the Euclidean definitions. I studied the words and let my attention pass from words to figures to see the illustrations of the general statements in the definitions—it may fairly be called a deductive method. There is, however, one important difference. The definitions given to the children are such as they themselves are capable of making up. That does *not* mean that every child, nor even any child, in the class could say precisely all these things by itself; it does mean that they are on the lines of the child's own evolutionary track.

The Definitions as Learnt Deductively.

They were written down with illustrative examples drawn underneath the definitions. The drawings were the same as those given before in the Prelimi-

nary Tests; they will not be reproduced here. The definitions were worded thus:

1. A square is a shape with four sides and four corners. The sides are straight and they are all of the same length; the corners are all of the same size.

2. A triangle is a shape with three sides and three corners.*

3. An oblong is a shape with four straight sides and four corners. Two sides are long and two are short. The two long ones are opposite to each other and are of the same length, and the two short ones are opposite to each other and are of the same length. All the corners are of the same size.

4. A diameter of a circle is a straight line going through the middle of a circle from one side of the circle to the other.

The Definitions as Taught Inductively.

The teacher, myself in this case, having the points of each definition in mind, taught up to them, but no instruction was given by the teacher otherwise than by questioning.

In another school, in which also I did the necessary teaching myself, a discussion arose afterwards between the two groups of girls because one group had done rather better than the other in the subsequent testing. "You ought to do best," said a girl of the deductive group to the inductive group; "the gentleman taught you the definitions; we had to learn 'em for ourselves."

"That's just where you're wrong," she was an-

*It will be remembered that one of the triangles drawn had curvilinear sides.

swered; "the gentleman never told us a thing; we told him all about it."

But perhaps a little more explicitness may be useful to those who may wish to repeat the experiment. Let me illustrate by means of the last example:

Pointing to all the diameters drawn, the teacher says: "What can we say about all these?" The answer 'lines' will be received. He can then ask the question: "What is a diameter of a circle?" He will be answered, if he chooses his questionée well: "A diameter of a circle is a line." One feature of the method is that the teacher accepts all the answers given to him and translates them into concrete form. He draws a curved line on the blackboard, but not within one of the circles, and asks: "Is that a diameter of a circle?" He is answered: "No, because it is not a straight line." He draws a straight line, still outside the circle, and asks: "Is that right?" The answer comes: "No, because it's not in a circle." "Very well," the teacher says, "let us try again. What is a diameter of a circle?" If he chooses a child's answer, as he should, from among the least proficient of the class, he will be answered: "A diameter of a circle is a straight line inside a circle." He accepts the answer and draws a straight line in a circle which neither passes through the center nor touches the circumference on either extremity. He again asks: "Is that right?" He will be told: "No, because the line does not go to the edge on both sides." He corrects his drawing, producing the line each way to the circumference. He will now be told his line is wrong, because it must pass through the middle or center of the circle. He then draws a fresh line passing through the center, but cutting the cir-

cumference, and he is now told the line must reach the edges, but not pass over them. At this stage he can rely upon receiving a correct definition from the great majority of his pupils; but it is essential, if we are to keep this method distinct from the other one, that he does not ask a number of children to give the correct definition. One or two may do so, and the teacher then passes on; otherwise the mnemonic repetitive factor comes in here as in the other method, and for the purposes of this experiment it is usually desirable to avoid this.

I do not propose to go in detail over the method employed for teaching the remaining definitions. Any experienced teacher—and this paper is not written for other than experienced teachers—will be able to ask analogous questions and get the answers corrected in analogous ways. With children taught frequently on this method it is quite possible to get the necessary drawings and corrections, or most of them, done by members of the class, so that the machinery of correction and amplification is mainly in the hands of the class, with the teacher there to see fair play and direct the discussion to profitable issues. But I do not press this latter point; the work of concrete exemplification of error was, in the case of all the experiments about to be described, solely in the hands of the teacher. Teachers who would not agree with the method of mutual correction may quite well agree with this.

There are one or two points of detail, however, which may cause difficulty. It is of little use for the teacher to point to the squares drawn and ask, "What are all these?" for he will naturally be answered, "Squares." Indeed, the word squares is written

above the drawings. But if he points to the squares, and the triangles, and oblongs as well, and asks the same question, he will get answers like "drawings," "shapes," "figures," "diagrams." He can then start his detailed questioning to bring out the definition of a square. Among his answers will very likely be, "A square is a shape with four lines and four corners." It is obvious that many figures which are not squares can be drawn to comply with this definition, and the correction will proceed as before. The size of a corner is a difficulty for young children; they confuse corner with edge. It helps to ask, (pointing to angles of different sizes) "If these were the corners of a room, how much sand or how many blocks could I put in that corner, and in that one?" In some such way the size of the corners becomes thinkable to the young child. The question has been raised whether items of considerable difficulty, like this one, should not carry more than one mark. The theoretical justification is conceded, but it is argued that in practice a mark for each unit gives sufficiently steady and reliable results.

Children will often give a quality which is true of only one or two of the squares or triangles. It is only necessary to point to the other ones in these cases.

I need, perhaps, hardly say that these children, like the others, knew they were going to be examined immediately afterwards on the work they were then doing.

5. *The Immediate Testing of the Two Groups.*

As soon as the teaching of the Inductive Group was completed, the group which had been learning the definitions in another room also stopped their work;

and in a third room, so that the environment of both the groups should be changed, both sets of pupils answered the following questions:

1. "What is a square?"
2. "What is a triangle?"
3. "What is an oblong?"
4. "What is a diameter of a circle?"

6. *The Marking of the Tests.*

The papers were marked exactly as in the case of the Preliminary Tests, so far as the positive units were concerned, but a new feature was added to the marking. It is well known that bad teaching and bad learning produce errors, and errors of a noxious kind. But some statements that we sometimes call errors in the work of school children are rather irrelevances and redundancies than errors. For instance, when a child, in defining a square, after mentioning the properties of a square quite correctly, says: "and the corners are opposite each other," the statement is worth no positive mark, but neither is it worth any negative mark. Or when a child, in defining a triangle, says: "Some of the lines are curved and some are straight," though this, strictly speaking, is no part of the definition (which includes only the qualities common to all the triangles given), yet it can hardly be called a bad error. But it is a bad error for a child to say: "A triangle is a shape with three *equal* lines and three corners." Such an answer receives five positive marks—one each for shape, lines, three (lines), corners, three (corners). But 'equal' receives a negative mark as a 'bad error.' Again, when a child, in defining a square, says, amidst much

which is correct: "The corners all come under each other," the statement is marked as a bad error. Or when a child, speaking of a diameter, says: "It must stand upright," the statement is regarded as a bad error. The first diameter in the drawings was upright, hence the error. The definition was being elaborated from a memory of the one drawing without comparison with the memories of the others.

In the case of the experiment in this school, therefore, besides giving the positive marks obtained by each group, I shall also give the negative marks, and, in addition, the positive marks with the negative marks subtracted from them.

It is interesting, before turning to the section showing the results, to try to guess from our general knowledge of children's minds of the given ages (roughly from eleven to twelve), and our opinions as to the methods of teaching and learning, which of the two groups gained the more positive marks and which group made the more bad errors.

7. *Subsequent Testing of the Two Groups on the Same Subject-Matter.*

In discussions among teachers the question is frequently raised as to the relation between the quickness and the permanence of knowledge. Teachers are prone in theory to back the tortoise rather than the hare, though in practice they repeatedly prod the tortoise up. How far does the present experiment throw light on the matter? Are we justified in supposing that a test given to two groups of children, immediately after certain knowledge has been acquired, supplies us with comparative results which

will be true, say, a week later, a month later, and so on?

To test this point the exercise above described was repeated a week later. The children were not aware that they would ever have to do this work again.

Then, once again, more than a month after the first test (the exact chronology of the experiment will be given later on), the test was repeated a second time. The papers were marked in both these cases exactly as in the first test, positive marks being given for the points remembered and negative marks for the bad errors. The results will enable us to see how far the comparisons between the groups based upon the immediate results are corroborated when the results for *deferred* reproduction are taken into consideration. Again, it is worth while to try to think the answer out on general principles before turning to the actual results.

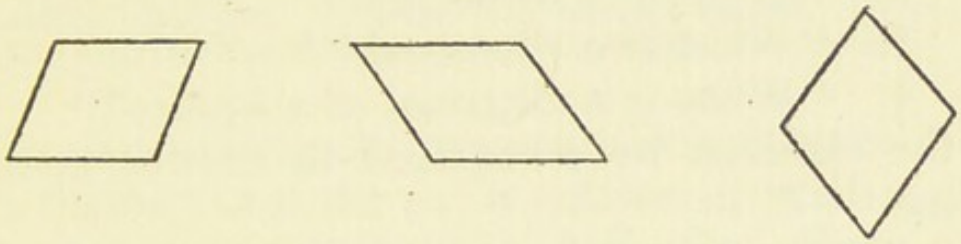
8. *The Testing of the Two Groups on New Material.*

It will be remembered that the children of the class were divided on the results of a test in which they were required to find definitions for themselves of a square, etc. How far has the teaching or learning by different methods affected their power to attack new material of an analogous kind? This is one of the most important questions that can be asked of any method of teaching or learning.

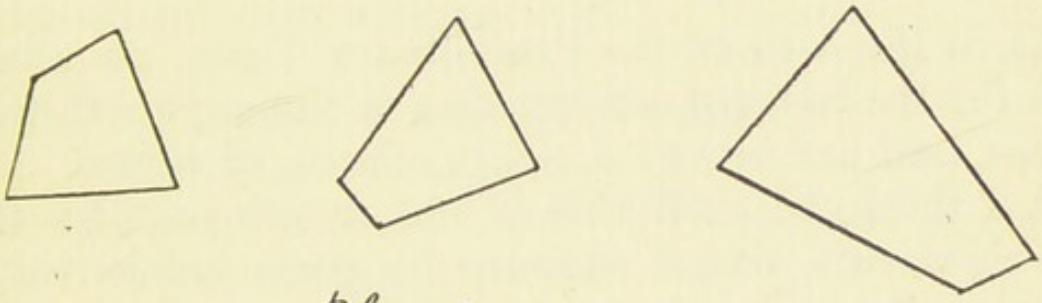
In ordinary pedagogical discussions it would be implied by assertions that children would be made more intelligent by one method than by another. To investigate this point experimentally the following tests were made:

Drawings were shown of rhombuses, etc., and their names written above them, thus:

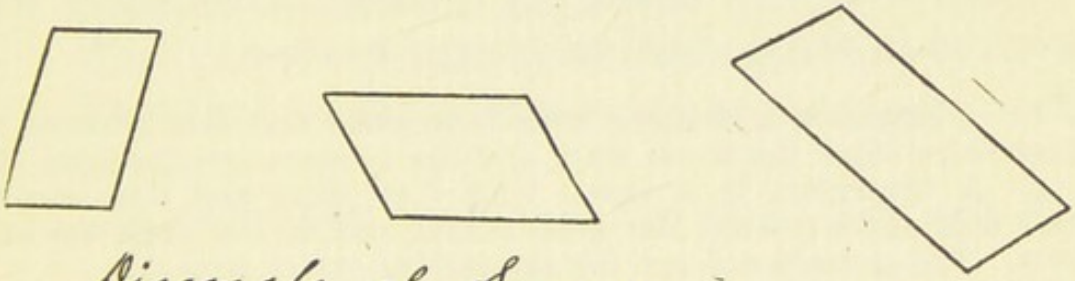
Rhombuses.



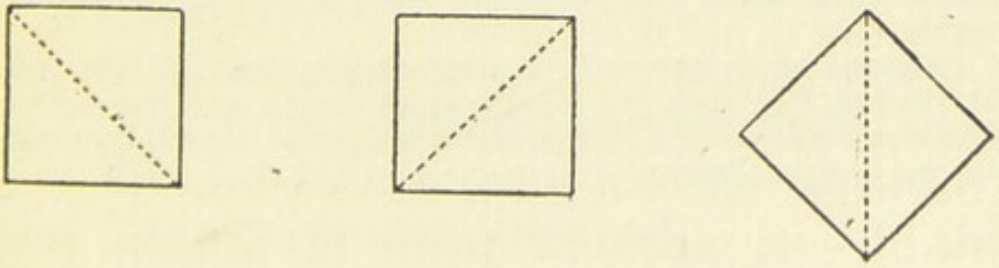
Trapeziums.



Rhomboids.



Diagonals of Squares
(the diagonals are drawn in dots).



In the drawings actually used the diagonals were continuous lines drawn in red.

Then the following questions were written on the blackboard:

1. "What is a rhombus?"
2. "What is a trapezium?"
3. "What is a rhomboid?"
4. "What is a diagonal of a square?"

and the children were required to answer them in writing.

9. *The Marking of the New Material.*

As in the case of the Preliminary Tests, we must look for the basis of our marking in the papers themselves, and not in any *à priori* scheme of values. It will, I think, be profitable if one or two samples of the children's actual answers be given before considering the units of marking which were adopted.

Laura B——, aged 13 years 11 months, who worked in the deductive group, wrote:

1. A rhombus is a shape with four sides and four corners, the four sides slant the same way, and the corners are the same size.
2. A trapezium is a shape with four sides and four corners. The four sides are not the same length and do not slope the same way. The corners are not the same size.
3. A rhomboid is a shape with four sides and four corners. It has two long sides and two short sides, the long ones are opposite each other, and the short ones are opposite each other, the corners are opposite each other and slant the same way, the corners are all the same size.
4. A diagonal of a square is a shape with four straight lines and four corners, the lines are all the same length and the corners are all the same size with a line going through the square from one corner to the opposite corner.

Rhoda T——, aged 12 years 11 months, who worked with the inductive group, wrote:

1. A rhombus is a shape. It is made up of four sides. The lines are straight, but are drawn slantling, and there are four corners joined exactly to one another.
2. A trapezium is a shape. It is made up of four sides and are

not the same length. The lines are straight but are drawn slantling. There are four corners, they join together exactly.

3. A rhomboid is a shape. It is made up of four sides two short sides being opposite, and the two long sides the same. The lines are straight but are drawn slantling. There are four corners they join exactly to one another.

4. A diagonal of a square is not a shape. It is a straight line, but drawn slantling. It joins two corners exactly opposite one another, and the line must not reach over the two corners.

Even from these two papers alone it would not be very difficult to make out a scheme of marking, and when taken in conjunction with the others, some forty or fifty in number, the following items of accurate description emerge:

1. A rhombus is a shape or figure, etc.
   It has sides or lines.
   It has four (sides).
   It has straight (sides).
   It has equal (sides).
   Two sides slant the same way (are parallel).
   The other two sides slant the same way.
   Two that slant the same way are opposite each other.
   The other two that slant the same way are opposite to each other.
   It has corners.
   There are four (corners).
   There are two big (corners).
   And there are two little (corners).
   The two big corners are opposite each other.
   The two little corners are opposite each other.
   The two big ones are equal.
   And the two little ones are equal.
   (A total of 17 points.)

2. A trapezium is a shape or figure, etc.
It has sides or lines.
It has four (sides).
It has straight (sides).
The sides are unequal.
It has corners.
There are four (corners).
The corners are unequal.
(A total of 8 points.)

3. A rhomboid is a figure or diagram or shape.
It has sides or lines.
There are four sides.
The sides are straight.
There are two long sides.
And there are two short sides.
The two long sides are equal.
And the two short sides are equal.
The two long sides are opposite each other.
And the two short sides are opposite each other.
The two long sides slant the same way.
And the two short sides slant the same way.
It has corners.
There are four (corners).
There are two big (corners).
And there are two little (corners).
The two big corners are equal.
The two little corners are equal.
The two big corners are opposite to each other.
The two little corners are opposite to each other.
(A total of 20 points.)

4. A diagonal of a square is a line.
   It is a straight line.
   It is drawn from one corner to another.
   The corner to which it is drawn is opposite the other.
   (A total of 4 points.)

If with this scheme of marking in view we turn to Laura B——'s paper, we shall see that she will receive 5 positive marks for her definition of a rhombus; but she has two 'bad errors.' The four sides of the rhombus do *not* slant the same way, and the corners are *not* the same size. Her definition of a trapezium receives 7 positive marks; there are no negative marks for bad errors in this definition.

It will now, doubtless, be quite easy for anyone with the aid of the table to assess the rest of the positive marks. But I might, perhaps, call attention to the fact that there are two more 'bad errors' in this paper. The corners of a rhomboid are *not* all the same size, and a diagonal of a square is *not* a shape.

Rhoda T——'s paper contains some errors in spelling, which, of course, are not counted in experiments of this kind. It can be quite easily marked on the system given above, and I think any teacher who marks it will agree that there are no 'bad errors.'

It must not be thought that children taught inductively make *no* bad errors when they apply their knowledge to new material. A comparison, however, between the number of bad errors involved in the use of the two methods will be found very useful later on. All the papers in all the tests and exercises in this school I marked myself, and the

marks were subsequently checked by the head mistress of the school.

10. *Chronology of the Experiment.*

All the tests and exercises, with the exception of the Preliminary Test, were taken on Tuesday mornings at 10.10 A. M. All instructions and teaching were given by myself, and I was present with the children during the whole time that each exercise was done.

The Preliminary Test for the division into two equal groups was worked on September 25th, 1911. The teaching and learning of the first set of definitions (which occupied 17 minutes for each group) and an immediate test in reproduction were done on October 3d. The second test in reproduction was given on October 10th. The test to see how far the children could apply their knowledge or method to new material was given on October 17th, and the last test on the first set of definitions—a further test in deferred reproduction—was given on November 7th. The lessons preceding the exercises were, in all cases, the same. For writing out what they knew no time limit was insisted on: the children were all allowed, nay encouraged, to put down as much as they could. A note was kept of the time taken on each occasion. In the test taken immediately after teaching and learning 25 minutes was the limit, and it was noticed that the Deductive Group, *i. e.*, those who had learnt the definitions, were much the quicker. The Inductive Group were, of course, to a large extent, making the definitions up as they went along. In the first test of deferred reproduction,

which took place a week later, 30 minutes were taken; in the test to show the power of application to new material, 32 minutes; and in the second test of deferred reproduction, which took place a month after the acquisition of the original definitions, 32 minutes were taken. It is possible to ascribe this lengthening of the time of the exercise to an increasing difficulty of remembrance. This may be a factor, but I am inclined to think there may be another; the children may be getting more thoughtful over the work, and consequently slower.

11. *Results.*

(*a*) *The Marks for the Preliminary Tests.*

Every mark which every child obtained in every exercise was carefully tabulated, though from the final table it was necessary to exclude three or four children who had been absent on several occasions. The first and last child in the Inductive Group were among these cases, so the corresponding children, namely, the first and last of the Deductive Group, were also omitted. There were then 21 cases left in each group.

In the Preliminary Test, in which the children tried by themselves to see what they could do in defining square, triangle, etc., the group which subsequently did the inductive work gained an average mark of 9.4, with a mean variation of 2.2; and the group which subsequently learnt the definitions deductively obtained an average mark of 9.5, with a mean variation of 2.3. The highest mark in each group was 15; the lowest mark was 6. In so far as it is possible to make a satisfactory division of a class on one test only the groups were well balanced.

(*b*) *The Marks for the Test Immediately After the Definitions Had Been Taught and Learnt.*

The Inductive Group gained an average mark of 22.6 of positive marks, with a mean variation of 2.5, whilst the Deductive Group gained an average mark of 25.6, with a mean variation of 2.2.* This is a clear and significant difference in favor of the Deductive Group—a difference which is accentuated when 'bad errors' are taken into account, for the former group makes 21 bad errors and the latter group only 12. Deducting the negative marks—the marks for bad errors—from the positive marks, it is found that the Inductive Group scores an average of 21.6 marks (mean variation 2.6), and the Deductive Group scores an average of 25.0 (mean variation 2.6)—a still clearer and more significant difference. It is highly probable from the above average marks and variabilities that this difference is a difference, so to speak, all along the line, *i. e.*, one which will be found between both the best pupils of each group and also between the worst. That relationship, however, will be shown more clearly by the following table:

Table I, showing the work of the Inductive and Deductive Groups compared, in the Preliminary Test, and in the First Test after the definitions had been learnt and taught (positive marks only).

