

CORRELATIONS BETWEEN BINET TESTS AND GROUP TESTS

W. T. ROOT

University of Pittsburgh

In the fall of 1920 the author supervised the testing of some 600 children in the schools of Monessen, Pa., with the Binet-Simon Tests, Stanford Revision. Early in 1921 it became possible, with the aid of Mr. Herman Gress¹ and Mr. Wade Blackburn,¹ to give a battery of mass tests to the same group. They consisted of the following: Otis Primary A (O. P. A); Otis Advance A (O. A. A); Haggerty Sigma I (H. Sigma I); Haggerty Delta I and II (H. Delta I and II); National A. I and B. I (N. A. I and N. B. I); Terman Group A (T. G. A); Menti-meter, (M.); Dearborn Series I and Series II (D. S. I and D. S. II); and Illinois I and II (Ill. I and Ill. II).

The usual statistical precautions were observed. All scores and correlations were rechecked. The greatest precautions were taken to secure uniform conditions throughout the testing. Only the splendid cooperation of the Monessen teachers made this possible. The mass tests were given weekly, on the same day at the same hour. The Dearborn Series I was given in two sittings. About 416 from the Grade I to Grade XII were given both the mass and the Binet tests. In correlating, Grades XI and XII are combined. Pearson's Product-Moment formula was used in correlating.

It has been assumed, in making the correlations, that the Binet Tests constitute the truest estimate of intelligence in so far as tests go. This may not always be the case with older (college) students as indicated in a recent article by De Camp,² but probably no one will take exception to the assumption that with children up to 15 or 16 years of age, the Binet Test constitutes the best single test estimate of the intelligence that can be made. Granting this, the correlation of any mass test with the Binet Test becomes of immense importance in estimating the value of the former. The administrator is anxious to know what mass test is most suitable for a particular grade or a partic-

¹ Mr. Gress is Superintendent of Schools, Monessen, and Mr. Blackburn is Supervisor of the grammar grades. They hope later to present a careful analysis of causes of variation in the correlations, and also an analysis of the causes leading to marked individual inconsistency in performance from test to test.

² De Camp, J. E.: Studies in Mental Tests. *School and Society*, Vol. XIV, pp. 253-258.

TABULATION OF CORRELATIONS

	Grade	Number	R	P. E. R
Binet with O. P. A.....	1	87	0.72	0.03
Binet with O. P. A.....	2	34	0.60	0.07
Binet with O. P. A.....	3	36	0.63	0.07
Binet with O. P. A.....	4	38	0.77	0.04
Binet with O. P. A.....	all	198	0.80	0.02
Binet with O. A. A.....	5	26	0.64	0.08
Binet with O. A. A.....	6	32	0.46	0.09
Binet with O. A. A.....	7	31	0.76	0.05
Binet with O. A. A.....	8	45	0.68	0.05
Binet with O. A. A.....	9	22	0.72	0.07
Binet with O. A. A.....	10	25	0.55	0.09
Binet with O. A. A.....	11-12	37	0.44	0.09
Binet with O. A. A.....	all	218	0.80	0.02
Binet with H. Sigma I.....	1	88	0.47	0.06
Binet with H. Sigma I.....	2	36	0.46	0.09
Binet with H. Sigma I.....	3	36	0.61	0.07
Binet with H. Sigma I.....	all	160	0.74	0.02
Binet with H. Delta I.....	1	88	0.71	0.04
Binet with H. Delta I.....	2	36	0.28	0.10
Binet with H. Delta I.....	3	37	0.57	0.07
Binet with H. Delta I.....	all	162	0.76	0.02
Binet with H. Delta II.....	3	36	0.62	0.07
Binet with H. Delta II.....	4	40	0.69	0.06
Binet with H. Delta II.....	5	25	0.58	0.09
Binet with H. Delta II.....	6	32	0.60	0.08
Binet with H. Delta II.....	7	31	0.82	0.04
Binet with H. Delta II.....	8	44	0.79	0.06
Binet with H. Delta II.....	9	22	0.44	0.12
Binet with H. Delta II.....	all	232	0.84	0.01
Binet with N. A. I.....	3	36	0.69	0.06
Binet with N. A. I.....	4	41	0.68	0.06
Binet with N. A. I.....	5	26	0.66	0.07
Binet with N. A. I.....	6	32	0.72	0.06
Binet with N. A. I.....	7	31	0.79	0.05
Binet with N. A. I.....	8	45	0.51	0.03
Binet with N. A. I.....	all	211	0.84	0.01
Binet with N. B. I.....	3	35	0.67	0.06
Binet with N. B. I.....	4	41	0.65	0.06
Binet with N. B. I.....	5	26	0.69	0.07
Binet with N. B. I.....	6	32	0.63	0.07
Binet with N. B. I.....	7	31	0.67	0.07
Binet with N. B. I.....	8	45	0.49	0.08
Binet with N. B. I.....	all	210	0.86	0.01
Binet with T. G. A.....	7	31	0.73	0.06
Binet with T. G. A.....	8	45	0.65	0.06
Binet with T. G. A.....	9	22	0.35	0.13
Binet with T. G. A.....	10	25	0.67	0.07
Binet with T. G. A.....	11-12	37	0.53	0.08
Binet with T. G. A.....	all	160	0.75	0.02
Binet with M.....	1	86	0.65	0.04
Binet with M.....	2	35	0.49	0.09
Binet with M.....	3	36	0.60	0.07
Binet with M.....	4	39	0.68	0.06
Binet with M.....	5	26	0.71	0.07
Binet with M.....	6	32	0.53	0.09
Binet with M.....	7	31	0.71	0.06
Binet with M.....	8	45	0.61	0.06
Binet with M.....	9	22	0.43	0.12
Binet with M.....	10	25	0.68	0.07
Binet with M.....	11-12	36	0.54	0.08
Binet with M.....	all	407	0.88	0.01

