

The Psychological Clinic

Copyright, 1915, by Lightner Witmer, Editor.

VOL. IX, No. 7

DECEMBER 15, 1915

REPRODUCTION OF SHORT PROSE PASSAGES: A STUDY OF TWO BINET TESTS.

BY A. TRAVIS,
Philadelphia, Pa.

GENERAL CONDITIONS AND DATA.

These tests were given in the Laboratory of Psychology of the University of Pennsylvania, as a regular class-room exercise in connection with Psychology 2c (undergraduate), Psychology 62 (graduate), and Psychology 30 (the Psychological Clinic). On the afternoon of April 15, 1915, the tests were given to 77 undergraduates; to 16 graduate students on the afternoon of April 16th; to 17 students in the College Courses for Teachers and 4 more graduate students on the morning of April 17th; and finally to 12 undergraduate and 2 graduate students in the Summer School on the morning of July 28th. The total number of students tested was 128, 69 men and 59 women. One man and one woman were colored. The distribution by college classes will be found in table I.

TABLE I
SUMMARY BY CLASSES, MEN AND WOMEN

	Men	Women
Arts, Senior	5	..
Junior.....	11	..
Sophomore.....	22	..
Freshman.....	7	..
Biology, Sophomore.....	..	2
Education, Senior.....	..	1
Sophomore.....	..	17
Freshman.....	3	9
Graduate.....	12	10
College Courses for Teachers.....	6	11
Summer School.....	3	9
Totals.....	69	59

PROCEDURE.

The class having come to a state of attention, the instructor said: "I want you to have a separate sheet of paper, please, which can be collected. Put your name at the top of the sheet, and the date. I am going to dictate a paragraph, or read a paragraph. I want you to pay attention to it, and then I want you to give the paragraph back again as well as you can. This is one of the Binet tests."

He then read selection B, from Dr. Town's translation of the Binet tests (1). This contains 80 words, and the reading occupied less than one minute. The students wrote down what they could recall of it, taking an average time of two and a half to three minutes, a few requiring as long as four minutes.

After allowing a few minutes of relaxation and rest, the instructor said: "On another sheet of paper will you make a similar record of this selection, which I shall read you. I shall not begin reading until you have all finished such preliminaries as writing your name. Everybody ready?"

Then he read selection A, which is taken from the 1915 Stanford revision of the Binet tests (10), with very slight changes. This selection contains 86 words, and the time of reading and subsequent writing was practically the same as in the preceding test. The papers were collected and filed, with the exception of selection B for 14 graduate students. These were given back to their writers to work over, and were not returned to the instructor. The paper for one student in the college courses for teachers appears to have been lost. Results are reported for 113 students in selection B. As will be seen later, the omission of these 15 records is probably not important.

MARKING.

Table II shows the results of both tests, by classes.

Selection B was marked in two ways. It was first analysed into seven essential ideas as follows, and the 113 papers were rated upon the number of points in seven which were reproduced by the student.

1. Many opinions have been given on the value of life.
2. Some call it good,
3. others call it bad.
4. It would be more just to say that it is mediocre,
5. for on the one hand our happiness is never so great as we would have it,
6. and on the other hand our misfortunes are never so great as others would have them.
7. It is this mediocrity of life which makes it just, or rather which prevents it from being radically unjust.

The passage was then subjected to a close logical (or rather grammatical) analysis after the manner set forth by Jevons (6) in his "Lessons in Logic," pages 94-97. This is very similar to the plan followed by Binet and Henri (2) in scoring the papers in their experiment upon children's memory span for phrases. As they remark, the division into phrases is bound to be somewhat arbitrary and any two experimenters will not agree about it, but the important thing is first to make as consistent a division as possible, and then to apply it impartially. By this method the paragraph yielded sixteen possible units, as follows:

Many opinions have been given	(1)	
on the value of life.	(2)	
Some call it good,	(3)	
others call it bad.	(4)	
It would be more just	(5)	
to say that it is mediocre,	(6)	
for { on the one hand	(7)	
our happiness is never so great		(8)
as we would have it,		(9)
and { on the other hand		(10)
our misfortunes are never so great		(11)
as others would have them.		(12)
It is this mediocrity of life		(13)
which makes it just,		(14)
or rather		(15)
which prevents it from being radically unjust.		(16)

The papers were rated again upon this basis, and the rating was reduced to a percentage. One man and six women made a perfect score, 100 per cent; and one student, the colored man, had only one point in sixteen, or 6.2 per cent. Three other men and six women made scores under 50 per cent. All the rest of the students, 96, made scores between 50 and 93.7 per cent. This in itself is sufficient evidence that the test was too easy to be an adequate means of ranking so large a group. It is significant, on the other hand, that out of 113 university students, ranging in age from, say 17 to 45 years, 10 should have failed to make a score of 50 per cent in a test which is set down among the Binet tests for 15 years and over. Kuhlmann (7) says the subjects should give back two-thirds of the ideas, but he does not tell us how to score the test or how to calculate the two-thirds. On this basis, 84 students made approximately 66 per cent or over, and 29 failed to pass. Must we regard these persons as in some respect subnormal? Or is it just one more bit of evidence for the unreliability of the tests?

Selection A yielded results far more interesting. There were 128 papers, and six (in fact eight) ways of marking were chosen for comparison and study:

(1) The papers were read quickly and sorted into five heaps,—distinguished, good, passed, not passed, and failed. This is a very satisfactory method for ranking college quiz papers when only two or three topics are assigned for discussion. There were 13 D's, 41 G's, 48 P's, 22 N's, and 4 F's. This distribution contains more D's and G's than a large class taking an examination in psychology, but there are fewer P's, more N's and F's. The percentage value of these marks will be discussed later.

Mr. Finkelstein, who has written a monograph (5) upon the marking system, based upon an analytic study of many thousands of marks given at Cornell University, believes that a five-division system is the best. His favored system differs from the one at Pennsylvania in that he would place four of these marks above the passing grade and one below. At Pennsylvania we use two marks below passing, and three above. By combining Mr. Finkelstein's two highest divisions, however, we can bring his scale into approximate correspondence with ours, 60 being the passing mark in both scales. Translating his divisions now into D, G, P, and N, and combining our two lowest divisions as well as his two highest, we can make the following comparison. He believes that the theoretically ideal distribution of grades in the hundred would be, approximately, 25 D's, 45 G's, 19 P's, 12 N's. The distribution in the hundred for this test was as follows, 10 D's, 32 G's, 38 P's, 20 N's (and F's). This substantiates the idea expressed in the conclusions of the present study, that tests of this sort are not a fair index of students' ability in college work. The better students at any rate, and probably all grades of students, tend to make low scores on tests when rated solely by points.

(2) The five piles representing the five grades were now assembled into one, and the papers were arranged in rank order from the best D down to the poorest F (see third column of table II).

(3) Selection A was analysed into its essential ideas, of which it contained fourteen as follows:

1. Tests such as we are now making
2. are of value
3. both for the advancement of science
4. and for the information of the person who is tested.

5. It is important for science
6. to learn how people differ
7. and on what factors these differences depend.
8. If we can separate the influence of heredity
9. from the influence of environment
10. we may be able to apply our knowledge
11. so as to guide human development.
12. We may thus in some cases correct defects
13. and develop abilities
14. which we might otherwise neglect.

The papers were scored on the basis of the possible fourteen points, and it is interesting to note that while seven students (one man and six women) made a perfect score on the slightly shorter selection, no one made a perfect score on selection A, and only one student, a woman in the Summer School, made as high as 12 points in the 14.