	Group A (Inductive).			Group B (Deductive).		
Marks in preliminary tests.	No. of girls.	Av. mark in prelim. test.	Av. mark in first final test.	No. of girls.	Av. mark in prelim. test.	Av. mark in first final test.
12 and over.....	5	13.0	25.4	5	13.4	25.4
9, 10, 11.....	7	10.0	22.0	7	10.0	25.1
7, 8.....	5	7.8	21.4	5	7.6	26.6
6.....	4	6.0	21.7	4	6.0	25.5

*The difference between the averages is about five times the probable error of their difference, even on the assumption that the series are not positively correlated.

An inspection of Table I shows that, while there is no difference in the results between the two sections of the ablest children, those at the top of each group, there are considerable differences in the remaining sections. The argument of the teacher whose letter I have given, namely, that routine methods are better for the immediate reproduction of the actual material learnt, must be conceded for children of this level of ability. A teacher once asked me rather scornfully: "What did you expect from your 17 minutes' teaching?" Not much, perhaps, but we shall see more about that later. And, of course, there were also 17 minutes' learning, so the comparison was a fair one in any case.

(c) *The Marks for Tests of Deferred Reproduction.*

It is, however, one thing to answer questions immediately after one has learnt, or has been taught, the answers; perhaps it is quite another thing to give those answers accurately by and by.

The second Final Test took place one week after the first Final Test, and the children, as I have already pointed out, did not know they were ever to have the exercises again. It is a matter of great importance to the teacher, and it is a matter of great importance to the experimenter to know how far the immediate results from the work of any group of children may be taken as fairly representing what that group of children will do later on. And, perhaps, a week is too short a time. "The children won't have forgotten all about it by then," as a teacher said, so a third Final Test—the second test of deferred reproduction—was given a month

later. Let me present a few comparisons. I will give first the positive marks only:

Table II, showing the relation between immediate and deferred reproduction (positive marks only).

		Preliminary test.	First final test.	Second final test.	Third final test.
Inductive Group.	Average....	9.4	22.6	23.2	22.0
	M. V.....	2.2	2.5	2.0	3.1
Deductive Group.	Average....	9.5	25.6	25.4	24.9
	M. V.....	2.3	2.2	2.1	1.9

I will next show the immediate and deferred results for the two groups, when the negative marks for the Final Tests have been subtracted from the positive marks:

Table III, showing the relation between immediate and deferred reproduction.

		Preliminary test.	First final test.	Second final test.	Third final test.
Inductive Group.	Average....	9.4	21.6	22.2	21.0
	M. V.....	2.2	2.6	2.2	3.0
Deductive Group.	Average....	9.5	25.0	24.9	24.1
	M. V.....	2.3	2.6	2.6	2.1

The results from both the tables are in marked agreement. Both groups of children have gone down somewhat, but the children taught inductively have lost less of their original knowledge than the group which worked deductively. It appears that for groups of children of this age and ability the immediate results of these methods of learning and teaching may be accepted as indicating, comparatively, not only what the children can do at the time, but also what the groups will do by and by.

This may be shown roughly in the following table:

(*d*) *Correlation Between Immediate and Deferred Reproduction.*

Table IV, showing the correlation between immediate and deferred reproduction in the two groups, section by section (positive marks only).

	Group A (Inductive).				Group B (Deductive).			
Marks in preliminary test.	No. of girls.	Av. mark first final test.	Av. mark second final test.	Av. mark third final test.	No. of girls.	Av. mark first final test.	Av. mark second final test.	Av. mark third final test.
12 and over....	5	25.4	25.6	25.2	5	25.4	24.8	25.0
9, 10, 11.....	7	22.0	24.0	23.4	7	25.1	24.4	23.6
7, 8.....	5	21.4	20.8	19.0	5	26.6	26.2	25.2
6.....	4	21.7	21.7	19.0	4	25.5	27.0	26.2

It is fairly obvious that positive correlation exists between the children's immediate work, their work after one week's interval, and their work after a month's interval. But there is some irregularity here and there, so that it will be better to set down each child's individual results and work out the correlations between them rather than trust wholly to the inspection of the averages of these corresponding sections. The lists were arranged thus:

Inductive Group (positive marks only).

Name. (Initials only.)	First final test.	Second final test. A week later.	Third final test. A month later.
E. S.....	29	24	26
M. J.....	27	27	22
D. S.....	21	24	23
E. B.....	25	25	27
.....	..	..	..
E. W.....	20	20	17
.....	..	..	..
			(21 cases.)

A similar table was made of the individual results from all the children who worked in the Deductive Group, which also contained 21 cases. Then the Pearson coefficient of correlation, or 'r' formula,

which runs thus: $r = \frac{\Sigma xy}{n\sigma_1\sigma_2}$, was applied to the individual cases, and the following results were obtained:

	Inductive group.	Deductive group.
Correlation between results of first and second final tests.....	+.537	+.532
Correlation between results of first and third final tests.....	+.589	+.331

It certainly appears that whilst the averages for the groups have remained remarkably steady, the coefficients of correlation show that the individual children have changed places considerably. But we must, I think, admit that an immediate test of the result of a method of teaching or learning is one which gives us reasonable ground to expect, from the group as a whole, a similar result later on. That is, for *reproductive* exercises, where fairly homogeneous groups of school children are concerned, tests of immediate reproduction and tests of deferred reproduction give very similar results. We must apparently concede the point argued for in the head master's letter previously quoted, viz., that a mechanical method is better either for immediate or deferred reproduction at examinations. Let us concede that point, bearing in mind that our examination results, so far, have been only of a kind in which *the exact reproduction has been asked for of precisely what was taught or learnt.*

(*e*) *Results When the Two Groups Are Tested on New Material.*

We have seen, hitherto, that the Deductive Group has scored higher marks than the Inductive on every

occasion. Certainly our preliminary division of the two groups on one test only is open to criticism, and we might suppose that we had not really succeeded in obtaining equal groups. And possibly we have not, though I think it likely that they were approximately equal in initial capacity. Whether that be the case or not, there is found a definite difference between the two groups on the results of a test on new material, and in the opposite direction from the previous difference. The group taught inductively now leads the way. Counting only positive marks, the difference is small: the Inductive Group scores an average mark of 20.9 (mean variation 3.8), whilst the Deductive Group scores an average mark of 20.0 (mean variation 2.7). But when the marks for 'bad errors' are subtracted there is a distinct and decided difference: the Inductive Group scores an average mark of 19.4 (mean variation 3.5), whilst the Deductive Group scores an average mark of 17.7 (mean variation 2.5).* The more mechanical method has produced a much larger crop of 'bad errors,' for there is an average of 2.3 per child in this group, against an average of 1.2 in the other:

Table V, showing the 'bad errors' of the two groups.

	Inductive group.	Deductive group.
No. of girls with 5 bad errors	0	2
No. of girls with 4 bad errors	0	3
No. of girls with 3 bad errors	3	3
No. of girls with 2 bad errors	5	7
No. of girls with 1 bad error	6	4
No. of girls with 0 bad errors	7	2

*The difference between the means is more than twice its probable error, even on the assumption that the series are uncorrelated.

For children of this age and capacity, therefore, we are entitled to urge that the inductive method is much less provocative of error when 'new material' is given for test purposes. On the *old* material the Inductive Group made more errors than the Deductive in every test, so that we cannot suppose that the result with new material is consequent upon a greater initial tendency to 'howlers' in the Deductive Group; indeed, the tendency seems the other way, if there be one. We are often warned that averages are prone to conceal important differences between individuals, but I cannot expect huge tables of individual results to be printed, and, indeed, no useful conclusions in experiments of this kind could be drawn from individual results as such if they were printed; they must be grouped if they are to be of service. But I can show how far the superiority holds or fails for the sections of corresponding initial ability in the two groups:

Table VI, showing the work of the two groups compared, section by section, in the last reproductive test, and in the test of application to new material, the marks for bad errors being subtracted from the positive marks.

	Group A (Inductive).			Group B (Deductive).		
Marks in preliminary test.	No. of girls.	Av. mark last reproductive test.	Av. mark application to new material.	No. of girls.	Av. mark last reproductive test.	Av. mark application to new material.
12 and over..	5	23.8	20.4	5	24.4	17.6
9, 10, 11.....	7	22.8	21.7	7	22.9	17.0
7, 8.....	5	18.6	17.4	5	25.2	19.0
6.....	4	18.2	18.2	4	25.7	18.5

In the case of every corresponding section the loss, when application is made to new material, is greater in the Deductive Group than in the Inductive Group.

12. *Pedagogical Conclusions.*

I have previously pointed out that, as a whole, the Inductive Group gains higher marks for application to new material than the Deductive. What pedagogical conclusions may we draw from this? Two outstanding conclusions seem to me to follow from this work: the first relating to teaching, the second relating to examination.

First, as to the method in teaching: I suppose no teacher would desire us nowadays to favor a method *merely* because it enabled us to produce a better result in the exact reproduction of what had been taught or learnt. Let us consider exact reproduction as so much to the good, but let us also remember that individual lessons form, or should form, part of a course, and the method which enables a pupil to make the best attack on new analogous material is, one may reasonably suppose, likely to emerge triumphant at the end of the course. Such a method does really train 'intelligence,' in the best sense of that much-abused word.

Secondly, as to method in examination. Whenever examiners set work to be done which is a mere reproduction of what the children have been taught or have learnt, they are favoring the mechanical method rather than the inductive one.

I do not say there is no place for examination of that sort, but high assessments for teaching should never be given on such a basis. The supreme test of good teaching is the power, on the part of the pupils, to attack 'new material.' One word, however, of caution. I do not mean material wholly new, as many psychological tests are. By the use of such tests as

those we are measuring natural ability rather than the result of pedagogical work. The material should be '*new*,' but it should be *analogous* to the work which the pupil may reasonably be expected to have done before.

I venture to suggest that examinations of this kind would raise the tone and method of teaching rather than, as too often has been and is the case, tend to depress them. I wish to exempt junior scholarship examinations. They should, in my judgment, be psychological in the sense given above.

But, perhaps, the reader may feel that I am building up a huge structure of theory on the basis of a very little experiment; so I will turn to the second series of experiments in this research and show how far the facts and conclusions resemble those of the first.

V. SECOND SERIES OF EXPERIMENTS.

1. *General Plan.*

As before, a whole class of children, under the same teacher, studying the same curriculum in accordance with the same time-table of school work, was divided into two equal groups on the basis of preliminary tests in geometrical definition, which the children attempted, untaught and unaided. Then the pupils of one group were taught inductively how to arrive at the definitions, whilst the other group learnt the definitions deductively. An immediate test was given to show which method was the better for the purposes of immediate reproduction, and subsequent tests were given on the same subject-matter to indicate which of the two groups was the more successful in deferred reproduction. Also, as before, a test was given on new analogous material to see which of the two methods showed the greater 'transfer' effect. One or two outstanding points of difference between the conditions of this experiment and those of the first experiment may, perhaps, usefully be mentioned here before proceeding to the details. The children who did the work were, as before, girls belonging to an elementary school in London. But in this case they were of a poorer social class; they were older than the girls in the previous school, their average age amounting to rather more than 13 years,

and, measured by school standards, they were more proficient mentally, for these children were graded as Standard VI, a, and VII, whereas, in the previous school, the children were graded as Standard V. The children of this class had done a great deal of manual constructive work, and were taught by a teacher from whom they had learnt to express their meaning in direct, simple language. Like the girls of the previous school, they knew nothing of geometrical definition; and 'Demonstrative Geometry,' or even 'Euclid,' were terms of no meaning either to themselves or to their parents.

There was, too, an important difference in the early part of the procedure of this experiment from that of the previous one. Instead of dividing the children on the basis of one test, *four* tests were given, in which the same material was employed throughout. It was believed that the division into equal groups would be much more satisfactory if it were effected on a wider basis than the results of one test; and, indeed, the greater regularity of the subsequent work showed that to be the case. Thirty-seven children started the experiment, but unavoidable absenses from school reduced the number available for the tabulation in 'equal groups' to 34.

2. *The Preliminary Tests and the Method of Marking.*

As in the previous case, drawings of squares, triangles, oblongs, and diameters of circles, with their names appended, were shown to the children, and the questions, "What is a square?" etc., were written on the blackboard. The units of marking, as before,

were obtained from a careful study of the answers actually given by the children.

One or two instances of the children's spontaneous attempts at definition may possess psychological interest.

Lily H——, aged 13 years 6 months, a girl graded as Standard VI, a, wrote:

1. A square is four lines each of the same length all joining one another and when they are joined they form a square with four angles the square may be straight up or slanting.
2. A triangle is a three line drawing, joining each at the ends and when it is drawn it forms a drawing with three angles each of the same size.
3. An oblong is a figure with four lines same as the square only there are two long lines and two of them are short lines with four angles of the same size.
4. A diameter is the line drawn through a circle to separate one half from the other only it must be drawn through the centre.

The above is one of the better papers which were worked during the Preliminary Tests, but it is certainly not the best. Let me now give an inferior one.

Ada B——, aged 13 years 1 month, also graded as Standard VI, a, wrote:

1. A square is a thing with four equal sides. A square can be all different shapes as long as the four sides are equal.
2. A triangle is something which has three sides and the sides must be as long as each other.
3. An oblong has four sides, two of the lines are short and two are long. The two long lines must face each other, and the short ones must be the same length as each other.
4. A diameter of a circle is a round ring divided into half.

It was seen, after a careful perusal of the children's papers, that the units of marking which had been worked out for use in the previous school were also quite suitable for this one. Turning to the answers of Lily H., and marking with these units, we find that in her first answer she has a mark for

'lines,' another for 'four' (lines), a third for 'of the same length,' a fourth for 'angles,' and a fifth for 'four' (angles). The lines do not exactly all join one another, but the statement was not considered an error; and the further statement, 'the square may be straight up or slanting' was considered irrelevant. In the second answer Lily H. receives a mark for 'drawing,' one for 'lines,' one for 'three' (lines), one for 'angles,' and one for 'three' (angles); a total of five marks. She describes the angles as being all of the same size, which is certainly a serious error, and was probably due to the confinement of her attention to the equilateral triangle, which was *one* of the triangular figures shown. For her definition of oblong she receives a mark for 'figure,' another for 'lines,' a third for 'four' (lines), a fourth for 'two long lines,' a fifth for 'two short lines,' and a sixth and seventh for 'angles' and 'four' (angles). For her fourth definition, namely, that of a diameter of a circle, she receives a mark for 'line,' and another for 'drawn through the center.' Lily H. thus receives a total of 19 positive marks.

Ada B——'s paper, marked in the same way, receives a total of 11 positive marks. Her third answer—the definition of an oblong—is quite unexpectedly good, considering the weakness of her definitions of 'triangle' and 'diameter.' With various triangles, mostly scalene, exposed before her eyes, it was certainly a 'bad error' to say "the sides must be as long as each other." Probably, with these explanations, the method of marking adopted will be readily applicable. I will now set out the chronology of the whole experiment.

3. *Chronology of the Experiment.*

First of all, a preparatory exercise was given at 9.40 A. M. on Friday, October 20th, 1911, under test conditions, to accustom the children to work of this kind, which was quite new to them. Then, on Tuesday, October 24th, Wednesday, October 25th, and Friday, October 27th, at 9.40 in the morning, immediately after Scripture lesson, a second, third and fourth Preliminary Test were given. On the results of the second, third and fourth tests the class was divided into two equal groups. Then on Friday, November 3d, at the same time in the morning, the pupils of Group B were taught the definitions inductively by me in the way previously explained, whilst Group A learnt the definitions by studying them as written out, and referring to the illustrative figures drawn beneath the written definitions. I took care that the children who were learning the definitions should receive all their instructions from me, and informed the children of both groups that they would be required to answer questions immediately afterwards about what they were learning or being taught, respectively. The children who had been trying for themselves without help for four exercises to see what they could do in the way of spontaneous definition were, of course, in a state of receptivity for instruction of either kind, inductive or deductive. Their marks had risen steadily day by day, so that they were still in the 'improving' stage for this work. The teaching lasted 17 minutes, and, of course, the same time was allotted to the study of the written definitions. The two groups were put together immediately, and the old questions: "What is a square?"

etc., were written on the blackboard. It was noticed that the children working in the deductive group had answered their questions some three to five minutes sooner than the girls in the inductive group.

Four days later, on Tuesday, November 7th, a second test of precisely similar nature was given to both groups at 9.40 A. M., after Scripture lesson, as before.

On Friday, November 10th, at the same time in the morning, and after the same lesson as on previous occasions, a test was given to both groups on new analogous matter to test the comparative 'transfer' values of the two methods of learning.

The last test in the series was given at the same time in the morning, and after Scripture, as before, on Friday, December 1st. In this test the previous questions on the material actually studied were given again—"What is a square?" etc. The object of this test—the second test in deferred reproduction—was to discover, if possible, which of the two groups had lost the more after a considerable interval; in other words, which method of teaching or learning favored the more permanent retention. I ought to say that, with the exception of the test given immediately after the teaching and learning, the children were not aware that they were going to do any of these tests before they were actually set to do them.

4. *The Final Tests and the Method of Marking.*

Three of the Final Tests were repetitions of the Preliminary Tests, and the same method of marking was adopted in them as in the Preliminary Tests. These were the tests given immediately after teach-

ing and learning and the two tests of deferred reproduction. The remaining test, in which the children, without further teaching, were required to attack new material, was identical with the corresponding test given in the previous school, and it was marked in the same way. Drawings of rhombuses, trapeziums, rhomboids and diagonals of squares, with their names appended, were shown to the children, and they were required to answer in writing the questions: "What is a rhombus?" etc.

5. *Results of the Experiments.*

(*a*) *Results of the Preliminary Tests.*

The marks obtained in the Preparatory Test will not be given; it was noticed that though the children at the top and bottom of the lists remained much the same, a considerable number of children changed places from the Preparatory to the first Preliminary Test. As it was very important that the work of the children should be 'steady' before the class was divided into two equal groups, two more tests—the second and third Preliminary—were given and the results correlated. The work started with 37 children, but two had been absent during the tests, so they were excluded from the lists.

The results of the three Preliminary Tests are shown compendiously in the following table:

Table VII, showing the correlation between the results of the First, Second and Third Preliminary Tests.

Marks in first preliminary test.	No. of children.	Av'age Marks in Preliminary Tests. First test.	Second test.	Third test.
19, 18	7	18.3	18.7	19.0
17, 16	5	16.4	18.4	17.6
15, 14	9	14.7	16.4	17.0
13, 12	7	12.7	13.9	14.7
11 to 8	7	9.6	10.1	12.3

It is obvious from Table VII that high positive correlation exists between the results of the first, second and third Preliminary Tests, and that we are measuring a mental function, or group of mental functions, which is working very steadily. A precise numerical value for the coefficient of correlation has been worked out from the individual cases by means of the Pearson formula $r = \frac{\Sigma xy}{n\sigma_1\sigma_2}$, and 'r' has been found to be + .80, with a probable error of .04, for Tests 1 and 2, and + .77 (probable error .05) for Tests 2 and 3. It seems very likely that a division of the class into two equal groups on the basis of such regular results as these will be satisfactorily effected.

The children were divided into two groups of 17 girls each, thus (N. G., the girl at the bottom of the list, was omitted):

Table VIII, showing the division into two equal groups.

Group A.

Name. (Initials only.)	Marks for Preliminary Tests. First test.	Second test.	Third test.	Total marks.
W. F.	18	24	21	63
H. L.	18	20	19	57
H. G.	17	20	17	54
.....	..	..	..	..
G. E.	11	11	11	33
T. F.	8	9	13	30
Averages	14.2	15.8	16.3	46.3
M. V.'s	2.3	3.0	2.6	

Group B.