TABULATION OF CORRELATIONS (Continued)

	Grade	Number	R	P. E. R
Binet with D. S. I.	1	85	0.79	0.03
Binet with D. S. I.	2	35	0.40	0.10
Binet with D. S. I.	3	36	0.72	0.05
Binet with D. S. I.	all	156	0.79	0.02
Binet with D. S. II.	4	38	0.65	0.06
Binet with D. S. II.	5	28	0.66	0.07
Binet with D. S. II.	6	31	0.74	0.05
Binet with D. S. II.	7	31	0.77	0.05
Binet with D. S. II.	8	45	0.65	0.06
Binet with D. S. II.	9	22	0.47	0.11
Binet with D. S. II.	all	195	0.87	0.01
Binet with III. I.	3	36	0.62	0.07
Binet with III. I.	4	38	0.66	0.06
Binet with III. I.	5	26	0.75	0.06
Binet with III. I.	all	100	0.74	0.03
Binet with III. II.	6	31	0.56	0.08
Binet with III. II.	7	31	0.72	0.06
Binet with III. II.	8	45	0.65	0.06
Binet with III. II.	all	107	0.68	0.04
N. A. I with T. G. A.	7	31	0.76	0.05
N. A. I with T. G. A.	8	45	0.72	0.05
N. A. I with T. G. A.	all	76	0.79	0.03
N. B. I with T. G. A.	7	30	0.73	0.06
N. B. I with T. G. A.	8	45	0.74	0.04
N. B. I with T. G. A.	all	76	0.73	0.04
O. A. A. with T. G. A.	7	31	0.83	0.04
O. A. A. with T. G. A.	8	43	0.77	0.04
O. A. A. with T. G. A.	9	21	0.73	0.07
O. A. A. with T. G. A.	10	25	0.87	0.03
O. A. A. with T. G. A.	11-12	35	0.72	0.06
O. A. A. with T. G. A.	all	160	0.85	0.01
H. Delta II with T. G. A.	7	31	0.86	0.03
H. Delta II with T. G. A.	8	44	0.83	0.03
H. Delta II with T. G. A.	9	21	0.85	0.04
H. Delta II with T. G. A.	all	97	0.85	0.02
M. with T. G. A.	7	29	0.79	0.05
M. with T. G. A.	8	43	0.75	0.05
M. with T. G. A.	9	20	0.60	0.10
M. with T. G. A.	10	22	0.79	0.05
M. with T. G. A.	11-12	34	0.63	0.07
M. with T. G. A.	all	159	0.82	0.02
H. S. I with O. P. A.	1, 2, 3	160	0.51	0.04
H. S. I with D. S. I.	1, 2, 3	153	0.67	0.03
M. with O. P. A.	all	188	0.93	0.01
M. with O. A. A.	all	216	0.75	0.02
N. A. I with N. B. I.	3-8	207	0.94	0.00
N. A. I with M.	3-8	211	0.93	0.01
N. A. I with O. P. A.	3-4	74	0.67	0.04
N. A. I with O. A. A.	5-8	134	0.88	0.01
N. A. I with D. S. I.	3	35	0.72	0.06
N. A. I with D. S. II.	4-8	167	0.89	0.01
N. A. I with H. D. II.	3-8	209	0.89	0.01
H. D. II with N. B. I.	3-8	206	0.92	0.01
H. D. I with M.	1-3	153	0.87	0.01
H. D. II with M.	4-9	192	0.89	0.01
H. D. II with O. A. A.	5-9	154	0.86	0.01
H. D. I with D. S. I.	1-3	118	0.87	0.02
H. D. II with D. S. II.	4-8	172	0.86	0.01

