CERTAIN ABILITIES FUNDAMENTAL TO THE STUDY OF GEOMETRY¹

J. H. MINNICK

University of Pennsylvania

Success in the formal demonstration of a theorem of geometry depends upon, at least, four abilities, namely:

1. The ability to draw a figure for the theorem.
2. The ability to state correctly and accurately the hypothesis and conclusion of the theorem.
3. The ability to recall additional facts about a figure when one or more facts are given, and
4. The ability to select from the available facts those that are necessary for a proof, and to arrange them so as to arrive at the desired conclusion.

The purpose of this investigation is three-fold:

1. To determine the relation of each of these four abilities to teachers' marks. This should, in turn, determine either the extent to which teachers value these abilities, or the degree to which they are able to base their marks on the things which they do value.
2. To determine the extent to which these abilities are developed in the various schools included in this investigation.
3. To develop tests which may be used for the purpose of diagnosis; that is, for the purpose of determining whether or not the weakness of a class is due to the lack of development of one or more of these abilities.

For these purposes four tests, (to be known as A, B, C, and D), have been developed, each testing one of the abilities in question. An effort has been made so to arrange each test that it will measure only one of the four abilities. Thus, Test A consisted of geometrical exercises for which the pupil was asked to draw the figure. In Test B the exercise was stated, and the figure drawn. The pupil was then asked to state the hypothesis and conclusion. In each exercise of Test C a figure was drawn, and certain facts about it were given, and the pupil was to add as many more facts as he could. The purpose of Test D was to measure the pupil's ability to select and organize facts so as to produce a proof. In order to eliminate

¹A complete report of this investigation will be issued in book form later.

the factors involved in the other three tests the figure was drawn, the hypothesis and conclusion were started, and a list of facts, including those necessary to the proof, was given. The pupil was then asked to give the proof.

This investigation was limited to classes which had recently completed the first two books of geometry. If more advanced classes had been included, it would have been necessary to limit the investigation to these advanced classes, or to use a separate series of tests. The time required in developing a second series of tests made the latter method of procedure impracticable. To limit such an investigation to the advanced classes is to deal with a specially selected group of pupils since extensive elimination had taken place in the earlier work. On the other hand, if the abilities in question are of value, it seems that they should be fairly well developed during the first two books of geometry. Also if the tests are to be used for the purpose of diagnosis it should be possible to apply them fairly early in the study of the subject, while there is yet time to remedy any difficulties which may be found to exist. Hence the tests were given to only those classes which had recently completed the first two books of geometry.

Each of the four tests was given to not less than one thousand pupils. In all, the tests were given in sixty-three schools, distributed throughout the East and the central West. The wide distribution of schools and the variation in the nature of the schools insured the inclusion of nearly every type of high school pupil in the investigation. Clearly it was impossible for the author to conduct the tests in person. However, the nature of the tests made it possible to formulate simple and definite directions which insured a fair degree of uniformity. In order to be sure that the pupils would understand these directions, the author tried them out with high school classes, and made any necessary changes before placing them in the hands of teachers. In only two cases did the returns from any of the schools indicate a variation from the directions, and the papers from these schools were rejected.

With the large number of schools so widely distributed it was impossible to prevent some variation in physical conditions, such as weather, noise, etc. However the teachers were asked to report any disturbing elements which might influence the results of the tests. From these reports it would seem that such variations were not serious.

Two scores were kept for each test. One was based on the number of necessary statements (or parts of figures in the case of Test A), given correctly, and the other was based on the number of incorrect and unnecessary statements given. The former will be referred to as the *positive* score, and the latter as the *negative* score. The author knows of no valid method of combining these two elements and certainly for the purpose of diagnosis it is desirable that they should not be combined. Weighted positive values have been assigned to the exercises of each test. These values were based upon the average per cent. of necessary statements given correctly for each exercise by all pupils taking the test, and they were so selected that a perfect score for a test would be one hundred. It did not seem possible to assign weighted negative values to the various exercises, as there was no upper limit to the number of errors that could be made, and there was, therefore, no basis for comparison. Hence the total number of unnecessary and incorrect statements in a paper was taken as the negative score.