Name. (Initials only.)	Marks for Preliminary Tests. First test.	Second test.	Third test.	Total marks.
D. A.	17	22	22	61
W. E.	19	20	19	58
H. L.	18	17	20	55
.....	..	..	..	..
A. R.	10	10	11	31
L. L.	11	7	13	31
Averages	14.5	15.5	16.3	46.3
M. V.'s	2.6	2.8	2.2	

Care was taken also that the children should be so arranged in the grouping that the ages of the one group should very closely approximate to those of the other. The average age of Group A worked out to 13 years 1 month (mean variation 7.2 months), and of Group B to 13 years 0.5 months (mean variation 5.6 months).

(*b*) *Results of the Tests in Immediate and Deferred Reproduction.*

It now remains to be shown what marks were obtained after the one group had been taught the definitions and the other group had learnt them. The total marks will be shown for the three Preliminary Tests, with the marks for immediate reproduction and for the two tests of deferred reproduction—the one given a few days later and the other about a month later than the test of immediate reproduction:

Table IX, showing the work of the Inductive and Deductive Groups compared, section by section, in the Preliminary Tests and in the Tests of Reproduction (positive marks only).

Group A (Deductive).

Marks for three preliminary tests.	No. of children.	Average Marks: Preliminary tests.	First reproduction.	Second reproduction.	Third reproduction.
Over 50.....	6	18.4	26.8	26.5	26.0
40 to 50.....	6	15.6	27.7	26.5	25.0
30 to 40.....	5	11.7	25.0	22.6	23.0

Group B (Inductive).

Marks for three preliminary tests.	No. of children.	Average Marks: Preliminary tests.	First reproduction.	Second reproduction.	Third reproduction.
Over 50.....	6	18.4	28.8	28.3	26.8
40 to 50.....	6	15.4	26.5	26.2	25.2
30 to 40.....	5	11.7	25.8	24.8	25.2

It seems clear that in this case the children taught inductively were just as successful as those taught deductively, even in immediate reproduction, and that after a month's interval they were rather more so; they had lost less of what they had been taught. This will, perhaps, be shown more clearly in the following tables:

Table X, showing the work of the two groups compared, in the Preliminary Tests, and in the Tests of Immediate and Deferred Reproduction (positive marks only).

	Average mark for three preliminary tests.	Average Marks: First reproduction.	Second reproduction.	Third reproduction.
Inductive group.....	15.4	27.1	26.5	25.8
M. V.'s.....	2.4	1.8	1.9	2.2
Deductive group.....	15.4	26.6	25.4	24.8
M. V.'s.....	2.4	1.9	2.2	2.4

Table XI, showing the work of the two groups compared in the Preliminary Tests and in the Tests of Immediate and Deferred Reproduction (when the negative marks have been subtracted from the positive marks).

		Average Marks.		
	Average mark for three preliminary tests.	First reproduction.	Second reproduction.	Third reproduction.
Inductive group.....	15.4	26.8	26.2	25.2
M. V.'s.....	2.4	1.9	1.8	2.3
Deductive Group.....	15.4	26.3	25.0	23.9
M. V.'s.....	2.4	1.8	2.3	2.0

The balance of advantage seems even more clearly on the side of the group taught inductively.*

(c) *Correlation Between Immediate and Deferred Reproduction.*

It seems likely from Table IX, already given, that there is considerable positive correlation between the results of immediate reproduction and those of deferred reproduction. That is to say, the girls who are best immediately after teaching and learning are also the best after an interval, and those who are worst immediately after teaching and learning remain the worst after some time has elapsed. But in Table IX the children are classified on the basis of their marks for the preliminary tests, and this classification tends to obscure much of the correlation which undoubtedly exists. In the following tables the classification is based on the marks obtained in the test of immediate reproduction:

---

*In Tables X and XI the difference between the means of the work of the two groups in deferred reproduction is about twice the 'probable error' in each case, even on the assumption that the series are not positively correlated.

Table XII, showing the results for Immediate and Deferred Reproduction compared, of the Inductive and Deductive Groups (positive marks only).

Deductive Group.

Marks in immediate reproduction.	No. of girls.	Average Marks in Reproduction. First.	Second.	Third.
Over 28	3	29.3	27.3	26.0
28	5	28.0	25.8	24.8
27, 26	5	26.2	26.0	24.6
Below 26	4	23.3	22.5	24.0

Inductive Group.

Marks in immediate reproduction.	No. of girls.	Average Marks in Reproduction. First.	Second.	Third.
Over 28	5	29.4	28.8	27.0
28	4	28.0	27.8	27.3
27, 26	5	26.6	25.0	25.2
Below 26	3	23.0	23.7	22.7

It is quite obvious, from the foregoing table, that considerable positive correlation exists between immediate and deferred reproduction, but such a table gives us no numerical equivalent for correlation. The correlation coefficients have, however, been worked out, and for the Deductive Group the coefficient for the first and second reproduction is + .62 (probable error .10), and for the second and third is + .58 (probable error .11).

In the Inductive Group the correlation coefficient between the first and second reproductions is + .76 (probable error .07), and between the second and third is + .76 (probable error .07). All the figures indicate high reliability for the results, and a comparison of the correlation coefficients for the Inductive and Deductive Groups shows the work of the former to be the more consistent.

(*d*) *Results of the Test on New Material.*

In the case of the previous school we found that with younger children of a lower standard the deductive method seemed the better for purely reproductive purposes. In this school the inductive method seems better, even for purposes of reproduction. We have now to see whether, when application is made to new material, the results for these children agree with or differ from those of the preceding school. First let me show the results for the two groups as wholes:

Table XIII, showing the work of the two groups compared in the Preliminary Tests and in the Tests of Application to New Material.

		Average Marks for New Material.	
	Average mark for three preliminary tests.	Positive marks only.	When negative marks have been subtracted.
Inductive group.....	15.4	25.5	24.2
M. V.'s.....	2.4	2.7	3.4
Deductive group.....	15.4	23.3	21.9
M. V.'s.....	2.4	2.9	2.9

We have a clear advantage, in both cases, on the side of the Inductive Group. The difference between the averages amounts to about three times its 'probable error,' even on the assumption that the series are not positively correlated. Once again, then, we find the inductive method triumphant when application is made to new material. Let me now show how far this is a difference which is to be found all along the line, *i. e.*, for the weaker as well as for the abler pupils:

Table XIV, showing the work of the two groups compared, section by section, in the Preliminary Tests and in the Test of Application to New Material (positive marks, and positive marks after deduction of the negative marks).

Marks for three preliminary tests.	Group A (Deductively Taught).			Group B (Inductively Taught).		
	No. of girls.	Marks for New Material. (Positive only).	Marks for New Material. (After deduction).	No. of girls.	Marks for New Material. (Positive only).	Marks for New Material. (After deduction).
Over 50	6	25.8	24.5	6	27.3	26.0
40 to 50	6	23.7	21.7	6	24.2	22.7
30 to 40	5	19.8	19.2	5	25.0	24.0

There seems little doubt that the group inductively taught shows a superiority which is general—a superiority which, somewhat unexpectedly to me, however, seems most clearly marked in the weakest (initially considered) of the three sections into which each group is divided.

VI. THIRD SERIES OF EXPERIMENTS.

1. *General Plan.*

As in the previous experiments, a whole class of children, working under the same teacher, with the same curriculum, and according to the same timetable of work, was divided into two equal groups on the basis of several tests in geometrical definition, which the children attempted without instruction and without help. Then, subsequently, one of the two groups was taught inductively and the other group learnt the definitions. There were tests of immediate reproduction immediately after the lesson, and another test, which might also be called a test of immediate reproduction, on the following day. About a week later there was a test of application to new material, and, three weeks after this, two further tests were given, which will be referred to as tests of deferred reproduction.

The work was done with fifty children, whose average age was 9 years 3 months. They were graded as Standard III of a municipal elementary school for boys, situated in a very good suburban neighborhood in the southeast of London. The inductive teaching was done in this case not by me, but by the teacher of the class; whilst the group which studied the written definitions was taken, during that particular lesson, by the head master of the school. All the tests were

administered by the class teacher, who had had some experience of research work in biology as well as in experimental pedagogy. One of the boys' fathers told him he was doing Euclid (which he wasn't), and gave him a 'tip' or two which affected some of his papers adversely; but, with that exception, the success of the experiment was not hindered by any previous knowledge on the part of the children. Whereas, with the Standard V class of girls, in the experiment just described, the teacher's methods were instructional rather than either definitely inductive, deductive, or memoriter, and with the Standards VI and VII class of girls, in the experiment which has just been described, the teacher's methods were both inductive and memoriter, according to the subject-matter dealt with; in this third case the reaction against unintelligent teaching had gone so far that, whilst the inductive teaching was extremely good, the memoriter work was decidedly novel to the children. Novelty has a stimulating influence, we all know, but it is unlikely that its influence is more effective in result than that of habitual practices. In any case it is essential to try the experiment with classes differently taught.

2. *The Preliminary Tests and the Method of Marking.*

Just as before, drawings of squares, triangles, oblongs and diameters of circles, with their names written against the drawings, were shown to the children; the questions, "What is a square?" etc., were written on the blackboard; the children were told to look at

the squares, triangles, etc., and to answer the questions in writing as well as they could.

The units of marking, as before, were obtained after a careful review of the answers actually given, and it was found that the units previously adopted were quite suitable. A few instances of the children's attempts at spontaneous definition may be worth quoting. It must be remembered that these children were considerably younger than either of the classes of girls whose work has previously been described, and that they were graded as Standard III as compared with Standards V, VI and VII. On the other hand, the school was much more favorably situated socially than either of the schools for girls. Moreover, it was a boys' school; and boys, whether through greater natural ability or more training, are more proficient, geometrically, than girls.

R. D., aged 9 years 1 month, wrote:

1. A square is a four sided figure with four points and the sides are all equal.
2. A triangle is a three sided figure with three points and the sides equal.
3. An oblong is a four sided figure with two sides long and two sides short.
4. A diameter is a strait line that goes anything like a circle and will go across any way.

If we mark this paper—R. D.'s first preliminary test—on the system of marking adopted in the previous experiments,* we see that for his definition of a square he receives one mark for 'figure,' one for the adjective 'sided,' one for the numerical adjective 'four,' and one for the equality of the sides. 'Points' are taken as equivalent to angles or corners, and

*The reader is recommended to turn to page 28 for the list of units.

therefore receives a mark, whilst the numerical adjective 'four' also scores. This gives a total of six marks for the definition of the square.

The definition of triangle receives one mark for 'figure,' one mark for 'sided,' one mark for 'three,' one for 'points,' and another for 'three' (points). 'The sides equal' receives a mark as a 'bad error,' but there were so few of these in the preliminary tests that they were not tabulated.

The boy's definition of oblong receives a mark for 'figure,' one for 'sided,' one for 'four,' one for 'two sides long,' and another for 'two sides short.'

His last definition is rather weak. He obtains a mark for 'line' and one for 'strait,' and that is all.

When one remembers that these are untaught, spontaneous definitions given by a boy 9 years of age, we shall, I am sure, regard them as affording evidence of considerable ability. Four times the boys answered these questions without help and without criticism, and advanced a little each time. This is what R. D. wrote on his fourth attempt—the fourth preliminary test—three days after the first:

1. A square is a four sided figure with four equal sides and four sharp points.

2. A triangle is a three sided figure with three equal sides to it and it has three sharp points.

3. An oblong is a four sided figure with four points but the sides are not all the same two sides one length and the other sides another length.

4. The diameter of a circle is a line that is going from one side to the other side of the circle and that is called the diameter of a circle and the line is quite strait.

Let us see how far this fourth paper is in advance of the first. The definition of a square receives the same mark as before; it is slightly more concise in

expression, but the units of correct description are the same in number in both cases.

The definition of triangle remains unaltered.

It is interesting and important to notice that even a clever boy may go on perpetrating a 'bad error' unless his attention is drawn to it, which, of course, the conditions of the experiment did not permit us to do in these tests.

That R. D. is clever for a nine-year-old boy is clearer from his next two definitions than from those of the square and triangle. He nearly doubles his previous mark for his definition of an oblong. He now receives marks for 'points,' for 'four' (points), for 'two sides long' and for 'two sides short,' and also for 'two long sides equal' and 'two short sides equal.'

His definition of diameter has also improved. He has now included the condition that it must go from one side of the circle to the other.

These papers of R. D. are extremely good ones, and do not represent the average mark of the class, which ranges from 11 to 13 units, rather than from 18 to 23, which R. D. obtains for his first and fourth papers, respectively.

Let me now give examples of some papers below the average.

J. C., aged 9 years 2 months, answered his first preliminary test as follows:

1. A square is four put into one shape with equal sides.
2. A triangle is a thing that has no equal sides, two are equal and one is not, and it has three sides.
3. An oblong is not a square, but it is a long one.
4. The diameter is a line drawn through the midle of a circle.

Side by side with this—J. C.'s first preliminary

test—let us compare the paper worked by him three days later—his fourth preliminary test:

1. A square is a shape of a block with four equal sides.
2. A triangle is a long square with only three sides, the two side ones are both the same and the top one is not.
3. An oblong is a square that is long, with two equal sides and two ends which are not the same size.
4. The diameter of a circle is a line drawn down the midle.

The marks for the definition of a square are in both cases the same: 'shape' receives a mark, 'sides' receives one, 'four' gets one, and 'equal' (sides) gets one.

The two definitions of a triangle receive the same mark: there is a mark for 'sides' and one for 'three' (sides), and that is all.

The first definition of an oblong receives no marks at all, whilst the one given later receives a mark for 'sides,' a mark for 'two equal' (sides), and one for "two ends which are not the same size as the others."

His definition of a diameter remains unchanged throughout the preliminary tests; in each case it receives two marks only, one for 'line' and one for 'drawn through the middle.'

J. C. advances from a mark of 8 in the first test to 11 in the fourth.*

Having given some indications of the work done in the Preliminary Tests, on the results of which the class was divided into two equal groups, let me set out in detail the chronological progress of the experiment.

---

*The average improvement from test to test is shown on page 91.

3. *Chronology of the Experiment.*

A first Preliminary Test was given at 9.40 A. M. on Monday, October 23, 1911, immediately after Scripture lesson; a second on Tuesday, October 24, a third on Wednesday, October 25, and a fourth on Thursday, October 26, at the same hour and after the same lesson on each occasion. On the results of these four tests the class was divided into two equal groups.

On Thursday, November 9, at 9.40 A. M., one of the two groups was taught the definitions inductively by the methods already described, whilst the other studied them as written out, with reference to the drawings appended to the verbal descriptions. Twenty-three minutes were taken by the teacher to teach the definitions inductively; the same time was allowed to the group which was studying the definitions with a view to remembering them. Both groups of children were aware that they were to be tested on their work at the close of the lesson. Accordingly, at 10.15 A. M., a test was given in immediate reproduction. In this school, since the children were young and the exercises very novel, we thought it best to take another test, identical with the test of immediate reproduction, at the same hour on the next day, Friday, November 10, to see how far the first day's test was reliable. These two tests will be referred to as the two Tests of Immediate Reproduction.

At 9.40 A. M. on Thursday, November 16, a test was given on the application of what had been learnt to new analogous material with the object of discov-

ering which of the two groups attacked the new material the more successfully.

Finally, two tests of deferred reproduction were given at 9.40 A. M. on Thursday, December 7, and Friday, December 8. The children were quite unaware that they would be required to take any of these tests, with the exception of the one immediately after the teaching and learning on Thursday, November 9.

4. *The Final Tests and the Method of Marking.*

The two tests of Immediate Reproduction were repetitions of the Preliminary Tests, as were also the two tests of Deferred Reproduction. The units of marking previously used in the Preliminary Tests were found quite satisfactory. The tests of deferred reproduction received negative as well as positive marks. One or two specimens of the worked papers may be of interest.

L. O., aged 9 years, who scored 13, 16, 18, 18 in his preliminary tests, and was taught inductively, for his first test of Immediate Reproduction on November 9, wrote as follows:

1. A square is a shape with four lines all the same size and for corners all the same size.
2. A triangle is a shape with three corners and three lines.
3. A oblong is a shape with two long lines the same size, and two shorter lines the same size.
4. A diameter of a circle is a line which goes from one part to the opposite part touching the middle of the circle and keeps inside the circle.

This is a good set of answers for a boy only nine years of age. Marked on the system of units previously used, the definition of a square receives seven marks, the definition of a triangle receives five

marks, that of the oblong receives seven marks, and that of the diameter of a circle receives three marks. It will be seen that, compared with the standard definitions, there is a loss of one mark in the definition of a square, since the description 'straight' is not applied to the 'lines' or 'sides.' The definition of triangle is correct.

Six marks are lost on the oblong. 'Four equal corners' are omitted, carrying three marks. 'Straight' is omitted in describing the lines or sides, and the two long lines and the two short lines are not described as opposite.

One mark only is lost on the definition of 'diameter;' the line is not described as 'straight.' The marks, totaled, amount to 22.

On the next day's test L. O. goes down one mark. His definitions of square and triangle remain unchanged. In his definition of oblong he omits the two points previously inserted, namely, that the two long lines are of the same length, and that the two short lines are of the same length. But in the definition of the diameter of a circle he inserts the description 'straight' which he had omitted the day before altogether. He thus scores 21 marks for his second test.

Let us now see what happens a month later when the same test is applied a third time. I give his paper in full.

L. O., aged 9 years 1 month, on December 7, 1911, in his first test of Deferred Reproduction, wrote as follows:

1. A square is a shape with four lines all the same length, and four corners all the same size.
2. A triangle is a shape with three lines and three corners.
3. An oblong is a shape with four lines two long lines both the same size, and two shorter lines both the same length.

4. A diameter of a circle is a line inside which goes from one part of the circle and touches the middle of the circle goes on to the opposite part of the circle to where it started and it must be a straight line.

This is a very good paper, and scores a total of 23 marks, an advance on the work of the month before. On the day following, on which was given the second test of Deferred Reproduction, L. O. scored 24 marks, for the description 'straight' of the sides of the square, omitted on December 7, was included on December 8. His average mark for his two tests of Immediate Reproduction, those, namely, of the 9th and 10th of November, was 21.5; his average mark for his two tests of Deferred Reproduction was 23.5.

It must not be thought that every boy obtains more marks a month after the lesson than he does for his immediate tests, but many of them do; and the average result shows only a slight decline, rather more marked in the group taught deductively than in the group taught inductively. This is explained by the fact that both the teaching and the learning were well within the comprehension of the boys. When this is the case, and they work in consequence with interest and enthusiasm, they forget surprisingly little.

It may now be of some value if I give the corresponding papers of a boy in the Deductive Group.

R. S., aged 9 years 2 months, who scored 14, 19, 19, 18 marks in his four preliminary tests, in his first test of Immediate Reproduction wrote:

1. A square is a shape with four sides and four corners. The sides are straight and all the same length. The corners are all the same size.
2. A triangle is a shape with three sides and three corners.
3. An oblong is a shape with four sides and four corners. The sides are straight and there are two long sides and two short sides.

The long sides are opposite one another and are the same length, and the two short sides are opposite and are the same length.

The diameter of a circle is a straight line which goes through the centre of the circle.

Only two units of definition are omitted: the equality of the angles is left out in the definition of the oblong, and the delimitation of the diameter by the opposite parts of the circumference of the circle is omitted in the last definition. It is an excellent paper, appearing on the face of it, if one compares it with the verbal expression of the definitions which were given to be studied, to owe a great deal to a highly developed rote memory. If that is so, it is memory for statements that are really understood, since they persist unchanged without the subsequent intrusion of stupid errors, and an unusually high mark is obtained by this boy for his power of application to new material. I propose to defer consideration of the latter issue, since just now we are concerned only with the tests of Immediate and Deferred Reproduction.

The next day R. S. obtained 29 marks, as compared with 28 of the previous day. There were slight changes of verbal expression. For instance, the triangle became "a three cornered figure with three sides." The equality of the angles was omitted in the square, but on this occasion, though not in the previous test, was included in the definition of the oblong. In the definition of the circle an improvement was shown; the point was included which the day before had been omitted; it was now mentioned that the line went from "one side to the opposite side of the circle."