ular group of grades. It is hoped the accompanying correlation tabulations are a step in this direction.

The following observations may be made directly from the tabulation:

1. *Grade I.* D. S. I correlates the highest with the Binet (0.79); O. P. A next (0.72); and H. Sigma I, last (0.46). (The Sigma test it will be recalled is a reading test and not strictly an intelligence test.) The D. S. I would seem then to be the most suitable for this grade but has the disadvantage of requiring 2 days to give, owing to its extreme length.

2. *Grade II.* Judged by their correlations with the Binet Test, none of the mass tests proved satisfactory with the Grade II. The O. P. A has the highest correlation (0.59); the M., next (0.49); H. Sigma I, next (0.45); the D. S. I, next (0.40); and the H. Delta I, last (0.28).

3. *Grade III.* All of the mass tests are here more satisfactory, giving a correlation coefficient within the neighborhood of 0.60. The D. S. I is apparently most suitable (0.71); with the N. A. I a close second (0.68). The H. Delta I is least suitable (0.57). It will be noted that H. Delta II yields a higher coefficient (0.62) than H. Delta I.

4. *Grade IV.* The O. P. A, (0.77), is decidedly higher than the next most suitable test, the H. Delta II (0.69). The remaining tests, it will be noted, all lie within the 0.60s.

5. *Grade V.* The Ill. I gives the highest coefficient (0.75), with the M. a little below, (0.71). The remainder of the tests lie within the 0.60s except the H. Delta II whose coefficient falls to 0.58.

6. *Grade VI.* There is a marked difference in the correlations for this grade. The D. S. II has a coefficient of 0.74 with the N. A. I next (0.72), while the O. A. A falls lowest (0.46).

7. *Grade VII.* The highest correlation is with the H. Delta II (0.82), with the N. A. I a little lower (0.79). All of the correlations are high for this grade, lying within the 0.70s, with the exception of N. B. I (0.67).

8. *Grade VIII.* The correlations for this grade cover a wide range. H. Delta I being highest (0.79); while N. A. I (0.51) and N. B. I (0.49) are the lowest. The National correlates highly with the Binet except for this one grade.

9. *Grade IX.* The correlations are here low, and the small number of cases make the P. E.s high. The O. A. A stands highest (0.72);

then a drop to a correlation of 0.47 with D. S. II. The T. G. A stands last (0.35).

10. *Grade X.* Only three of the mass tests given cover this grade. The Mentimeter stands highest (0.68); the T. G. A a close second (0.67); and O. A. A, decidedly lower (0.54).

11. *Grades XI and XII.* The M. and T. G. A each give a correlation of 0.53. The O. A. A falls to 0.43. None of the tests are as satisfactory as with the lower grades.

12. *Grades I to IV inclusive.* Considering uniformity of high correlation the O. P. A seems best for these grades. It should be noted though that the Otis falls low on the Grades I and III. No test is entirely satisfactory.

13. *Grades III to VI inclusive.* It is sometimes desirable to consider these grades together. D. S. I and II make the highest and most uniform correlations; while the N. A. I and N. B. I make a close second.

14. *Grades V to VIII inclusive.* For these grades D. S. II is most desirable with H. Delta II of nearly equal value.

15. *Grades VII and VIII.* If these two grades are grouped together, H. Delta II is seemingly far superior.