In order to insure uniformity of scoring all papers were marked by the author. In each test the first exercise was scored for the entire school, then the second, and so on, until all the exercises had been marked for a given school. Where an exercise admitted of a variety of solutions, a copy of each solution found among the papers, together with a memorandum of how it was scored, was kept as a guide in all similar cases.

In order to determine the relation between the teachers' marks and the abilities in question, the coefficient of correlation between the school and test grades was computed. For this purpose it was impossible to combine the data from the various schools as they had different standards for grading. Hence the coefficient was computed separately for each school. In order to obtain this coefficient, it was necessary to array the pupils of a school according to the grades which they received for the first two books of geometry and again according to the test scores. The latter array of pupils had to be based on both the positive and negative scores. In order to meet this condition the pupils were arrayed according to the positive scores, and again according to the negative scores. The two ranks thus assigned to each pupil were added, and the pupils were again arrayed according to these sums. The formula :²

²See WILLIAM BROWN, *The Essentials of Mental Measurement*, pp. 42-53.

$$\rho = 1 - \frac{6 \sum d^2}{N(N^2 - 1)}$$

was then used to compute the coefficient and the result obtained by this formula was corrected by the formula

$$r = 2 \sin \left(\frac{\pi}{6} \rho \right)$$

TABLE I
Coefficients of Correlation for Test A.

School	r	P. E.	Relation of r to 3 P. E.
XXIII	0.313	0.068	r > 3 P. E.
XXV	0.395	0.071	r > 3 P. E.
XXXIII	0.111	0.180	r < 3 P. E.
XXXV	0.628	0.089	r > 3 P. E.
XXXVI	0.303	0.125	r < 3 P. E.
XXXIX	0.426	0.140	r > 3 P. E.
XL	0.487	0.079	r > 3 P. E.
XLI	0.303	0.093	r > 3 P. E.
XLII	0.697	0.051	r > 3 P. E.
XLIII	0.436	0.088	r > 3 P. E.
XLIV	0.364	0.085	r > 3 P. E.
XLV	0.364	0.086	r > 3 P. E.
L	0.588	0.123	r > 3 P. E.
LI	0.240	0.188	r < 3 P. E.
LII	0.528	0.074	r > 3 P. E.
LIII	0.588	0.055	r > 3 P. E.
LIV	0.150	0.147	r < 3 P. E.
LVI	0.199	0.085	r < 3 P. E.
LIX	0.688	0.069	r > 3 P. E.
LX	0.578	0.087	r > 3 P. E.
LXI	0.548	0.099	r > 3 P. E.
LXII	0.436	0.053	r > 3 P. E.
LXIII	0.292	0.069	r > 3 P. E.

TABLE II
Coefficients of Correlation for Test B.

School	r	P. E.	Relation of r to 3 P. E.
XIV	0.467	0.127	$r > 3 \text{ P. E.}$
XV	0.548	0.127	$r > 3 \text{ P. E.}$
XVI	0.447	0.079	$r > 3 \text{ P. E.}$
XVII	0.140	0.115	$r < 3 \text{ P. E.}$
XVIII	0.188	0.056	$r > 3 \text{ P. E.}$
XIX	0.230	0.076	$r > 3 \text{ P. E.}$
XX	0.219	0.051	$r > 3 \text{ P. E.}$
XXI	0.323	0.075	$r > 3 \text{ P. E.}$
XXV	0.568	0.059	$r > 3 \text{ P. E.}$
XXX	0.395	0.081	$r > 3 \text{ P. E.}$
XXXI	0.271	0.090	$r > 3 \text{ P. E.}$
XLVII	0.668	0.065	$r > 3 \text{ P. E.}$
XLVIII	0.261	0.096	$r < 3 \text{ P. E.}$
XLIX	0.538	0.054	$r > 3 \text{ P. E.}$
LVII	0.588	0.075	$r > 3 \text{ P. E.}$
LVIII	0.344	0.086	$r > 3 \text{ P. E.}$

TABLE III
Coefficients of Correlation for Test C.