One month later R. S. scored 28 marks. He

omitted the description 'straight' in his definition of a diameter of a circle which he had before included. On the day following he made the same omission. Otherwise his definitions are just as good as those which he had given a month before. His average mark for Immediate Reproduction is 28.5, and for Deferred Reproduction is 28.0.

Let me give one more illustration, the work of a boy who obtained 6, 6, 7 and 7 marks in his four preliminary tests, and who also was taught in the Deductive Group.

H. W., aged 10 years 1 month, in his first test of Immediate Reproduction wrote:

1. A square is a shape with four corners and four sides the same size.
2. A triangle is a shape with three corners and three sides.
3. An oblong is a shape with two small sides, and two big sides opposite one another.
4. The diameter of a circle is a line passing through the middle of the circle.

Marked on the same units as before, the definition of a square obtains six marks; the definition of triangle obtains five marks; the definition of oblong receives six marks, for it is called a 'shape,' its 'four' 'sides' are implied, its 'two long sides' and its 'two short sides' are noted, and the fact that its 'two long sides are opposite each other.' The definition of diameter receives two marks. This is not a strong paper; it scores 19 marks only as a total, but it implies a very great advance on this boy's preliminary tests. One point of interest lies in this. Whereas, in the papers of R. S., recently given, there was an appearance of rote learning in the answers, there is, in the case of this boy, no direct indication of that.

H. W.'s next test of Immediate Reproduction,

worked on the following day, receives the same number of positive marks, namely, 19. A 'bad error' has crept in, for the corners of the triangle are described as all the same size. The 'two small sides' of the oblong are now called "two small tops," but this and the 'bad error' are the only changes. One month later, for his first test of Deferred Reproduction, H. W. wrote:

1. A square is a shape with four corners and four sides opposite one another and they are all of the same length.
2. A triangle is a shape with three corners and three sides they are not opposite one another.
3. An oblong is a shape with three corners and three sides, it is a zig-zag shape not all the same length.
4. A diameter of a circle is a line passing through the middle of it.

Considerable changes are evident in this paper. There are, as before, six positive units of correct description in the definition of the square; but the statement "four sides opposite one another" has been adjudged a 'bad error.' It is, of course, the confused application of some phrase remembered, but not understood. Let it not be supposed, however, that no child inductively taught makes similar errors.

The definition of triangle receives five marks as before. The memory of the oblong has largely gone. It is still remembered that it is a 'shape' and has 'corners' and 'sides,' and thus three positive marks are obtained. But to give an oblong 'three' corners and 'three' sides and to call it 'zig-zag' shape is held to involve three bad errors. The definition of diameter remains unchanged, and scores two marks. The paper as a whole receives 16 positive marks, and there are four marks for bad errors.

But on the next day, in his second test of Deferred Reproduction, H. W. made a decided recovery. He then wrote:

1. A square is a shape with four corners and four sides they are opposite one another, with all the sides and corners an equal size.

2. A triangle is a shape with three sides and corners it is a zig-zag shape.

3. An oblong is a shape with four sides, two long sides and two short tops they are opposite one another.

4. A diameter of a circle is a line passing through the middle of it.

This is undoubtedly H. W.'s best paper. He scores the highest marks he has yet scored for the definition of the square, namely, seven positive marks, since, for the first time, he has mentioned the equality of the corners, but he retains his 'bad errors.' The definition of a triangle remains unchanged in correct units; it is held inadmissible to call the triangle a zig-zag shape. The definition of oblong has returned to its first condition; indeed, it is rather better, for it is easier now to regard H. W. as implying that the 'two long' and 'two short' sides are pairs of equals. The mark for the double equality is, however, not given, as the meaning is somewhat doubtful. The definition of diameter remains unchanged. H. W. scores 20 positive marks for his paper and one negative mark for a 'bad error.' Again we find the marks for Deferred Reproduction not much inferior to those of Immediate Reproduction in this case; indeed, the last paper is the best the boy did throughout the series.

I trust that the inclusion of these papers will be of service in giving life and body to the rather bloodless array of figures, which I give subsequently, dealing

with the results of the tests in Immediate and Deferred Reproduction.

The Test of Application to New Material was identical with that used in the experiment previously described. Drawings of rhombuses, trapeziums, rhomboids and diagonals of squares, with their names appended, were shown to the children, and they were required to answer in writing the questions: "What is a rhombus?" etc. It may add to the facility with which the progress of this experiment is understood if I give verbatim one or two of the worked papers. In the test of application to new material negative marks were assigned as well as positive marks.

L. O., aged 9 years, a boy who worked in the Inductive Group, whose work in Immediate and Deferred Reproduction has already been quoted, wrote the following paper in this test:

1. A rhombus is a shape containing four lines all the same length, so that if you looked at it one way it seems to bend backward, and if you look at it again it looks to bend forward.
2. A rhomboid is a shape also containing four lines, two long lines both the same length, and two shorter lines both the same length.
9. A trapezium is a shape with four lines three long ones, and one short one.
4. A diagonal of a square is a streight line going from one corner to its opposite one.

L. O. receives four marks for his definition of rhombus—one for 'shape,' one for 'lines,' one for 'four,' and one for all the same 'length.' He receives three marks for his definition of trapezium, one for 'shape,' one for 'lines,' and one for 'four.' His statement that there are three long lines and one short one was not held to be equivalent to the statement that the sides were unequal, but it was not considered a 'bad error.' For his definition of rhomboid

he obtains seven marks—one for 'shape,' one for 'lines,' one for 'four,' one for 'two long lines,' one for 'two shorter lines,' and two for the pair of equalities in the length of the lines.

The definition of diagonal receives four marks, one for 'line,' one for "streight," and two for "from one corner to its opposite one."

The paper scores a total of 18 positive marks, and there are no 'bad errors;' the *average* mark obtained by the boys of the Inductive Group is rather lower than this.

R. S., aged 8 years 5 months, who also was taught in the Inductive Group, wrote:

1. A rhombus is a figure with four straight sides and four equal corners.
2. A trapezium is a figure with four corners which are equal with four sides.
3. A rhomboid is a figure with two small sides which are horizontal and two bigger parlerlell lines equal.
4. A diagonal of a square is a straight line from one corner to another corner.

The first definition receives a mark for 'figure,' a mark for 'sides,' one for 'four,' one for 'straight,' one for 'corners,' and one for 'four;' a total of six positive marks; but there is one 'bad error'—the corners are *not* equal: boys taught inductively can obviously make the same sort of blatant error as boys taught deductively.

The second definition receives a mark for 'figure,' one for 'corners,' one for 'four,' one for 'sides,' and one for 'four' (sides); a total of five positive marks; but, again, there is a 'bad error'—the angles of the trapezium are *not* equal. The definition of rhomboid obtains seven positive marks—one for 'figure,' one for 'sides,' one for 'two small' (sides), one for 'two

bigger' lines, a mark for saying the two bigger are equal, and one for saying the two bigger lines are parallel. There is one 'bad error;' the two small sides were in one case only drawn horizontally. The definition of diagonal receives three positive marks —one for 'line,' one for 'straight,' and one for "from one corner to another:" the further specification of 'opposite' corner is omitted.

The paper, as a whole, gains 21 positive marks, with three negative marks for 'bad errors.'

Let me quote one more illustration from among the boys who were taught inductively.

H. B., aged 9 years 2 months, wrote:

1. A rhombus is a fugare which is like a square and has fore corners.
2. A trapezium is something like a triangle only it has fore corners.
3. A rhomboid is a sought of fugare which is something like an oblong.
4. A diagonal of a square is a square with a line across the midal.

This is a very weak paper; it was worked by a boy who was almost at the bottom of the Inductive Group in the preliminary tests, and he seemed to justify his position. It is psychologically interesting that he apprehended the similarity between the work now required and the work he had been taught, but was unable to specify the points of similarity and difference between the figures of the first set and the figures of the second set. He had but little knowledge and could not apply much of that. His marks are: three for his definition of rhombus, two for his definition of trapezium, one for his definition of rhomboid, and one for the definition of the diagonal of a square. "Across the midal" is not held to be

wrong, though it might be; in any case it is not regarded as sufficiently definite to obtain a mark. It is regarded as a 'bad error' to say that the diagonal of a square is a square. H. B. receives a total of seven positive marks, with one negative mark for bad errors.

It is, perhaps, worthy of note that this boy falls from 17 in his test of Immediate Reproduction to 9 in his test of Deferred Reproduction. He cannot apply his old knowledge and he cannot remember it for more than a day or two.

Let us now turn to some illustrative examples of the work of the group taught deductively.

G. M., aged 8 years 1 month, wrote:

1. A rhombus is a figure with two slanting sides and two straight ones arranged so that two of the sides are facing each other and the other two opposite each other and also four corners.
2. A trapezium is a figure with four slanting sides arranged so that there are two sides nearly the same length, these two are generally touching each other. Then there is a smaller one and yet a smaller one still, so that there are four sides and two equal ones the others ofcourse are not.
3. A rhomboid is a figure with two slanting sides and two straight ones and also four corners two of the sides are longer than the other two and also are opposite one another and so are the two shorter sides. There can be ones upright and lying down and also slanting ones.
4. A diagonal of a square is a line drawn from one corner to the other it need not have to be drawn from a corner for it could be from the middle of the top to the middle of the bottom, but you can't have it so that it is from the middle of the one side to the middle of the bottom or to the middle of the top. For the diameter is the greatest and longest line you can have across it or down it and that wouldn't be the longest, not nearly.

This is an excellent paper for a boy of eight years of age. He was taught in the Deductive Group, but evidently he is quite capable of applying what he has learnt. It would be a serious error to suppose that because a boy has *learnt* a set of definitions therefore

he cannot apply them. In a very large number of cases he certainly can. The contention raised in this monograph is that inductive teaching produces a higher transfer to new material than deductive, not that deductive teaching involves no transfer at all. This first-rate paper may do something to prevent an exaggerated conclusion which the subsequent figures may not succeed in adequately moderating. Let us mark the paper on the usual system of units. G. M. is evidently using the word straight to mean, as it often does with boys, horizontal and vertical; he does not mean that only two of the lines are 'straight' in the proper sense. And he is wrong on his own meaning, for one of the rhombuses drawn had neither vertical nor horizontal lines, but two of them had, and to these he has apparently confined his attention. He receives a mark for 'figure,' a mark for 'sides,' and one for 'four' (sides), which is involved in his pair of twos, and one for 'corners' and one for 'four.' He gets two marks for seeing that the opposite sides are paired. This marking yields a total of seven positive marks, whilst he receives a negative mark for being wrong on his own meaning of 'straight.' In his definition of trapezium he receives a mark for 'figure,' one for 'sides,' and one for 'four.' His first description of the sides is held to be equivalent to saying they are unequal, so he receives a mark for that. Later he is marked for a 'bad error' in saying that two of the sides are equal. They are so in one of the trapeziums only. For the definition of trapezium, then, he gets four positive marks, with one negative mark for a 'bad error.' Again, in his definition of a rhomboid we find a misuse of the word straight, and again he is wrong, even on his own meaning.

But he obtains positive marks for 'figure,' for 'sides,' for 'four' (sides), for 'two long' (sides), for 'two shorter' (sides), and two marks for noting the pairs of opposites. He also notes the 'four corners.' He thus receives nine positive marks and one mark for a 'bad error.'

His definition of diagonal is extremely interesting. He receives two positive marks only—one for 'line' and one for "from one corner to another." After that, alas! the transfer from diameter (the corresponding definition which was learnt) has been too thorough. No diameters were drawn in the squares which were before the boy's eyes, and it is not unfair to call the lapse into diameter a 'bad error.' This definition receives therefore two positive marks and one negative mark. The paper, as a whole, receives a total of 22 positive marks, and there are four bad errors; it is considerably above the average of the papers worked by the Deductive Group generally.

H. W., aged 10 years 1 month, whose work in Immediate and Deferred Reproduction has already been quoted, wrote the following in his test of application to new material:

1. A rhombus is a shape something like the shape of a diamond.
2. A trapezium is a shape with four corners not opposite one another their are different shapes of trapeziums they are a zigzag shape some corners longer than others, they are not squares.
3. A rhomboid is a shape with two small tops both opposite one another, and with two long sides with the corners exactly opposite one another.
4. A diagonal of a square there is a square and a line passes right through. Sometimes they pass from side to side other times from corner to corner.

H. W.'s definition of rhombus receives one mark only—a mark for the description 'shape.' For the definition of trapezium three positive marks are

gained—one for 'shape,' one for 'corners,' and one for 'four.' There are no 'bad errors.' It was not thought admissible to regard the expression "some corners longer than others" as involving the inequality of the angles. His definition of a rhomboid receives a mark for 'shape,' one for 'sides,' one for 'corners,' and one for 'four' sides, for the number of sides is involved in the rest of his answer. He also receives a mark for "two small tops," one for "two long sides," and one for noting that the two small sides are 'opposite' each other. The opposition of the angles has not been allowed for in the system of marking. This definition therefore receives a total of seven positive marks. The definition of diagonal receives two positive marks only—one for 'line' and one for 'from corner to corner.' It was regarded as a bad error to say that "sometimes they pass from side to side." The total marks for this paper amount to 13 positive marks, from which one has to be deducted for 'bad errors.'

Let me now pass to the work of a boy who scored 13, 12, 12 and 11 in his four preliminary tests. It seems likely from these figures that we are dealing with a boy of little educability, and this suggestion is confirmed by his later work. In his two tests of Immediate Reproduction he scores an average of 18.5 marks; in both tests of Deferred Reproduction he scores 13 marks, so that a month afterwards he is back again to the position he occupied *before* he learnt the definitions, and he completely fails in applying what he has learnt.

A. R., aged 8 years 6 months, the boy whose work has just been described generally, wrote:

1. A rhombus is a square which is not strate up.

2. A trapezium is a four-sided thing which sides are not all strate.
3. A rhomboid is like an oblong but its lines are not strate up.
4. A diagonal of a square is a diameter of a circle only it is a squear.

A. R. has seen some general resemblance between the 'figures' of his first set of definitions and those of his second set, but the resemblances have hindered rather than helped him, for a rhombus is *not* a square, and a diagonal of a square is *not* a diameter of a circle. The meaning of the word "strate" is misconceived; his reference to the sides of the trapezium is not incorrect on the basis of his own meaning. Of positive marks, on the system of marking adopted, he can obviously obtain very few. He scores no marks for his definition of rhombus, two for his definition of trapezium, one for his definition of rhomboid, and none for his definition of diagonal. His three positive marks are subject to a deduction of two for the 'bad errors' previously indicated. Boys of this kind are the despair of the teacher, but the evidences yielded by his work do not point so much to stupidity as to ineducability.

Possibly the reader may already have gathered from a perusal of the papers which I have used as illustrations some opinions of his own as to the relative applicability of the two methods of teaching and learning. But all such opinions need to be confirmed or modified by a consideration of the tables of results which are set out in the next section.

5. *Results of the Experiments.*

(*a*) *Results of the Preliminary Tests.*

The marks for the four preliminary tests were fairly steady, decidedly so, when the age of the chil-

dren was taken into consideration. Very few of the boys made any violent jumps, and there was a general improvement from exercise to exercise.

In the first test the average mark was 11.1, in the second 12.3, in the third 12.9, and in the fourth 13.1. The correspondences between the results of the first, second, third, and fourth Preliminary Tests are shown compendiously in the following table:

Table XV, showing the correlation between the results of the four Preliminary Tests.

Marks in the four preliminary tests.	No. of boys.	Average Marks in Preliminary Tests. First test.	Second test.	Third test.	Fourth test.
70 and over.....	4	16.5	19.8	18.3	18.3
60 to 70.....	8	14.8	14.4	16.5	17.3
50 to 60.....	11	12.5	13.4	13.9	14.3
40 to 50.....	15	10.4	11.3	12.1	11.9
Below 40.....	12	6.7	7.9	8.9	9.0

There is obviously high positive correlation between the results of the successive preliminary tests. The mental functions we are testing appear to be working very steadily. Exact numerical values for the coefficients of correlation have been worked out from the 50 individual cases on the Pearson formula. Between the results of Tests 1 and 2 the correlation coefficient is + .76 (probable error .04), between Tests 2 and 3 is + .79 (probable error .03), and between Tests 3 and 4 is + .80 (probable error .03). These high correlations between the results of the successive tests give us reasonable expectations that a division into two equal groups may be satisfactorily effected. The boys were divided into two equal groups containing 25 children each. The following table will indicate how the division was made:

Table XVI, showing the Division into Two Equal Groups.

Group A.

Name (Initials only).	Marks for Preliminary Tests. First.	Second.	Third.	Fourth.	Total.
R. D.	18	22	18	22	80
A. C.	17	18	19	16	70
L. O.	13	16	18	18	65
.....	..	..	..	..	..
W. G.	4	9	6	12	31
.....	..	..	..	..	..
G. K.	3	5	7	10	25
Averages	11.1	12.4	12.9	13.3	49.7
M. V.'s	3.5	2.9	2.5	2.4	

Group B.

Name (Initials only).	Marks for Preliminary Tests. First.	Second.	Third.	Fourth.	Total.
H. B.	17	20	17	17	71
R. S.	14	19	19	18	70
C. L.	16	17	16	18	67
.....	..	..	..	..	..
S. B.	9	8	11	5	33
.....	..	..	..	..	..
A. W.	4	8	6	6	24
Averages	11.1	12.3	13.0	12.8	49.3
M. V.'s	2.7	2.8	2.7	3.3	

The average mark per boy per test for Group A was 12.4 (mean variation 2.6), and for Group B was 12.3 (mean variation 2.6). The average age of Group A was 9 years 3 months, and of Group B was also 9 years 3 months.

(*b*) *Results of the Tests in Immediate and Deferred Reproduction.*

It now remains to be shown which of the two groups was the more successful when tested on precisely what they had been taught or learnt.

First, let me give the marks of the two groups as wholes, together with their variability:

Table XVII, showing the work of the Inductive and Deductive Groups compared, in the Preliminary Tests and in the Tests of Immediate and Deferred Reproduction (positive marks only).

	Average Marks.				
	For all four preliminary tests.	First immediate reproduction.	Second immediate reproduction.	First deferred reproduction.	Second deferred reproduction.
Inductive group...	12.4	18.8	18.6	18.0	18.1
M. V.'s.....	2.6	2.6	2.5	3.0	3.2
Deductive group...	12.3	20.5	20.6	18.8	19.4
M. V.'s.....	2.6	3.4	4.1	3.4	3.5

In the tests for deferred reproduction, it will be remembered, negative marks were given as well as positive marks. The marks for the two groups are given below after the negative marks have been subtracted from the positive marks:

Table XVIII, showing the marks (after deduction) for the Inductive and Deductive Groups compared, in the Preliminary Tests, and in the Tests of Deferred Reproduction.

	Average mark for four preliminary tests.	Average Marks. First deferred reproduction.	Second deferred reproduction.
Inductive group.....	12.4	17.7	17.8
M. V.'s.....	2.6	3.2	3.4
Deductive group.....	12.3	18.6	19.1
M. V.'s.....	2.6	3.6	3.7

There seems no doubt that, when the tests are given on precisely the subject-matter which has been learnt or taught, the group which learnt the definitions did better work than that which was taught inductively, and this is true both in immediate and deferred reproduction, and for both positive and negative

marks. This conclusion must, of course, be drawn subject to the age and mental proficiency of the pupils. It now remains to be seen whether the difference between the groups is one which is common to the more proficient as well as to the less proficient pupils:

Table XIX, showing the marks of the two groups compared, section by section, in the Preliminary Tests and in the Tests of Immediate and Deferred Reproduction (positive marks only).

Marks in four preliminary tests.	Inductive Group. No. of boys.	Inductive Group. Average mark of two tests immediate reproduction.	Inductive Group. Average mark of two tests deferred reproduction.	Deductive Group. No. of boys.	Deductive Group. Average mark of two tests immediate reproduction.	Deductive Group. Average mark of two tests deferred reproduction.
70 and over...	2	19.0	19.2	2	24.0	21.0
60 to 70.....	4	21.2	19.5	4	22.5	20.9
50 to 60.....	5	19.7	19.2	6	22.9	21.0
40 to 50.....	8	18.0	18.5	7	18.8	18.3
Below 40.....	6	17.0	14.8	6	17.8	16.0

It seems clear that there is a balance of advantage all along the line in favor of the group which learnt the definitions, so far, at least, as the positive marks are concerned. It now remains to be shown whether this is also true when the negative marks are deducted from the positive marks:

Table XX, showing the marks of the two groups compared, section by section, in the Preliminary Tests and in the Tests of Deferred Reproduction (after the negative marks have been deducted).