16. *Grades VII to IX inclusive.* With the increase of junior high schools this grouping is now frequent. The O. A. A is apparently best.

17. *Grades IX to XII inclusive.* Grouping these grades, the O. A. A is perhaps most satisfactory.

18. *Grades VII to XII inclusive.* Considering these five grades together, there is little choice between O. A. A, T. G. A, and M. The author favors the T. G. A because it is very easy to administer, requires but 35 minutes to give, and is the simplest to score.

19. It will be noted that when the grades are pooled and correlated with the Binet, a high correlation for "all" in the tabulation is not indicative for any particular grade.

20. *Grades I to XII inclusive.* The highest general tendency to correlation with the Binet is with the Mentimeter (0.88). The Dearborn test is a close second (0.87). As these two tests cover somewhat different abilities (M. placing a premium on language ability, and D. on non-language ability) the writer suggests this combination as being the best, if two tests can be given to the entire 12 grades. The writer has been able to get more from these two series (Mentimeter and Dearborn) when he is desirous of making individual analysis than with any other combination of two mass tests.

However, if only one mass test can be given, the varied character of the Otis test makes it more valuable in analysis than either the Dearborn or the Mentimeter *alone*.

The Dearborn proved difficult to give and needs shortening and simplifying, but when this is done the author feels that this will prove to be one of our very best tests. All of the difficulties could be easily rectified. At present it is not easy for the average teacher to give, and errors in scoring are much more frequent than with other mass tests. Even as it stands, it is certainly a superior test with certain foreign children who have not yet mastered the English idiom—and this failure of mastery (with a foreign language in the home) is a bigger problem than is usually realized by those giving mass tests.

21. The correlations between the various mass tests are higher and more uniform than between the mass tests and the Binet Series. The following correlations are conspicuously high:

N. A. I with N. B. I, grades 3-8.....	0.94
N. A. I with M, grades 3-8.....	0.93
M. with O. P. A, grades 1-3.....	0.92
H. D. II with N. B. I, 3-8.....	0.92

CAUSES FOR SIGNIFICANT VARIATION IN CORRELATIONS

The following are probably the chief causes for variation from test to test, and from grade to grade, in the correlations presented here:

1. Probably the greatest single factor is the difference in weight that different mass tests attach to different abilities. If a few rough captions are made, such as linguistic ability, arithmetical ability, etc., and the percentage of value attached to each caption in the various tests listed, it will be found that a marked difference exists in the relative value attached to any caption as we go from one mass test to another. It would often seem that when the maker of the test had 10 arithmetic problems *that* ability got 10 points in score; if he had 5 puzzles, puzzles scored 5; and if he happened to have on hand 20 completion sentences, completion of sentences got 20 points to the score. Be that as it may, it is certain that *chance*, rather than any *knowledge* of the relative merits of different elements in the intelligence-complex or compound, determines the *proportion* of any particular kind of psychological or pedagogical test. The *difference in proportion* is on the whole more noticeable and probably more significant as the cause of variation in score

from test to test than differences in the kind of test used by different mass-test compilers referring, of course, to the omnibus type of psychological test.

It will also be found that not only does a difference exist in the weighting of the test as a whole, but taking a certain region of the test likely to be answered, say by a Grade VI pupil, one mass test will differ radically from another both in the weight attached to different psychological captions and in the captions themselves.

2. As indicated in the line above, the mass tests differ from one another not only in the weighting of various captions but also in the actual captions included in the omnibus, or in the region of a particular grade.

3. H. Sigma correlates relatively poorly with the Binet, probably because it is essentially a reading test, and also because it demands a certain degree of reading ability. In cases where the child could read the line from the test but had attention riveted on the mechanics of reading, no action followed. Another cause of failure to respond to the test seemed to be an aversion to making a mark on the printed page.

4. It is conceivable that certain local grade conditions can play an important part; methods of teaching, predominance of certain foreign elements of a particular race, stress on certain school subjects, etc.

5. The Binet Test is largely independent of the element of time; mass tests must of necessity rest on a time basis. We do not know to what extent different subjects are benefited in one case and injured in the other, or *vice versa*.

6. Finally, marked change in the rank-order of an individual from one mass test to another or from mass test to the Binet may rest upon various individual differences. An analysis of such cases with a close study of the causes operating in an individual case, is a much needed task but beyond the scope of this preliminary report.