School	r	P. E.	Relation of r to 3 P. E.
VII	0.209	0.123	$r < 3 \text{ P. E.}$
VIII	0.548	0.030	$r > 3 \text{ P. E.}$
IX	0.178	0.080	$r < 3 \text{ P. E.}$
X	0.385	0.108	$r > 3 \text{ P. E.}$
XI	0.416	0.110	$r > 3 \text{ P. E.}$
XII	0.538	0.054	$r > 3 \text{ P. E.}$
XIII	0.333	0.090	$r > 3 \text{ P. E.}$
XXVIII	0.436	0.093	$r > 3 \text{ P. E.}$
XXXII	0.042	0.138	$r < 3 \text{ P. E.}$
XXXIV	0.487	0.058	$r > 3 \text{ P. E.}$
XXXVII	0.406	0.116	$r > 3 \text{ P. E.}$
XXXVIII	0.508	0.099	$r > 3 \text{ P. E.}$

TABLE IV
Coefficients of Correlation for Test D

School	r	P. E.	Relation of r to 3 P. E.
V	0.351	0.066	$r > 3$ P. E.
XXII	0.528	0.077	$r > 3$ P. E.
XXIII	0.209	0.078	$r < 3$ P. E.
XXIV	0.325	0.077	$r > 3$ P. E.
XXV	0.395	0.013	$r > 3$ P. E.
XXVI	0.303	0.035	$r > 3$ P. E.
XXVII	0.126	0.057	$r < 3$ P. E.
XXIX	0.568	0.056	$r > 3$ P. E.
XLVI	0.323	0.044	$r > 3$ P. E.

The coefficients of correlation for each of the tests are given in Tables I-IV. If the coefficient is greater than three times the probable error it has scientific significance. In the majority of the schools, this is true, and we may conclude that usually the teachers' marks do bear a positive relation to the abilities with which the investigation is concerned. In considering these coefficients it is to be remembered that the school grades should depend upon factors not included in this investigation, and therefore the correlation should not be perfect. In several cases the coefficients are fully as high as can be expected. However they are frequently low and in many cases they are so low that, if they can be taken as indices of the value of abilities, then the schools are scarcely justified in giving as much time to this phase of geometry as they now do.

TABLE V
Summary of Positive Scores

Test	Medians for all Pupils tested	25-per- centile	75-per- centile	Lowest Median by any School	Highest Median by any School
A	62.5	51.3	72.9	50.5	78.7
B	69.3	51.8	82.2	38.5	80.9
C	50.6	36.5	65.2	29.0	67.0
D	73.3	55.4	87.0	54.7	80.5

TABLE VI

Summary of Negative Scores

Test	Medians for all *Pupils tested	25-per- centile	75-per- centile	Lowest* Median by any School	Highest Median by any School
A	7 1	10.7	4 2	11.8	4 8
B	3 5	5.9	1 5	4 5	2 0
C	4 1	7 3	1 9	7.3	2.4
D	2 6	5.4	0.9	3 7	1.5

The median positive and negative scores made by all pupils taking each test are given in Table V and VI. Since we do not know that the tests measure the respective abilities in the same way, we can not compare the results of the different tests. Thus the low median score of Test C does not necessarily indicate that the pupils have less ability to recall facts about a figure than to do any of the things called for in the other tests. However, we may compare the results obtained by giving the same test in different schools. Such a comparison shows that some of the schools are developing these abilities much more effectively than others. In the case of each test the marks of some schools are quite satisfactory, while those of others are very low. This variation in achievement may be due, in part, to differences in local conditions, rather than differences in methods and in teaching ability. Nevertheless, it is difficult to see how local conditions could result in the extremely low scores made by some schools. If the abilities tested are essential to success in the study of geometry, then the results indicate that progress is almost impossible in some of the schools until these abilities have been further developed. On the other hand, the achievements of other schools indicate that it is altogether possible to develop these abilities to a fair degree during the study of the first two books of geometry.

*As the negative scores represent the numbers of incorrect and unnecessary statements, the larger numbers represent the lower scores.

Since, as we have seen, these abilities may be satisfactorily developed, but in many cases they are not, then if they are of value it is important that we have some means for determining when they are satisfactorily developed. It is believed that these tests or others similar to them will prove valuable for this purpose. The standards given in Tables V and VI furnish a means of comparison. If a class is above the median scores in these tables, and especially if it is near or above the 75-percentile marks, the teacher may be fairly sure that the abilities have been sufficiently developed to insure the success of the class. If, however, the class falls below the median score in any one of the tests, and especially if it falls near or below the 25-percentile mark, special effort should be made to develop the ability in question.