Marks in four preliminary tests.	Inductive Group. No. of boys.	Inductive Group. Average mark in two tests deferred reproduction.	Deductive Group. No. of boys.	Deductive Group. Average mark in two tests deferred reproduction.
70 and over.....	2	19.3	2	21.5
60 to 70.....	4	19.1	4	20.5
50 to 60.....	5	19.1	6	20.9
40 to 50.....	8	18.4	7	18.4
Below 40.....	6	14.2	6	15.4

Again, there seems a decided balance of advantage on the side of the group which learnt the definitions deductively.

(c) *Correlation Between the Results of Immediate and Deferred Reproduction.*

It would seem likely from the tables given above that the tests given immediately after the teaching and learning may be regarded as fairly significant of the relative position of the two groups even after considerable time has elapsed—in this case after a month. As this is a very important issue for experimental pedagogy, it may be well to subject the hypothesis to further determination. The following tables will show in a general way how far the suggestion may be taken as valid:

Table XXI, showing the correlation between the marks obtained in the various Tests of Reproduction (positive marks only).

Inductive Group.

Marks for the first test of immediate reproduction.	No. of boys.	Average Marks per Boy in the Reproductive Tests. First.	Second.	Third.	Fourth.
Over 25	2	27.0	26.5	23.5	25.5
20 to 25	3	22.3	19.7	20.7	20.7
18 to 20	7	19.6	19.7	19.1	19.7
16 to 18	6	17.7	18.1	16.0	15.7
15 to 16	4	16.0	15.0	16.0	16.0
15 and under	3	14.3	15.3	16.0	14.3

Deductive Group.

Marks for the first test of immediate reproduction.	No. of boys.	Average Marks per Boy in the Reproductive Tests. First.	Second.	Third.	Fourth.
Over 28	2	29.0	27.0	25.5	26.0
25 to 28	3	26.7	26.0	22.3	22.7
21 to 25	4	23.5	23.2	22.2	23.2
17 to 21	8	19.7	20.4	18.5	19.7
16 to 17	4	17.0	18.7	16.2	15.5
16 and under	4	13.7	13.2	12.5	12.7

It is obvious that considerable positive correlation exists between the results of the successive exercises. A more precise determination may, of course, be made by means of a correlation coefficient. Worked out by the standard formula from the individual cases, the following are the coefficients: For the Inductive Group the results of the first Test of Reproduction correlate with those of the second to the extent of + .78, the second with the third to the extent of + .57, and the third with the fourth to the extent of + .85. For the Deductive Group the correlation coefficients are: first and second tests, + .86; second and third tests, + .68; third and fourth tests, + .94.

It is obvious that tests given immediately may be fairly regarded as indicative of what will happen later on, at least in such exercises as these, when we are making comparison of one group with another.

(*d*) *Results of the Test on New Material.*

We have seen already, when the tests required an *exact reproduction* of what had been learnt or taught, that the children in this school who learnt deductively, like the Standard V children of the girls' school in the experiment first described, did better work than the group taught inductively. But in both the experiments previously described it was found, when the test given was on *new material,* that the children taught inductively did better work than those who learnt their definitions. Is this advantage also to be found on the side of the inductive group in this school? These children are younger and are less proficient mentally, according to school grading,

than either of the girls' classes whose work has been described. Moreover, they are boys, not girls. Are these variations in conditions such as to produce a difference in the results? It will further be remembered that the inductive group, in this case, was taught by its own teacher, and not by me, so that any intensity of impression due to personal novelty was thereby eliminated.

Perhaps I may be pardoned for a sentence of apparent digression. I hold it extremely important for the science of experimental pedagogy that no result should be taken as valid for general application unless the use of it is justified by its success in the hands of the usual teachers of the school. Its success in the hands of the specialist or other exceptional person is quite insufficient to recommend it for general adoption. Let us, then, see what the results were when the whole experiment was conducted by the teachers themselves. I shall show the work of the two groups compared both in the Preliminary Tests and in the Test of Application to New Material. First, let me give the results of the two groups as wholes:

Table XXII, showing the work of the two groups compared, in the Preliminary Tests and in the Test on New Material.

	Average mark for four preliminary tests.	Average Mark for New Material.	
		Positive marks.	Marks after deduction.
Inductive group.....	12.4	16.3	15.6
M. V.'s.....	2.6	3.5	4.0
Deductive group.....	12.3	15.7	14.9
M. V.'s.....	2.6	5.8	5.2

Again we find, notwithstanding the decidedly superior acquisition of the material studied (see Table

XX) on the part of the deductive group, that they are inferior to the other in their power to attack new material of an analogous nature. Four boys in the deductive group completely failed to make a reasonable application of their old knowledge, obtaining only 6, 7, 2 and 4 marks, respectively, whilst only one boy in the inductive group failed to do so, and he obtained 8 marks.

Let us now see, as we have in previous cases, how far this difference is to be found for the weaker as well as for the abler children of each group:

Table XXIII, showing the work of the two groups compared, section by section, in the Preliminary Tests and in the Test of Application to New Material (positive marks, and positive marks after the deduction of the negative marks).

Group Inductively Taught.

Marks in four preliminary tests.	No. of boys.	Marks for Application to New Material.	
		Positive only.	After deduction.
70 and over	2	18.5	18.0
60 to 70	4	18.3	17.0
50 to 60	5	17.8	17.6
40 to 50	8	15.5	15.0
Below 40	6	14.0	13.2

Group Deductively Taught.

Marks in four preliminary tests.	No. of boys.	Marks for Application to New Material.	
		Positive only.	After deduction.
70 and over	2	18.0	16.5
60 to 70	4	15.0	14.2
50 to 60	6	17.8	17.2
40 to 50	7	14.0	13.3
Below 40	6	15.3	14.6

Except in the case of the least proficient section of boys at the bottom of each group, there seems to be

an advantage all along the line in favor of the inductive group. When, therefore, the tests are tests of the application of knowledge rather than an exact reproduction of it, we are, perhaps, entitled, on the whole, to conclude that inductive methods are the better. It may be noted that whilst the marks for the inductive group proceed regularly downwards from the highest section to the lowest, those for the corresponding sections of the deductive group do not. The variability for this group is disproportionally high, due, doubtless, to the psychological fact that for some children of this age the step from knowledge to the application of it is a very considerable one; whereas, of course, the children of the other group had been through a process of applicable method when they had received their inductive lesson. The variability of the work is, however, decidedly high, and the difference between the means of the work of the two groups is very small; and did this experiment stand alone, I should hesitate before putting much confidence in the conclusion which I have indicated above. But its consilience with the previous results lends strength to the conclusion, especially when the differences in the conditions under which it was obtained are taken into account.

VII. FOURTH SERIES OF EXPERIMENTS.

1. *General Plan.*

In the experiment now to be described, just as in those previously recounted, the work was done with all the children of one class, under one teacher, with the same curriculum of study, and working according to the same time-table of instruction. The experiment was carried out in a municipal higher grade school for boys, an elementary school situated in a somewhat mixed neighborhood. The class chosen for the experiment was the First Class in the school, containing 35 boys, graded as Ex. VII on the English standard system of school grading, of an average age of approximately 13½ years.*

The teacher of the class had a theoretical acquaintance with psychological work, and had already carried out some observations in educational psychology. He was, especially, capable of temporary dissociation between the pedagogic and psychologic attitudes—a necessary capacity in an experimenter. Beyond this, he was a first-rate teacher who varied his methods according to the subject-matter with which he was dealing.

As in previous cases, the class was divided into two equal groups on the results of tests in spontane-

*In America this would constitute Grade VIII, or rather, perhaps, the First Year of High School.

ous definition, but the test on which the division was effected was not the same as that used for the purpose in the former tests. But, as before, there were tests of immediate and tests of deferred reproduction, and a test of application to new material of an analogous kind. Further relevant conditions will be given in the details which follow.

2. *The Preliminary Tests and the Method of Marking.*

The first test in this series was the spontaneous definition of squares, triangles, oblongs, and diameters of circles, which were drawn in the way already indicated, and the questions (with which by now the reader will be quite familiar): "What is a square?" etc., were set for written answers. The papers were marked on the system of units which has already been described, and an average mark was gained of 19.1 out of a possible maximum of 30. This, as might have been expected, was by far the highest mark that had been obtained by any class doing this test. It was *not* proposed to use this test on squares, triangles, etc., for the purpose of dividing the class, but it served a useful purpose as a preparatory exercise.

Next week the teacher taught all the children of the class how to arrive at the definitions inductively in the way that I have described in the first series of experiments (p. 33). This lesson also rendered valuable service. It gave full opportunity to all to understand quite clearly what they had to do when they were set to attack the preliminary test on which the class *was* to be divided.

The questions used for the preliminary test were

the same as those which, in the previous schools, had been used as a test of the power of application to new analogous material. In one sense it is, of course, in this case, also a test of application to new material, for one inductive lesson on the square, etc., had already been given. We may, indeed, look upon our division into two 'equal groups' in this case as being effected during the course of a series of lessons instead of at the very beginning of it.

The questions were: "What is a rhombus?" "What is a trapezium?" "What is a rhomboid?" and "What is a diagonal of a square?" The answers were marked on the system of units which has already been described.*

One or two of the papers worked may be of service in enabling an experienced teacher to gauge the mental level of the boys taking these tests.

Edward S——, aged 14 years 8 months, wrote:

1. A rhombus is a figure, it has 4 equal straight lines, it has angles, there are 4, 2 equal large ones and 2 equal small ones.
2. A trapezium is a figure, it has 4 straight lines of different lengths, it has angles, there are 4, all of different sizes. any shape.
3. A rhomboid is a figure, it is enclosed by 4 straight lines, 2 equal long ones, two equal short ones. It has 4 angles, 2 equal small ones and two equal large ones.
4. A diagonal is a straight line going from one corner to the other of a square terminating at both ends dividing the square into 2 triangles.

This is not the best paper; there are four boys in the class who get higher marks, but it is obvious that we are here dealing with a very different mental level, geometrically, from those at which we have previously worked. With the table of units at hand,

*The reader is recommended to turn to page 41 for the list of correct units of description.

it is quite easy to mark this paper. The only difficulty occurs in the case of the last definition, in which the qualification *opposite* is omitted when the corners of the square are mentioned. It is held, however, that the statement "dividing the square into 2 triangles" is equivalent to the limitation of from one corner to the 'opposite' corner. A total of 37 positive marks was gained—11 for the definition of rhombus, 8 for the definition of trapezium, 14 for the definition of rhomboid, and 4 for the definition of a diagonal of a square. There are no 'bad errors.'

Charles B——, aged 13 years 9 months, wrote:

1. A rhombus is a figure or drawing consisting of 4 straight lines. All the lines are the same length and the two lines opposite one another are parrallel to one another. It has 4 corners and four equal angles.
2. A trapezium is a drawing. It has 4 straight lines, 4 angles, all the lines are of different lengths. It has two long and two short sides. The angles are all different.
3. A rhomboid is a drawing consisting of 4 lines which are straight. It has 2 long and 2 short. The 2 short are parrallel to one another, and the 2 long are parrallel to one another. It has 4 corners and four angles. The 2 long sides are the same length and the 2 short are the same length. All the angles are not equal.
4. A diagonal is a line, must be straight. It is drawn from one corner to the one opposite. It passes through the centre of the figure. It does not go outside the figure. It must touch the corners.

This also is a good paper, gaining one mark more than the average for the whole class. The marks are quite easily given. The definition of rhombus gains 9 positive marks. There is one 'bad error'—the four angles are *not* equal; but throughout this experiment we worked with positive marks only. The definition of trapezium gains all the positive marks possible on our scale of marking, namely, 8. It is interesting to note that the term 'corners' is not in C. B.'s mind synonymous with 'angles.' In every one of his defi-

nitions in which the confusion *can* occur he makes the same duplication, but these duplications are not 'bad errors' according to our system of marking. The definition of rhomboid receives 12 positive marks, and the definition of a diagonal of a square receives full marks, namely, 4. The total marks for this paper amount to 33. It is scarcely necessary to multiply examples; sufficient have been given to show how much more competent, geometrically speaking, these boys are than those with whom we worked in the previous boys' school. I turn now to the chronology of the whole of the experiment.

3. *Chronology of the Experiment.*

The first test in this series was given on Friday at 9.30 A. M., September 29, immediately after Scripture lesson. In this test the boys were asked, untaught and unaided, to define square, triangle, etc. Exactly one week later *all* the boys in the class were taught inductively how to arrive at the definitions of square, triangle, etc. On Tuesday, October 17, at 11 o'clock, immediately after recreation, the test used as a Preliminary Test in this school, "What is a rhombus?" etc., was given, on the results of which the class was divided into two equal groups. In this experiment one test only was given for purposes of division. It was hoped that the preparatory work with the squares, triangles, etc., together with the greater age and proficiency of the children, would result in the necessary steadiness, and that the boys' variability would be so small that one test would suffice.

On Thursday afternoon, October 19, from 2.30 to 2.50, one of the groups was taught inductively how to

arrive at the definitions of rhombus, trapezium, rhomboid, and diagonal of a square. Exactly the same method was followed as that used by me in the first and second experiments. The teacher of the class had heard me 'teach' the definitions, so there was no danger that he would vary the method essentially; but a minor variant was employed. He jotted down on the blackboard (which I did not), in an abbreviated form, the 'units' of description as the boys supplied them. His argument for the variation was that the boys who were going to study the definitions deductively would have visual memories of verbal descriptions to help them, and that the inductively taught group ought also to have some visual verbal memories to assist them. Whilst the boys of one group—Group A—were being taught the definitions, the other group—Group B—went into the school hall and had a reading lesson under a student-teacher. From 2.55 to 3.15 the boys of Group A went into the school hall and took the reading lesson, whilst Group B came back to their own teacher and studied the definitions of rhombus, etc., which had been constructed from the spontaneous descriptions of the Preliminary Test and had been already written in preparation upon a blackboard, with the appropriate drawings.

Definitions of Rhombus, Trapezium, Rhomboid and Diagonal of a Square in the Form in Which They Were Given to the 'Deductive' Group to Study.

A *rhombus* is a figure enclosed by 4 equal straight lines. Two sides opposite are parallel, and the other two sides opposite are parallel. It has 4 angles, 2

large and 2 small. The 2 large angles are equal and opposite, and the 2 small angles are also equal and opposite.

A *trapezium* is a figure enclosed by 4 unequal straight lines. It contains 4 unequal angles.

A *rhomboid* is a figure enclosed by 4 straight lines, 2 long and 2 short. Two long sides are equal, opposite and parallel, and the two short sides are equal, opposite and parallel. It has 4 angles, 2 large and 2 small. The 2 large angles are equal and opposite, and the 2 small angles are equal and opposite.

A *diagonal of a square* is a straight line which starts at an angle and passes across the square to the opposite angle.

The boys were told to study the definitions, and they, as well as the boys of the inductive group, were made aware that they would be required to answer questions on them. The time from 3.15 to 3.30 was spent by the boys of both groups in the playground. At 3.30 all the boys returned to their classroom; the questions, "What is a rhombus?" etc., which had been written on the blackboard, were exposed to view, and the boys wrote the answers. There was one other variant from the method which I had used myself, for the drawings of the figures were placed before the boys whilst they were answering the questions in their tests of reproduction.

Exactly one week later, on Thursday, October 26, at 3.30 P. M., the boys of both groups worked a test in *deferred* reproduction, following immediately upon the recreation interval, as in the test of *immediate* reproduction.

Two weeks after this test, at 10 o'clock in the morning, on Wednesday, November 8, following two

short lessons on Scripture and French reading, the test of Application to New Material was given.

Perhaps a summarized note showing the main chronological issues involving differences from other experiments may be of service. First, both groups had inductive teaching, as well as inductive practice, *before* the Preliminary Test. There was one preliminary test, and only one. The test of deferred reproduction was given *one* week after the test of immediate reproduction. The test of application to new material was given three weeks after the teaching and learning which we were relying on to differentiate the groups, and two weeks *after* the test of deferred reproduction.

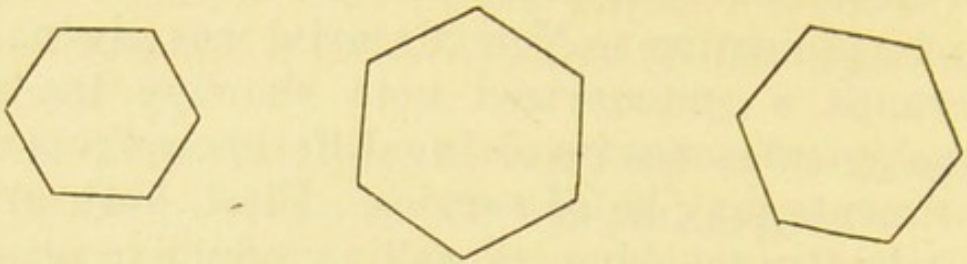
4. *The Tests of Immediate and Deferred Reproduction.*

In these tests all the boys in the class answered in writing the following questions: "What is a rhombus?" etc. The questions were written on the blackboard, and the drawings of the rhombus and other figures were again shown to the boys. I have already pointed out that this was a variation on the method previously adopted.

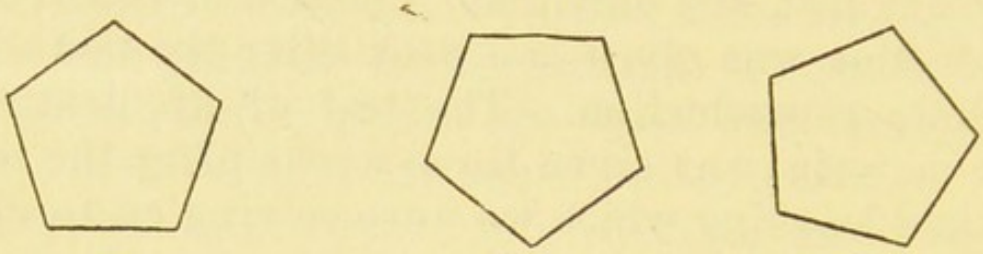
5. *The Test of Application to New Material.*

Drawings of hexagons, pentagons, tangents and quadrilaterals (similar figures), with the names appended, were shown to the boys thus:

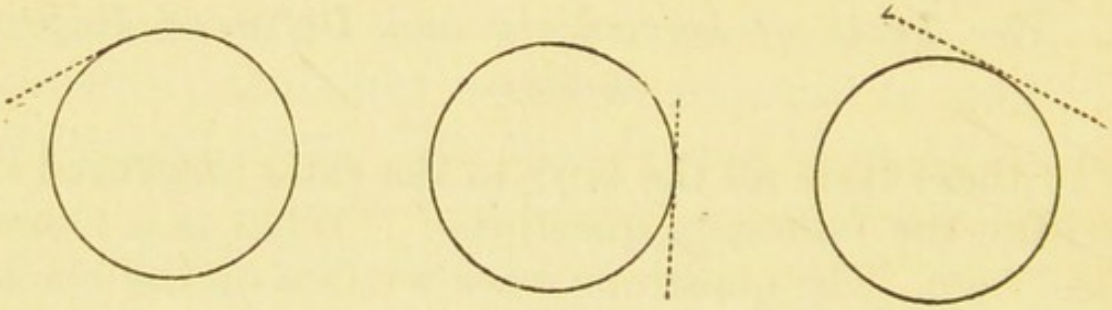
HEXAGONS.

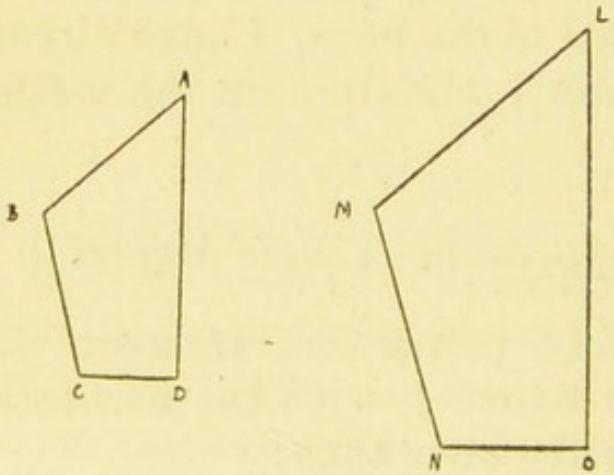


PENTAGONS.



TANGENTS TO CIRCLES.
(The tangents are drawn in dots.)





The sides of LMNO were each 1½ times the corresponding sides of ABCD, so that no easily recognizable ratio should appear. The figures were drawn so that CD and NO were not quite in the same straight line.

In the diagrams actually used the tangents were continuous lines drawn in red.

Then the following questions were written on a blackboard and the boys required to answer them in writing:

1. "What is a hexagon?"
2. "What is a pentagon?"
3. "What is a tangent to a cicle?"
4. "In how many ways is ABCD like LMNO?"

The boys were allowed, nay encouraged, to give thought and time to their answers. It will, doubtless, be remembered that no time limits were imposed in any of the tests and exercises in these experiments. The system of marking the papers could, no doubt, be inferred by analogy from the units of correct description which the boys and girls have given in other cases and which we have adopted. But it is unnecessary for us to infer what our units ought to be; they emerge quite clearly from a consideration of the papers actually worked.

Let me give one or two by way of illustration before listing the units on which the boys' papers were marked.

Frederic R——, aged 13 years 11 months, who worked in the deductive group, wrote:

1. A hexagon is a figure enclosed by six equal straight lines. It has six angles, all equal. The two opposite sides are parallel in the three cases.

2. A pentagon is a figure enclosed by five equal straight lines. It has five angles all equal. None of the sides are parallel to each other.

3. A tangent to a circle is a straight line any length, which must touch the side of the circle anywhere, but must not cut it.

4. The first thing why ABCD differs from LMNO is its size. The 4 angles are the same in both figures. The 4 straight lines are the same only in proportion, LMNO is about half the size again as ABCD. M.N.O. angles are the same as B.D.C. only the sides are

different lengths. A. angle is exactly the same as L. angle. Both the figures are exactly the same shape. The only thing why one is different from the other is in size.

Even without a list of units of correct description it is not difficult to assess this paper. The definition of hexagon receives a mark for 'figure,' four marks for "six equal straight lines," three marks for "six angles equal," and six marks for noting that there were three pairs of opposite sides, and that three pairs were parallel. F. R. thus obtains a total of 14 marks for his definition of hexagon. The definition of pentagon receives eight marks—one for 'figure,' four for "five equal straight lines," and three for "five angles equal." The definition of tangent receives three marks—one for 'line,' one for 'straight,' and one for "touching the side of the circle." A boy's conception of touching would be satisfied if the line impinged upon the circumference of the circle in such a way that, if produced, it would cut the circumference. Consequently it is necessary to add the limitation 'if produced, will not cut the circle.' The fourth answer is a good one, but it is unfortunate that the boy is bothered by the notion that he has to find *differences,* which every now and again intervene among the *similarities.* He calls ABCD and LMNO both 'figures,' for which he receives a mark; for '4 angles' he receives two more; for noting that the four angles are equal, each to each, in the two figures, he receives four marks; for "4 straight sides" three more, and for the similar proportionality of the sides he obtains four more. Finally, he notes that the figures are alike in shape, for which he receives a mark. F. R.'s total mark for this answer is 15, and his mark for his whole paper 40. His marks were 38 for his

preliminary test, 49 for immediate reproduction after teaching, 49 for deferred reproduction, and 40 —the present mark—for application to new material. If these are compared with the average marks given later, it will be seen that he is five or six marks ahead of the average throughout the entire series. I will give one or more worked papers before setting out the units of correct definition which were accepted as the basis of marking.

Robert S——, aged 14 years, who worked in the inductive group, wrote:

1. A hexagon is a figure enclosed by six, straight, equal, sides. The opposite sides are equal and parallel. One side is exactly balanced by the opposite one. It has six angles, which are all equal. Three are on one side and three on the other.

2. A pentagon is a figure enclosed by five straight sides. They may be equal or unequal. No sides are opposite and no sides are parallel. It have five angles. They may be equal or unequal.

3. A tangent to a circle is a straight line. It may be drawn at any angle. It must touch the circle but not cut it.

4. Both have five sides. The base in each case is horizontal. They have five angles each. The angles are the same number of degress in each case. There are two large ones and two small ones. The two large ones are formed by the base and sides and the two small ones from the top and sides. The smallest angle is A in 1 and correspondonds with L in 2. The largest C in 1 and N in 2.

The definition of hexagon obtains 14 marks—one for 'figure,' four for "six straight equal sides;" six for noting that there are three pairs of opposite sides, and that they are parallel, each to each, and three marks for "six equal angles." The definition of pentagon obtains six positive marks—one for "figure," three for "five straight sides," and two for "five angles." The statement 'no sides are opposite

and no sides are parallel' is held to be of too negative a nature for inclusion within the definition. To say that the sides and the angles may be equal or unequal would be accounted 'bad errors,' though, as I have said before, we did not tabulate the 'bad errors' in this fourth experiment. The definition of tangent receives three marks—one for 'straight,' one for 'line,' and one for 'touch the circle.' "It may be drawn at any angle" is too vague to be regarded as either positive or negative. The point is missed that the tangent, *if produced,* will not cut the circle. In the fourth answer there are two curious errors. The figures have 4 sides and 4 angles, and not 5, as R. S. says. He obtains marks for mentioning 'sides' and 'angles' as pertaining to both. Nearly all the rest of the answer is occupied with the equality of the angles each to each, for which 4 marks are obtained. One mark is gained for noting that the base lines in each case are horizontal; that is regarded as equivalent to noting that their inclination is the same. This marking yields a total of 30 positive marks, with 4 'bad errors.' I give this paper because I wish to make it quite clear that boys inductively taught could quite well make 'howlers' as well as boys deductively taught, though these boys, in both groups, make extremely few. R. S.'s other marks were 26, 42 and 43; in all cases, except that of the Deferred Reproduction Test, well below the average. Probably the perusal of the papers given above may make clearer the usefulness of the units of correct definition which are now appended.

Units of Correct Description or Definition of Hexagon, etc.

1. A hexagon is a figure.
   It has sides or lines.
   It has straight sides or lines.
   It has equal sides.
   It has six equal sides.
   Two sides are opposite.
   Two other sides are opposite.
   And the two other sides are opposite.
   Two opposite sides are parallel.
   Other two opposite sides are parallel.
   And the other two opposite sides are parallel.
   It has angles.
   Its angles are six in number.
   And they are equal.
   Its angles are greater than right angles.
   (A total of 15 points.)
2. A pentagon is a figure.
   It has sides or lines.
   Its sides are straight.
   The sides are equal.
   There are five sides.
   It has angles.
   Its angles are five in number.
   The angles are equal.
   And they are greater than right angles.
   (A total of 9 points.)
3. A tangent to a circle is a line.
   It is a straight line.
   The line touches the circle.
   And, if produced, does not cut it.
   (A total of 4 points.)

4. ABCD and LMNO are both figures.
They both have sides.
They both have straight sides.
Their sides are in both cases unequal.
And they are 4 in number in both figures.
They both have angles.
Their angles are 4 in number.
And are in both cases unequal angles.
BA is the same fraction of LM.
As BC is of MN.
As CD is of NO.
As AD is of LO.
BA has the same slant or is parallel to LM.
BC has the same slant or is parallel to MN.
CD is parallel to NO.
And AD is parallel to LO.
The angle at A equals the angle at L.
The angle at B equals the angle at M.
The angle at C equals the angle at N.
The angle at D equals the angle at O.
The figures have the same shape.
(A total of 21 points.)

It is, of course, not urged that the common properties of the figures have been exhaustively enumerated, but only that the units of correct description are such as are actually used by boys and are serviceable for marking papers in such experiments as these.

6. *Results.*

First, let me give the coefficients of correlation between the results for the various exercises in so far as they may be of value. The marks for the Prelimi-

nary Test in the A Group were very closely correlated with those of the B Group; the boys were most successfully paired in the two groups, from the best downwards to the worst. Worked out on the regular formula, the coefficient of correlation amounted to + .98. The results of the test in immediate reproduction correlated with that of deferred reproduction to the extent of + .752 in the inductive group and + .777 in the deductive group. There was a falling off on the average of about one unit in the marks. There were 7 cases out of 34 in which the mark for deferred reproduction was higher than for immediate reproduction, 10 cases in which it was the same, and 17 cases in which there was a decline. The decline of the whole class was from an average mark per boy of 44.94 to 43.50, with mean variations approximating to 4 in both cases, and a correlation coefficient between the results of immediate and deferred reproduction of + .78. Though the difference is small, we are entitled statistically to say that there is a *general tendency to decline,* since the difference between the means is from three to four times the probable error of that difference. A *general* slight decline seems, therefore, clear. The inductive group falls from 45.2 to 43.5; the deductive from 44.7 to 43.5. But the fall is too irregular to enable us to conclude that there is any greater tendency to loss on the part of the children inductively taught than of those deductively taught.

Let us now consider the results of the test on new material. It is clear that the difference between the results of the two groups is very small in this school, though it favors the inductive group, as, indeed, is the case in all the experiments. But the variability

is such that without very high positive correlation between the two series the probable error of the difference between the means will be considerable.

Now let me give the average results in gross, treating the groups as wholes. There were 17 boys in each group:

Table XXIV, showing the work of the Inductive and Deductive groups compared, in the Preliminary Test, in the Tests of Immediate and Deferred Reproduction, and in the Test of Application to New Material.

	Preliminary test.	Test of immediate reproduction.	Test of deferred reproduction.	Test on new material.
Inductive Group:				
Average mark...	32.06	45.18	43.53	35.65
M. V.....	3.13	3.67	3.97	3.24
Deductive Group:				
Average mark...	32.12	44.71	43.47	35.00
M. V.....	2.98	4.69	4.80	3.53

The boys of the Inductive Group appear to hold the advantage throughout, though they were slightly weaker in the Preliminary Test. A closer analysis of the results is given in the next table:

Table XXV, showing the work of the Inductive and Deductive Groups compared, section by section, in the Preliminary Test, the Tests of Immediate and Deferred Reproduction, and the Test of Application to New Material.

Inductive Group.

Marks in preliminary test.	No. of boys.	Average Mark per Boy. Preliminary test.	Immediate reproduction.	Deferred reproduction.	New material.
Over 35.....	4	37.50	47.00	45.50	37.50
30 to 35.....	6	32.50	47.00	45.83	37.66
25 to 30.....	7	28.57	42.57	40.42	32.85

Deductive Group.

Marks in preliminary test.	No. of boys.	Average Mark per Boy. Preliminary test.	Immediate reproduction.	Deferred reproduction.	New material.
Over 35.....	4	37.50	47.25	47.75	37.00
30 to 35.....	6	32.33	44.16	43.16	36.66
25 to 30.....	7	28.85	43.71	41.28	32.42

Only in one test—that of application to new material—does there appear to be a regular sectional advantage on the side of the inductive group, both for the weaker as well as for the stronger boys. In both reproductive tests the balance of advantage shifts from side to side.

We may justifiably conclude that the results of this experiment, having regard to the greater age and mental ability of the children, are consilient with those of the former researches. The inductive method has shown itself the better when application to new analogous material is the test employed. We are unable to say with any confidence which of the two groups has been the more successful in immediate and deferred reproduction. The average results are slightly in favor of the inductive group, but the balance of advantage fluctuates from side to side, and is decidedly uncertain. But this is the fourth case in which the inductive method has shown itself superior in application to new material, and the second case in which the inductive method has equaled the other, even for purposes of reproduction. In both these classes there had been much previous inductive teaching. But it must be remembered as well that the class of much *younger boys,* in which the *deductive* group scored heavily in reproductive work, were *also* accustomed to much inductive work. Age appears to be a factor; perhaps it is the *younger* children who reproduce better on a deductive and memoriter method. This hypothesis will be put to the test in the last of this series of experiments.

VIII. FIFTH SERIES OF EXPERIMENTS.

1. *General Plan.*

Just as in the previous experiment, a whole class, working under one teacher, and according to the same syllabus of instruction, with the same time-table of work, was divided into two equal groups on the results of a test in the definition of geometrical forms, which the boys attempted, untaught and unaided. Then one group worked inductively and the other deductively. Tests were given immediately after the teaching and learning, also in deferred reproduction a week later, and in reproduction, still further deferred, about seven weeks after the first test of deferred reproduction. About two weeks after the teaching and learning a test of application to new material was given. The boys who did the work were graded as Standards VI, a, and VII; their average age was 12 years 9½ months, and their teaching generally had been clear and efficient, but had tended rather towards deductive than inductive methods. The school was situated in a poor neighborhood in the southeast of London, and the average mental ability of its pupils was low;* but the boys of the highest class, with whom the experiment was made, were by no means without ability; in fact, in

*The natural mental ability of the pupils of this school was well known to me from the results of a number of mental tests which had been applied to every child over eight years of age.

consequence of certain exigencies of organization, the class contained more children of ability than would ordinarily be found in a top class of such a size in a school of this social type.

2. *The Preliminary Tests and the Method of Marking.*

The Preliminary Test, on the results of which the boys were divided into two equal groups, was the same as that used in the experiment just described, which took place in the higher grade school. The teacher had already used the questions: "What is a square?" "What is a triangle?" "What is an oblong?" and "What is a diameter of a cicle?" (with the consideration of the appropriate drawings) as a kind of general propaedeutic to the experimental series, and the boys had already been shown inductively how to work out the definitions of square, triangle, oblong and diameter just as they had in the higher grade school.

The Preliminary Test, given two or three weeks later, consisted in the questions: "What is a rhombus?" "What is a trapezium?" "What is a rhomboid?" and "What is a diagonal of a square?" The appropriate drawings were shown and the boys attempted to answer the questions. I give below one or two of their worked papers.

William L——, aged 13 years 8 months, wrote:

1. A Rhombus is a figure with all side equal two sides slope at 60° and the other two run parallel.
2. A Trapezium is a figure with four unequal sides, and it as a right angle in it.
3. A Rhomboid is a figure with two long sides equal and two short sides equal, but none of the corners have right angles.
4. A Diagonal of a Square is the distance across from corner to another corner which slopes at 45°.

W. L., in his first definition, gains a total of six marks. His second definition receives four marks. "It as a right angle in it" is not true as applied to all the trapeziums; it is a 'bad error,' but there are so few of these that they are not tabulated. The definition of rhomboid receives eight marks. The negative statement that there are no right angles, though correct, receives no mark, as we could hardly have made allowance for all the *negative* statements which may be truly made about the figures. The definition of diagonal receives two marks only—one for 'distance' and one for "from corner to another corner." W. L.'s paper receives a total of 20 positive marks. It is one of the best papers worked in the class, and is assessed considerably above the average mark, which is 12.25 for this preliminary test.

Let me now give the paper of a boy who is among those toward the bottom of the lists.

Frank B——, aged 12 years 4 months, wrote:

1. A Rhombus is a square turned in shape.
2. A Trapezium is a figure all sides unequal.
3. A Rhomboid is an oblong with the two smallest perpendicular lines slanting.
4. Diagonal of a square is a line drawn from top to bottom of the corners.

As we have seen in former cases of 'unintelligent' children, the similarities between these figures and those which they have previously dealt with are apprehended, even to the extent of error, for a rhombus is not a square. That a square is one shape and a rhombus is another shape is probably dimly understood by the boy; he is giving, perhaps, what he conceives to be a genetic definition of a rhombus, but he receives no marks for it on our system of marking.

For his definition of trapezium he obtains three marks. The rhomboid he defines genetically; his definition is worth, perhaps, two marks—one for 'lines' and one for "two smallest lines." His definition of diagonal is worth two marks; he describes it as a 'line' and notes that it goes from one corner to another. F. B. thus receives a total of seven marks, which is a little above half the average mark for the class.

The two examples of worked papers, given above, will enable teachers to see the limits of the mental level with which we are dealing. These boys are very obviously much below the first-class boys of the higher grade school whose work we considered in the experiment previously described.

3. *Chronology of the Experiment.*

First of all came the inductive work with the squares, triangles, oblongs and diameters of circles. This was done with all the class. A week or so later, on Wednesday, October 11, at 10 A. M., following immediately upon Scripture lesson, the Preliminary Test was given, on the results of which the boys were divided into two equal groups. Most of the boys had finished their work in 20 minutes, though no one was hurried, and one or two took a few minutes longer. On Thursday, October 12, immediately after registration, the teacher of the class taught one of the groups how to arrive inductively at the definitions of the geometrical figures which they had attempted in the Preliminary Test. The teacher had heard me teach similar definitions and was well acquainted with the method as I employed it. The teaching took 22 min-

utes, from 2.14 to 2.36 P. M. Whilst the one group was being taught by their own teacher, the other group, under another master, studied the definitions with reference to the drawings of the figures which were appended. They knew that the exact words of the definitions were not to be required, but that they might use them if they chose. Both groups of boys were aware that they were to be tested immediately afterwards on what they had been taught or learnt. The definitions given to the 'deductive' group ran as follows:

*Definitions of Rhombus, Trapezium, Rhomboid and Diagonal of Square to Which Appropriate Drawings Were Appended.**

1. A Rhombus† is a figure with four straight equal sides; the opposite sides are parallel. It has four corners, two big ones opposite and equal, and two smaller ones opposite and equal.

2. A Trapezium is a figure or shape with four straight unequal sides and four unequal corners.

3. A Rhomboid is a figure with four straight sides. The two long sides are opposite, equal and parallel. The two short sides are opposite, equal and parallel. It has four corners, two big and two small. The two big ones are equal and opposite, and the two small ones are equal and opposite.

---

*The drawings may be seen on page 39.

†Perhaps a slight amendment might usefully have been made in this definition of the rhombus; it is not clear on this wording that there are *two pairs* of opposite sides which are parallel; the form of words used in the previous experiment seems more satisfactory.

4. A Diagonal of a Square is a straight line drawn from one corner to the opposite corner.

At 2.40 P. M., a few minutes after the teaching and learning, both groups answered in writing the questions: "What is a rhombus?" etc. No time restrictions were laid down; each boy was permitted to go on until he could do no more, but the superiority of the *pace* of the boys who had *learnt* the definitions was evident in this and in all succeeding exercises. After a lapse of one week, at the same hour in the afternoon, namely, 2.40, and on the same day of the week, Thursday, October 19, both groups answered the questions: "What is a rhombus?" etc. This will be referred to as the first test of deferred reproduction. None of the boys were aware that they were ever again to be required to answer these questions; it was only the test of immediate reproduction of which they had been forewarned.

One week later, again on Thursday at 2.40 P. M. (October 26), the boys worked a further test—a test of application to new material—and on Thursday, December 7, at 2.40 P. M., two months after the test of immediate reproduction, a second test of deferred reproduction was given, in which the old questions, "What is a rhombus?" etc., were repeated; and, as before, the boys answered them in writing.

4. *The Tests of Reproduction.*

These were in all cases the same. They consisted, as previously stated, of answers in writing to the questions: "What is a rhombus?" etc. One or two papers to indicate what these boys could do after teaching and learning may be of interest.

Thomas G——, aged 13 years 6 months, one of the best boys who worked in the deductive group, in his exercise in immediate reproduction, wrote:

1. A Rhombus is a figure with four equal straight lines. The opposite lines are parallel. It has four corners, one pair of opposite corners being equal and the other pair of opposite corners being equal.
2. A Trapezium is a figure with four unequal sides, and four unequal corners.
3. A Rhomboid is a figure with four straight sides, two long sides and two short ones. The two long ones are equal and opposite each other, and the two short ones are equal and opposite each other. The figure has four corners, two big ones and two little ones, The two big ones are equal and opposite, and the two little ones are equal and opposite.
4. A diagonal of a square is a straight line drawn from one corner to the opposite corner.

This is an excellent paper; the definition of rhomboid, for example, where it differs from the wording of the definition studied, is better than the definition we provided. The boy rightly says "two long and two short" sides, *before* speaking of "The two long" sides. Our own definition is faulty in that respect. The word 'The' is not only distinguishing, but relative, and, indeed, very often distinguishing because relative. Let us mark the paper in accordance with the list of units of correct description given on page —:

The definition of rhombus receives single marks for 'figure,' 'four,' 'equal,' 'straight,' 'lines,' 'four,' 'corners,' and eight marks for noting two pairs of opposite, parallel sides and two pairs of opposite equal corners—a total of 15 marks. For his definition of trapezium he receives obviously every mark but one. He has omitted 'straight' in his description of the sides, thus receiving seven marks out of eight. Every possible point on our system of mark-

ing is scored by his definition of rhomboid, with the exception of two; he omits the two pairs of parallels, thus receiving 18 marks. The definition of a diagonal of a square receives full marks, namely, four. F. G.'s total mark for his immediate reproduction is 44, which is much above the average of his group.

In his next test, one week later, he loses four marks on his first definition, for he now omits to note the two pairs of opposite parallel sides. His mark for trapezium remains unchanged. In his definition of rhomboid he now notes the two pairs of parallels, which he omitted to do in his test of immediate reproduction, and on this occasion receives full marks, namely, 20. The definition of a diagonal of a square remains unchanged. F. G., therefore, has gone down two marks in one week. Let us see how many he has lost seven weeks after this. The definition of rhombus suffers most; the parallelism of the opposite sides does not reappear, and it is doubtful whether F. G. remembers that there are *two* pairs of opposite equal angles, for his expression is dubious. He has now lost four of the marks he originally obtained for this definition. The definition of trapezium remains unchanged. In the definition of rhomboid two marks are lost, for he now omits to note that there are two obtuse angles and two acute angles. The total mark for this definition is 18. The definition of a diagonal of a square remains unchanged, and receives four marks as before. Two months after learning the definition F. G. receives 40 marks for his reproductive test, against 44 marks in his test of immediate reproduction, and 42 marks in his first test of deferred reproduction, which took place one week after he learnt the definitions. He loses very little; he had

evidently understood the definitions as well as learnt them. Indeed, his understanding is shown by his power of 'transfer,' for he receives a very good mark for his application to new material.

Bearing in mind that this pupil worked in the deductive group, let us compare his work with the corresponding papers of one of the best boys in the inductive group.

George H——, aged 13 years 9 months, in his test of immediate reproduction, wrote:

1. A rhombus is a 'figure,' sides, four sides, all equal, four angles, two opposite sides are parallel, other sides are parrallel, all sides are straight.
2. A trapezium is a figure, sides, of four, all unequal, four angles, angles unequal, all sides are straight.
3. A rhomboid is a figure, of sides, four sides, two opposite sides are equal, parrallel, and has four angles, two opposite angles are equal, sides straight, two opposite sides straight, two sides are longer than the other pair of sides.
4. A diagonal of a square is a line from one corner to the opposite corner, it is also straight.

G. H.'s paper is, like F. G.'s, an excellent one. There is a certain staccato utterance which is a little irritating, but it is a peculiarity of his own and is not shared by the members of the inductive group generally. For the definition of rhombus he receives single marks for 'figure,' 'four,' 'sides,' 'equal,' 'straight,' 'four,' 'angle;' two marks for noting a pair of parallel sides, and that they are opposite sides; and one mark for noting the other pair of parallel sides; but he fails to note that the other pair of parallel sides are opposite also. He also receives two marks for noting that one pair of angles are equal and opposite. His total mark, therefore, for this definition is 12. His definition of trapezium receives full marks. The definition of rhomboid is not

so good. He receives single marks for 'figure,' 'four,' 'sides,' 'straight,' 'four,' 'angles;' three marks for noting that one pair of opposite sides are 'equal,' 'opposite' and parallel; two marks for noting that one pair of angles are equal and opposite, and two marks for stating that two sides are longer than the other two—a total of 13 marks. The definition of diagonal scores full marks. G. H. thus receives a total of 37 marks for his exercise in immediate reproduction.

In a week's time, when he takes his first test in deferred reproduction, he obtains one mark less. In his definition of rhombus he omits the parallelism of the 'other sides,' losing a mark which he had gained the week previous. His definition of trapezium remains unchanged. In the definition of rhomboid, though it is expressed with some slight differences, he obtains all the marks which he received before, namely, 13. The definition of diagonal remains unchanged; for this he obtains four marks, as before, making a total of 36 marks.

Seven weeks later there is a somewhat more serious loss. He still receives 11 marks for the definition of rhombus, which has remained unchanged. His definition of trapezium has improved, for, though it contains no further units of correct description according to our scale, he notes that the smallest angle is opposite the smallest side and the biggest angle is opposite the biggest side. These statements are not quite clear, but indicate the commencement of a fresh idea about the trapezium. Two marks on his previous record are lost in his definition of rhomboid; he omits now that there are two long and two short sides. The last definition remains unchanged. For

the second paper in deferred reproduction, therefore, G. H. receives 34 marks.

These papers written by F. G. and G. H., though much superior to the average work, are typical in the slowness with which points like these of definitional description are forgotten when they have been duly understood, and expressed in a way which is really a result of work on the part of the pupil himself.

5. *The Test of Application to New Material.*

This, after all, is the supreme test of what teachers call 'intelligence.' We have seen in the two papers given above that the boy who learnt deductively knew more of what he had actually studied than the boy taught inductively, not only immediately after the work, but after two months had elapsed; and with the boys of this class we shall find this difference to be true generally between the boys of the deductive and the boys of the inductive groups. The older children hitherto—girls and boys graded as Standard VII and upwards—have not shown this difference, though the younger and less proficient children have. I incline to attribute this to the relative predominance of deductive work in the usual teaching of this class, whereas in the two preceding classes of elder children, both boys and girls, the teaching was predominantly inductive. Are we about to find that these boys give us results which differ from those of the older children previously experimented with, and, indeed, from *all* the children previously experimented with, when test is made of their power to apply their knowledge to new material?

The test of application was the same as that used

with the Ex-VII class in the Higher Grade Boys' School. Drawings of hexagons, tangents to circles, pentagons and quadrilateral similar figures were shown. The questions: "What is a hexagon?" "What is a tangent to a circle?" "What is a pentagon?" and "In how many ways does ABCD resemble EFGH?" were written on the blackboard and the children answered them in writing.*

I will illustrate what the boys did by means of two papers, both above the average, one from the 'deductive' and one from the 'inductive' group.

Harry W., aged 13 years 6 months, who worked in the 'deductive' group, wrote:

1. A hexagon is a figure with six straight sides all of which are equal, it has also six equal corners or angles.
2. A tangent to a circle is a straight line, drawn so that it touches the circumference of the circle.
3. A pentagon is a figure with five straight sides and five angles, all sides being equal and all angles being equal.
4. A. b. c. d. is the same as E. f. g. h. They vary by the sides, and the angles, if you look at them closely and then measure the angles they will all be different on one and all the same as the first on the other. They look different by the size.

With the exception of the last definition, this is an easy paper to mark. The definition of hexagon receives a total of 8 marks. The definition of tangent receives 3 marks. The definition of pentagon receives 8 marks. In the last answer about the similarity of the quadrilateral figures, it is clear that H. W. wishes to express the inequality of the angles in both figures and the equality of the angles, each to each, of one figure with those of the other, for which he receives 5 marks. Thus H. W., taught deductively,

*The drawings may be seen on page 108. One of the two similar quadrilaterals was lettered EFGH on this occasion.

scores 24 marks for his test of application to new material.

Frank C——, aged 13 years 2 months, who was taught inductively, wrote:

1. A Hexagon is a six straight sided figure, having all sides equal, it has six angles equal, larger than right Angles.
2. A Tangent to a circle is a line which is straight and is just touching the boundary of a circle.
3. A Pentagon is a five, equal, straight sided figure, it has five equal angles larger than right angles.
4. Both have four sides.
   Both have four angles.
   Both have four angles which are larger than right angles.
   A angle equals E angle.
   B " " F angle.
   C " " G angle.
   D " " H angle.
   Both have sides with the same slope.
   Both are placed on the same side.

The definition of hexagon receives a total of nine marks; the definition of tangent three marks; and that of pentagon nine marks. The last answer is more difficult to mark. Both figures have 'sides;' this carries one mark. There are four sides in both figures; this carries another mark. Similarly, "Both have four angles" carries two marks. The next statement is wrong; it is not true that both have four angles which are larger than right angles. Then there are four marks for noting the equality of the angles, each to each, and four marks for noting that the sides of the figures are parallel. One further mark is gained by F. C.'s statement that both the figures are on the same side (of the base). This answer, therefore, receives a total of 13 marks. The paper is an excellent one, and carries a total of 34 marks; it is, in fact, one of the best papers worked in either group in the test of application to new material.

Lest the reader should carry away a quite exaggerated notion of the power of application of these pupils (I am using the expression 'application' in the strictest sense), I propose to give one further paper by a boy who worked in the deductive group and made very little application of his knowledge. He obtained 37 marks in his test of immediate reproduction and 34 marks a week after. But he obtained a very poor mark when he worked on new material, and seven weeks later he sank to 23 marks when tested on his old knowledge. There are evidently *some* boys who learn quickly and forget quickly. The pedagogical error, now happily being rectified by psychologists, has been to regard these boys as the rule rather than the exception. This boy, George L., aged 12 years 4 months, wrote:

1. A Hexagon is a six sided figure. Each of the six sides are straight equal and opposite and Paralled.

2. A tangent of a circle is a straight line drawn which is slanting and the circle stands on it.

3. A Pentagon is a figure with five sides, they are all straight. The Three small ones are equal and opposite, and the two long ones are equal and opposite.

4. a. b. c. d's. has two straight long sides equal and the other two sides unequal E. f. g. h's has two long straight sides equal, and the other two unequal only a. b. c. d is smaller than E. f. g. H.

"Equal and opposite" has, unfortunately, transferred itself too successfully. For his definition of hexagon he receives 5 marks. "Each of the sides are opposite and parallel" is considered to be too confused to gain positive marks, but is not regarded as involving 'bad errors.' The definition of tangent gains 2 marks only; the latter part of his definition was drawn from one figure only. The statements that the tangent is slanting and that the circle stands

on it were not true of all the tangents drawn, and are considered 'bad errors.' In his definition of pentagon he receives 4 marks only. It is considered a 'bad error' to say that "three small ones are opposite." No positive marks are allowed for saying that "three are equal" and "two are equal," and it is counted an error to say there are "two long" and "three small" sides. In his last answer G. L. receives 2 marks; both the figures have 'sides,' and in each case two are longer than the remaining two. But none of them were equal; though, as two of them were not very different in length, the statement was not accounted a 'bad error.' The statement as to the size of the two figures is irrelevant; the boys were asked for 'resemblances,' not for differences. This is one of the worst papers in the class. The boy *had* acquired the knowledge of the definitions of rhombus, etc., but he could not apply it, and he speedily forgot it.

Possibly, with these examples before him, the reader may find greater interest in the tabulated results, which I now give.

6. *Results.*

(*a*) *Of the Preliminary Tests.*

In the Preliminary Test the highest mark gained by any boy was 19, the lowest was 6, and the average mark was 12.25. There were 16 boys in the group deductively taught and 16 boys in the group inductively taught. The average mark of the boys of the first group was 12.25 (mean variation 3.0), and of those in the second group was 12.25 (mean variation 3.0). The correlation between the total results of the corresponding boys in the two groups was practi-

cally perfect, amounting to + .97 on the product-moment formula. In so far as *one* test can in any way be satisfactory as a basis of the division of a class into equal groups, it seems fair to suppose that an adequate division has been made. These boys, it will be remembered, had had some special inductive teaching concerning the square, triangle, etc., though I should not describe the general methods of their teacher as predominantly inductive. I incline to think this special inductive propaedeutic may have been an advantage to us in making the division, but it may, I fear, serve to throw some bias on the inductive side and unduly favor the inductive group. We may, however, remember that we have three experiments already described in which no such propaedeutic was given.

(*b*) *Of Immediate Reproduction.*

What marks did the two groups obtain immediately after the teaching and learning? In two previous experiments with older children, girls as well as boys, the group taught inductively appeared to advantage from the first. Is that also the case with these Standard VII boys? We can say quite definitely that it is not. The average mark obtained by the boys of the deductive group was 34.2 (mean variation 6.0), and of the inductive group 31.4 (mean variation 4.5).

This difference between the means and its probable error justify us statistically in asserting the

existence of a general tendency in favor of the 'deductive' group. The superiority of the work of the deductive group in immediate reproduction may also be shown compendiously in the following table:

Table XXVI, showing the work of the Deductive and Inductive Groups compared, section by section, in the Preliminary Test and the Test of Immediate Reproduction.

	Deductive Group.			Inductive Group.		
Marks in preliminary test.	No. of boys.	Average mark in preliminary test.	Av. mark in immediate reproduction test.	No. of boys.	Average mark in preliminary test.	Av. mark in immediate reproduction test.
Over 15.....	4	17.0	38.0	4	17.0	34.5
10 to 15.....	7	12.4	32.7	7	12.5	29.6
5 to 10.....	5	8.2	33.2	5	8.0	31.4

(c) *Correspondence Between Immediate and Deferred Reproduction.*

But, after all, the important question in education is not so much what can be done by pupils immediately after they have just been taught, but what they can do some time afterwards. Do they remember what they once knew, and how far can they apply their knowledge? To the second of these questions I hope to give an answer when dealing with the results of the test on new material. Let me turn for a while to the first, and let me break it up into a number of constituent questions. The boys gain certain marks immediately after teaching and learning. What do they gain a week later, and, more important still, what do they gain two months later?

In a rough way we can find the answers to our questions in the following table:

Table XXVII, showing the work of the Inductive and Deductive Groups compared, section by section, in the Tests of Immediate and Deferred Reproduction.

Deductive Group.

Marks for immediate reproduction.	No. of boys.	Average Marks. Immediate reproduction test.	Deferred reproduction, first test.	Deferred reproduction, second test.
40 and over	5	41.4	38.6	38.4
35 to 40	3	37.7	38.0	32.7
30 to 35	4	30.8	28.3	25.8
25 to 30	3	26.7	22.3	23.0
Below 25	1	22.0	20.0	19.0

Inductive Group.

Marks for immediate reproduction.	No. of boys.	Average Marks. Immediate reproduction test.	Deferred reproduction, first test.	Deferred reproduction, second test.
40 and over	0	....	....	....
35 to 40	6	37.0	33.7	31.7
30 to 35	3	32.0	29.3	27.0
25 to 30	6	27.3	22.8	24.5
Below 25	1	20.0	17.0	24.0

The conclusions seem clear. The Inductive Group contains no boys at all equal to the highest section of the Deductive Group. The best boys in the Inductive Group correspond to the second section of the Deductive Group, but even then they are inferior to that section, both in the immediate and deferred tests. The work done in immediate reproduction may be very well taken as representative of what the work will be later on in exercises of this kind, for the various sections into which the groups are divided

retain their relative positions throughout the whole experiment. Calculated exactly, the correlation coefficients between the results of Immediate Reproduction and those of the first Deferred Reproduction Test in the Deductive Group is + .804, and between Immediate Reproduction and the second Deferred Reproduction Test (two months later) is + .859. The corresponding figures for the Inductive Group are + .616 and + .619.

Summarizing the results and treating the groups as wholes, the averages and variabilities are as follow:

Table XXVIII, showing the work of the Inductive and Deductive Groups compared in the Tests of Immediate and Deferred Reproduction.

	Immediate reproduction.	First deferred reproduction.	Second deferred reproduction.
Deductive Group:			
Average mark.....	34.2	32.5	30.1
M. V.....	6.0	6.3	7.0
Inductive Group:			
Average mark.....	31.4	27.8	27.6
M. V.....	4.5	5.1	4.1

The Deductive Group has outdistanced the Inductive Group quite clearly, both in immediate and deferred reproduction, not only in positive marks, for, perhaps, I ought to add, it has also made fewer 'bad errors.' It is the third result in which this has been found to be the case. We shall, therefore, again have to admit the contention urged against inductive methods in the earlier chapters of this monograph. We must certainly conclude that, in examinations on precisely what has been taught or learnt, children taught by what we have called deductive methods may be more successful than children taught

inductively. Also we see that children need not be young to be taught successfully by deductive methods. Let us now turn, however, to the Test of Application to New Material and see whether the same relation between the two groups holds there.

(*d*) *Results of the Test on New Material.*

We have seen that for purposes of immediate, and even of deferred, reproduction the more mechanical method has shown itself superior to the less mechanical. Is the same relationship retained between the two groups when the test is no longer one of simple reproduction, but requires a transfer of knowledge or method to analogous material? We can say at once that the same relation is not maintained. The inductive group now comes to the front, but the difference between the means of the two groups is a small one and the variability of the averages is high. The deductive group scores an average mark of 20.5 (mean variation 5.9), and the inductive group an average mark of 21.1 (mean variation 4.4). But let us look a little more closely into the composition of these averages:

Table XXIX, showing the work of the Inductive and Deductive Groups compared in Immediate Reproduction and in the Test on New Material.

	Deductive Group.			Inductive Group.		
Marks in immediate reproduction.	No. of boys.	Average Marks for Immediate reproduction.	Average Marks for New material.	No. of boys.	Average Marks for Immediate reproduction.	Average Marks for New material.
Over 35.....	8	40.0	23.0	6	37.0	24.8
30 to 35.....	4	30.8	17.8	3	32.0	19.3
25 to 30.....	3	26.7	19.7	6	27.3	17.5
Below 25.....	1	22.0	12.0	1	20.0	15.0

The figures certainly suggest a superiority on the side of the inductive group in three of the corresponding sections into which the groups are divided; and the regular decline of the figures in both groups (with the exception of the average of 19.7 in the third section of the Deductive Group) would appear to indicate that there is a general tendency in favor of correlated transfer in the Inductive rather than in the Deductive Group. The coefficient of correlation between the results of the Inductive and Deductive Groups, when tested on new material, is, however, not very high. With high variability as well, this involves a high probable error. So that we may conclude in this case merely that the Inductive Group does better work on the whole than the Deductive Group, but we have not the usual statistical justification that there is a strong general tendency in that direction. We shall, however, hardly feel disposed to attribute the superiority of the Inductive Group to chance, since in every one of the five experiments, with different teachers, with children of different ages, of different abilities and of different sexes, we have found the inductively taught group the more competent when tested on the power of application to new material.

IX. GENERAL SUMMARY.

In five different schools in different parts of London, attended by children varying in social class, experiments have been made to test the relative values of 'inductive' and 'deductive' methods of teaching as applied to geometrical definition. Both girls and boys, of ages ranging from 8 to 15 years, were set to do the work. The main problems were two in number. In the first place, an attempt was made to discover which of the two methods gave the better results when the children were tested on *precisely what they had been taught or had learnt.* In the second place, an endeavor was made to find out which of the two methods gave the better results when the children were tested on *new material.*

The answer to the first of these two questions was not the same in all of the five schools tested. In three of them, two of the three boys' schools and one of the two girls' schools, the conclusion was unambiguously in favor of the 'deductive and memoriter' method. This was the case with the younger and less proficient boys and girls, and at first sight it looked as if age were an important factor in the production of this result, but the same result was obtained with a class of boys who were much older, so that age was certainly not the *only* factor of differentiation. In two classes, the oldest class of boys and the oldest class of girls who did the work, the inductive method

was just as successful as the 'deductive,' even for purposes of exact reproduction, immediately afterwards, of what had been taught or learnt. There were some indications that the children inductively taught lost rather less of what they had known than those deductively taught when they were tested *some time afterwards;* but, on the whole, the tests of *deferred* reproduction gave the same comparative results as those of *immediate* reproduction. The importance of this consideration in testing school methods *where exact reproduction is required* is obvious..

The answer to the second of the two main issues was the same in all of the five schools tested. The children who were taught 'inductively' did better work than those taught 'deductively' in every case when they were required *to apply themselves to new material.*

This research, therefore, offers an experimental justification of what are known, among teachers, as 'intelligent' methods of teaching, and of the superior 'transfer' effect of certain methods.

Many pedagogical corollaries may be drawn from the experiments, but it will be sufficient in this place to emphasize a consideration already alluded to in the body of the text.

Examinations, whether internal, that is, conducted from within by the school authorities, or external, that is, conducted by external educational authorities, should always include questions on subject-matter which is *not* identical with that set down in the syllabuses of instruction *if the examination is to test good method in teaching.* But if the tests are to serve any useful pedagogical purpose, the new mate-

rial, though it should not be identical, ought to be analogous to that which has been dealt with in the school curriculum. Questions on *new analogous material* are probably the best questions of all (if the same set of questions be required to serve a double purpose), for they test, with fair adequacy, whether the work set down in the syllabuses has been efficiently done, and they also test, with admirable adequacy, whether the methods by which the school work has been done were such as to give the pupil power to apply his knowledge.

INDEX

'Bad' errors, meaning of, 36.
and mechanical method, 51.
Chance or Variability, 7.
Children's Definitions, spontaneous, 27, 28, 57, 71, 72, 73, 74, 102, 103, 120, 121.
after teaching and learning, 76, 77, 78, 80, 81, 82, 125, 127.
of new analagous material, 40, 83, 84, 85, 86, 88, 90, 109, 111, 130, 131, 132.
Circle, definition of diameter of, 29, 32.
drawing of diameter of, 24.
Classes taking the experiment, 20, 23, 55, 69, 100, 119.
Co-operation of Teachers, 4.
Correlation coefficients, 9, 10, 50, 62, 66, 91, 96, 115, 137.
value of, 30.
Deductive Method, method of learning by, 37.
method of testing, 19.
shown to be the better, 46, 93, 134, 136, 137.
Deferred Reproduction, 44, 47, 63, 65, 92, 107, 124, 135.
Definitions, 'real,' 26, 27.
arguments in favor of deductive treatment of, 17.
arguments in favor of inductive treatment of, 18.
units of marking of, 28, 29, 41, 42, 43.
as learnt deductively, 31.
as learnt inductively, 32.
children's spontaneous, after teaching and learning, of new analagous material, see Children's Definitions.
of diameter of circle, 29, 32.
of hexagon, 113.
of oblong, 29, 32.

of pentagon, 113.
of rhomboid, 42, 106, 123.
of rhombus, 41, 105, 123.
of square, 25, 28, 32, 35.
of diagonal of square, 39, 106, 124.
of tangent to circle, 113.
of trapezium, 42, 106, 123.
of triangle, 28, 32.
Demonstrative Geometry, introduction to, 17, 27.
Diagonals of Squares, 39, 106, 124.
Diameter of Circle, definition of, 29, 32.
drawing of, 24.
Durability of knowledge, 19, 37, 48, 49, 64, 66, 95.
Education, method of settling disputed questions in, 16.
Educational Science, 3.
Equal groups, how formed, 30, 45, 56, 61, 92, 133.
use of, 52.
Errors, 'bad,' meaning of, 36.
and mechanical method, 51.
Errors, method of correcting inductively, 34.
'probable errors,' method of determining, 8-10.
in spelling not counted, 43.
Experiment, use of, 16.
Experimental Pedagogy, 1.
Geometrical Definitions, see Definitions.
Geometrical Teaching, alleged cause of 'chaos in,' 18.
Geometry, Demonstrative, introduction to, 17, 27.
Groups, equal, how formed, 30, 45, 56, 61, 92, 133.
use of, 52.
Hexagon, definition of, 113.
drawings of, 108.
Immediate Reproduction, 19, 44, 47, 63, 65, 92, 107, 124, 135.
Inductive Method, an objection to, 14.

arguments for, 18.
method of learning by, 32.
method of correcting by, 34.
method of testing, 19.
shown to be the better, 51, 52, 64, 66, 67, 98, 116, 117, 138.
Intelligence, meaning of, 18.
test of, 38, 129.
training of, 53.
Knowledge, durability of, 19.
relation between quickness and permanence, 37, 48, 49, 64, 66, 95.
Marks, positive and negative, 36, 38.
Material, 'new,' see 'New Material.'
Negative marks, 36, 38.
'New material,' meaning of, 53, 54.
application to, 19, 38, 50, 67, 96, 107, 116, 129, 138.
'New methods,' general tendency of, 13, 14.
alleged disadvantages of, 14.
Novelty, influence of, 70.
Oblong, definition of, 29, 32.
drawings of, 24.
Pedagogy, Experimental, 1.
Pentagon, definition of, 113.
drawings of, 108.
Positive Marks, 36, 38.
Practice versus Theory, 11, 13, 15.
'Probable Errors,' method of determining, 8-10.
Reproduction, deferred, 44, 47, 63, 65, 92, 107, 124, 135.
immediate, 19, 44, 47, 63, 65, 92, 107, 124, 135.
Rhomboid, definition of, 42, 106, 123.
drawings of, 39.
Rhombus, definition of, 41, 105, 123.
drawings of, 39.

School Classes taking the experiment, 20, 23, 55, 69, 100, 119.
Science, Educational, 3.
'Science' of Education, 4.
Spelling errors not counted, 43.
Spontaneous definitions, 27, 28, 57, 71, 72, 73, 74, 102, 103, 120, 121.
Square, definition of, 25, 28, 32, 35.
    diagonals of, 39, 106, 124.
    drawings of, 24.
Tangent to Circle, definition of, 113.
    drawings of, 108.
Teaching, the divergence of Theory and Practice, 11.
    breach between Theory and Practice, 13, 15.
    unintelligent, reaction against, 70.
Teachers, co-operation of, 4.
Theory versus Practice, 11, 13, 15.
Time taken for the exercises, 44, 59, 75, 104, 106, 122.
Trapezium, definition of, 42, 106, 123.
    drawings of, 139.
Unintelligent teaching, reaction against, 70.
Unsophisticated material, 20.
Variability or Chance, 7.

WELLCOME LIBRARY INSTITUTE

Educational Psychology Monographs

Edited by Guy Montrose Whipple

WARWICK & YORK, Inc.

Baltimore, U. S. A.

Moto-Sensory Development

Observations on the First Three Years of a Child.

By

GEORGE V. N. DEARBORN

Price:
12*mo*,
215 + *vi pages*,
frontispiece.
$1.50.

Few subjects are of greater interest to the parents of young children or to school teachers with the truly scientific spirit of their profession than the evolution of a child's mechanism of efficiency. To the psychologist, and to a less extent to the physiologist, acquaintance with the average course of this human unrolling is clearly a technical necesstiy. All these surely should welcome every competent new account of the first three years of human life.

This book, as its name implies, discusses both the motor and the sensory development of an average child. It consists of careful observations of the steps in individual evolution with the addition of numerous notes and brief theoretic discussions of the observations. The chief emphasis has been put on the beginnings of voluntary movement and on the forerunning phenomena. These are considered from both the physiologic and psychologic points of view.

The affective side of child-development is more fully treated than are the purely intellectual processes, although the moto-sensory evolution of ideation as exhibited in learning to talk is as amply considered as circumstances allowed and as was expedient.

A feature of the book is a careful chronologic epitome of the observed development, perhaps more detailed than in any work since the pioneer treatise of Preyer. This is given in two tables of considerable length, one of them arranged alphabetically and the other by weeks. For purposes of reference these tables will be found of value.

Throughout the book there is continual reference to the temporal and other relationships of mental development as noted in similar accounts by Preyer, Darwin, Shinn, Moore, Major and others. These notes facilitate the use of the book for pedagogical purposes, and they also enable parents to judge more accurately of the natures of their children in comparison with the average.

WARWICK & YORK, Inc., BALTIMORE, MD.

Spelling Efficiency in Relation to Age, Grade and Sex, and the Question of Transfer

An Experimental and Critical Study of the Function of Method in the Teaching of Spelling.

By

J. E. WALLACE WALLIN

Price:
12mo, cloth, vi, 91 pages.
$1.25.

There are few elementary school subjects in which inefficiency is more surely detected and reprobated in later life, and in the teaching of which the elementary schools are charged with more extravagant waste of time, than spelling. 7.22 per cent. of the time of the child in the elementary schools in ten of our largest cities is devoted to the study of spelling, and yet the complaint continues to be almost universally voiced that the elementary and secondary school graduates have not learned how to spell.

School superintendents and teachers have felt the justice and sting of these criticisms, and have attempted to provide a remedy either by increasing the time devoted to spelling or by changing the methods of teaching. The results, however, have not in all cases proved satisfactory.

Dr. Wallin, who has been offering courses in educational psychology and the principles of teaching in schools of education for a number of years, points out briefly in this monograph some of the fallacies involved in the exclusive use of the incidental method of teaching spelling, based upon the psychological principles which condition the reduction of mechanical subject-matter to the plane of automatism (spelling is of an instrumental nature). By means of the results of the very researches made in the past to demonstrate the adequacy of the incidental method, it is shown that its use has not justified the claims made in its behalf. On the other hand, the superiority of a spelling drill technique, based upon the laws of habit formation, is shown, partly by the author's own investigation and partly by the results of a thoroughgoing application of the method under control conditions during four years in a large school system.

The book also discusses the relation of spelling efficiency to age, grade and sex; the facts derived from the tests are supported by numerous tables, a number of practical conclusions are offered, and a bibliography is appended.

When Should a Child Begin School?

An Inquiry Into the Relation Between the Age of Entry and School Progress.

By

W. H. WINCH

Few educational questions have excited more general interest in recent years than that of the age at which children should commence their attendance at school. On the one side we have the rule-of-three conclusion, felt rather than expressed as an inference, that the more teaching the child gets and the sooner he begins school the more progress he is sure to make. On the other we have had a strong feeling, now growing in intensity and range, that attendance in school, particularly in England, begins too early and that there is an educational disadvantage in commencing as soon as the children of Great Britain do. While this investigation by Mr. Winch has special reference to England, where the school life begins at a much earlier period than in either America or Germany, the results set forth by the author are of vital interest to all who have to do with the education of children.

The effect of age of entry is considered from several points of view: 1. Does early entry at school enable the pupil to make more rapid advancement in school standing than entry at a later age? In other words, in a given grade are those pupils who entered school earlier found to constitute the younger portion of the class? 2. In the same grade some pupils may be doing work of a high degree of efficiency, others work of an inferior quality. To what extent does early entry correlate with high efficiency when tested by examinations? 3. How far does early entry depend upon social circumstances? 4. What is the influence of early entry upon the subsequent behavior of pupils and upon their attentiveness to school work?

The results of Mr. Winch's inquiry are now published for the first time. Some of them have been privately circulated, and a few of the tables, together with the methods employed, were discussed some years ago at a meeting of the Inspectors of the Education Committee for London.

Price:
12mo, cloth,
108 *pages.*
$1.25.

WARWICK & YORK, Inc., BALTIMORE, MD.

Mental Fatigue

"Die Geistige Ermüdung."

By

MAX OFFNER,

Translated from the German by

GUY MONTROSE WHIPPLE

Price:

12mo, cloth, viii, 133 *pages.*
$1.25.

This noteworthy monograph is a comprehensive exposition of the nature of mental fatigue, of the methods proposed for measuring it, and of the results that have thus been obtained, with special reference to their application to classroom problems.

The text is an amplification of a lecture delivered before the Munich association of gymnasial teachers, and its primary purpose is not to contribute to the experimental investigation of fatigue, but to inform and to interest teachers.

The following are among the topics discussed: The nature and forms of fatigue, the symptoms of fatigue, the measurement of fatigue by physiological and by psychological methods, the factors other than fatigue that affect efficiency of mental work—practice, adaptation, warming-up, spurts, enthusiasm, etc.—and the laws of fatigue.

In considering the application of these laws to school-room problems, attention is given to the dependence of fatigue upon individual differences, upon age, puberty, the length of lesson periods, the number of lessons per day, the day of the week, the introduction of various rest pauses (recesses, holidays, vacations, etc.), change of occupation, the fatigue coefficient of the different studies, also to hygienic arrangement of the school program and other practical problems. A selected bibliography closes the monograph.

The translation is offered with the conviction that it will meet a very general demand on the part of the teacher of educational psychology and of the hygiene of instruction for a clear and systematic presentation of the problem of mental fatigue and its relation to school work.

WARWICK & YORK, Inc., BALTIMORE, MD.

Relative Efficiency of Phonetic Alphabets

An Experimental Investigation of the Comparative Merits of the Webster Key Alphabet and the Proposed Key Alphabet Submitted to the National Education Association.

By

GUY MONTROSE WHIPPLE,

Price:
8vo, 60 pages.
35c. paper binding.

This monograph will exert a two-fold appeal to those who aim to keep abreast of present-day movements in education. First, in that it offers an excellent example of the application of the experimental method to a pedagogical problem, and in this respect will take its place as a contribution to experimental pedagogy; secondly, in that it deals with an important topic just now a matter of general discussion in educational circles.

The National Education Association has under consideration the adoption of a new key-alphabet for phonetic notation. The merits of the proposed alphabet have been the subject of extensive and lively debate, but no one has hitherto done the obvious thing and tried out the new alphabet under experimental conditions. This Dr. Whipple has accomplished, and the results will interest every teacher who uses a phonetic alphabet in his class work as well as every educator who believes with the author that, in the school as well as in other realms of life, "you can tell by trying."

In view of the fact that the subject of phonetic alphabets will be given much attention by educators during the next year, this work is offered at a price which will place it easily in reach of teachers in city and rural schools, and also the members of clubs and reading circles.

WARWICK & YORK, Inc., BALTIMORE, MD.

Backward and Feeble-Minded Children

A Series of Studies in Clinical Psychology.

By
EDMUND B. HUEY

Price:
12*mo*,
200 *pages*,
illus.
$1.40.

Each of the more populous States has several thousand mental defectives, large numbers of whom are attending the public schools. They usually make little progress and are distressingly disturbing factors in the regular classes. In Germany, and recently in France, and in some of our own cities, these children are being placed in special classes or in special schools, according to the degree of defect. Teachers and school experience immediate relief, and the children themselves are the greatest beneficiaries. All the schools have these defectives, and the problem of recognizing and caring for them is an immediately pressing one in all our cities, towns and rural districts.

Following a year in the clinics of Paris, Dr. Huey's position at Lincoln for nearly a year and a half involved making a mental examination of each new admission to this, one of the largest state institutions for the feeble-minded.

As research psychologist to the institution Dr. Huey made careful psychological study of 35 selected cases which represent the transition zone between feeble-mindedness and non-feeble-mindedness. These are just the border cases that puzzle the school principal or the clinician. In this volume he presents case after case representing various types and groups of backward and feeble-minded children. The mental and physical characteristics of each child and the salient features of different groups are clearly stated, with charts which graphically present the results of various measurements and tests.

The methods of making examinations and tests and of making observations and gathering data needed for the interpretation of any given case are illustrated in detail. The concreteness of the material and the abundance of illustrative examples will be appreciated by all, and make the studies intelligible even to those unfamiliar with psychological technique.

WARWICK & YORK, Inc., BALTIMORE, MD.

Experimental Studies of Mental Defectives

A Critique of the Binet-Simon Tests and a Contribution to the Psychology of Epilepsy.

By

J. E. WALLACE WALLIN, Ph.D.

About 150 *pages.* $1.25.

The Binet-Simon tests have been hailed by popular writers and even by some scientific workers as a wonderful mental X-ray machine, which will enable us to dissect the mental and moral mechanisms of any normal or abnormal individual. But those who have had extensive experience with these tests know that, despite their very great practical value, they have numerous imperfections and definite limitations. These imperfections and limitations can be made known only by thoroughgoing trial on large groups of individuals by expert investigators. Dr. Wallin is well qualified by training and experience to undertake this work, and he has presented in this, the seventh of the series of Educational Psychology Monographs, a systematic critical study of the results of the Binet Scale when applied to a colony of epileptic children, and has included a guide for the conduct of the tests.

In the course of his study certain facts have been revealed concerning the mental status of the epileptic which should interest the schoolman as well as the alienist and the physician, for epileptic children constitute a numerous class which grades nearer the public school laggard than do feeble-minded children, and which cannot be reached by the cut-and-dried methods of the schools, but requires a special educational regime. Moreover, epilepsy, despite the investigations of many alienists, still remains a little understood pathological condition with marked disturbance of mentality.

We commend this contribution to the attention of physicians, alienists and all schoolmen who are interested in the scientific examination of mental deficiency.

WARWICK & YORK, Inc., BALTIMORE, MD.

Variations in the Grades of High-School Pupils

By
CLARENCE TRUMAN GRAY.

12*mo*,
Cloth ca,
120 *pages*.
$1.25.

Ten years ago no serious attempt had been made to study scientifically the relative merits of various systems of grading students, despite the fact that statistical methods for undertaking such studies were fully available and that grading plays so large a rôle in the school career of hundreds of thousands of school children. In the last five years, however, this inviting field has been the scene of numerous important investigations, so that we have at least arrived at a better understanding of the nature of the problem and of the general line along which progress must be made.

In the present monograph Mr. Gray reports the methods and results of his investigation of one phase of the general problem, viz., the nature, degree and causes of the variations occurring in the grades of high-school pupils. The general aim of his study is to base an educational investigation upon school grades. It is usually argued that such marks are inaccurate, that they are complex, that they are not scientific, and, above all, that it is impossible to measure mental traits by such cold statistics as grades afford. In direct contrast to these arguments stands the fact that all promotions from the kindergarten through the university are based upon this so-called inaccurate, complex, unscientific and cold estimates of progress and achievement. One of the most vital and fundamental principles of any school system is its plan of promotions, and because of the close relation between promotions and grades there is the most urgent need that schoolmen become interested in the problems of grading. *Variations in the Grades of High-School Pupils* should interest all teachers, and more particularly all school administrators, because the author not only shows clearly how unreliable are the grades commonly given by teachers, and makes evident the need of instruction and training in grading, but also presents a relatively simple method by means of which any high-school principal can study the condition of the grading in his own school and take due steps to remedy the faults that he may find.

WARWICK & YORK, Inc., BALTIMORE, MD.

How I Kept My Baby Well

By

ANNA G. NOYES.

12mo, Cloth, Illustrated, ca, 180 pages. $1.25.

The fact that the *Journal of Educational Psychology* has defined its scope to include the consideration of child psychology and hygiene justifies the inclusion in the allied series of *Educational Psychology Monographs* of the material set forth in the present volume.

Mrs. Noyes has made a contribution of real interest to physicians and nurses, to mothers and fathers, and to students of childhood generally. The value of her work is twofold. On the one hand, it points the way to a method and type of observation that any intelligent mother can undertake with profit to herself and to others, and in so far disproves the contention of some critics of the child-study movement that observations of young children by their own mothers can never yield data of real value; on the other hand, it furnishes generalizations in the shape of principles or rules governing feeding, clothing and the general control of infant development that will be of direct utility to those who, like the author, face that vital problem—how to keep the baby well. Mrs. Noyes has displayed commendable caution in drawing these generalizations. It is not asserted that what applied to her own baby will apply invariably to any other baby, but only that it undoubtedly will apply to many babies, and that her method of attacking the problem is, at any rate, a method that other mothers may follow to advantage when confronted with the same situation.

The conservation of human life by the reduction of infant mortality is a noble undertaking, and it is hoped that this little contribution may in some measure further that undertaking.

The volume is profusely illustrated. The author and Mr. Noyes followed the life of the child through his first two years with a camera just as faithfully as the mother followed him with her charts and memorandum pad. As a consequence there appear as illustrations more than sixty pictures of the baby, most of them full-page cuts. The book also contains some forty or fifty full-page charts. Both photographs and charts greatly enhance the value of the book.

WARWICK & YORK, Inc., BALTIMORE, MD.