(4) The papers were ranked again, this time strictly in order of points, and the second rank order is recorded in column 5 of table II and column 6 of table III. Table III gives the results according to the first rank order, and a comparison of columns 1 and 6 of table III will show the changes in rank order caused by giving chief importance to points instead of to other qualities which were considered in assigning the first rank order,—such as judgment in organizing and condensing material, terseness and vigor of expression, and finally what is hard to describe but easy to recognize, an effect of skill and mastery in the execution of the task.

(5) The papers being still in the second rank order, were given percentages ranging from the highest, twelve-fourteenths or 86 per cent, down to number 125, three-fourteenths or 21 per cent, with 20.7, 20.3, and 20.0 added to accommodate the last three papers, numbers 126, 127, and 128, which were even poorer. It was not assumed that the decimal fractions of one per cent had any significance whatever except as aids in maintaining an unbroken series from highest to lowest and avoiding duplications of rank.

This sort of percentage ranking used to be more common in the public schools than it is at present. If taken too seriously it is open to the criticism formulated by Mr. Finkelstein in his monograph on marking. He says, "Theoretically this scale implies that distinctions of a fineness of one-hundredth may be made, and in practice such distinctions are constantly attempted. But what is the difference, if any, between a mark of 75 and one of 76? What, for that matter, does 75 mean? Has the student accomplished 75

per cent of some ideal accomplishment? . . . The very fact that its divisions [the hundred-point scale's] are so minute is doubly insidious; it promises precision, but it cannot afford it." In the case of selection A, a score of 86 per cent means that the student has actually reproduced 86 per cent of the ideas given out. A score of 81 per cent, however, represents 11/14 of the ideas, which strictly should be 79 per cent, but some papers with 11 points were better than others, and some were worse, so that in ranking the papers by percentages the seven papers with 11 points each were distributed over the range from 81 per cent down to 76 per cent inclusive.

(6 and 7) The selection was analyzed a second time, after the method of Jevons, and now yielded twenty-five points. The papers were marked on this basis, and the score reduced to a percentage by multiplying by four.

(1)	Tests	such as we are now making	(2)
	(3)	{ both for the advancement of science	(4)
	are of value	{ and for the information of the person	(5)
		(6) who is tested.	
(7)	It is important	for science	(8)
		{ how people differ	(10)
	(9) to learn	{ and on what factors	(11)
		these differences depend.	(12)
(13)	(14)	{ the influence of heredity	(15)
	If we can separate	{ from the influence of environment,	(16)
	(17)		
	we may be able	to apply our knowledge	(18)
		so as to guide human development.	(19)
	We may	{ thus	(20)
		{ in some cases	(21)
(22)			
	correct defects	}	
(23)		{ which we might otherwise	(24)
	and develop abilities	} neglect.	(25)

(8) The percentages now ranged from the highest, 22 points or 88 per cent, down to four points or 16 per cent as the lowest. The results will be found in columns 7 and 8 of table II, and in columns 8 and 9 of table III. The distribution by points will be found summarized in table IV.

REPRODUCTION OF SHORT PROSE PASSAGES. 195

TABLE II
RESULTS OF TESTS A AND B. SUBJECTS GROUPED BY COLLEGE CLASSES

CLASS	TEST A							TEST B			Col- lege Rank- ing	Term Mark Psych. 2
	Grade	Rank order	Points in 14	2d rank order	Per- centage for 2d rank order	Points in 25	Per cent	Points in 7	Points in 16	Per cent		
A '15	G	27	8	41	57.0	12	48.0	7	11	68.7	94.5	D
	G	15	9	18	64.7	14	56.0	7	12	75.0	77.0	P
	P	101	6	99	38.3	7	28.0	7	11	68.7	77.7	P
	G	33	8	49	54.3	10	40.0	6	10	62.5	93.0	G
	N	122	4	122	22.5	10	40.0	6	8	50.0	88.0	P
A '16	D	12	10	12	70.0	16	64.0	7	12	75.0	P
	N	113	4	114	26.5	7	28.0	6	11	68.7	N
	G	42	8	51	53.7	12	48.0	6	11	68.7	62.9	N
	G	23	9	25	62.3	16	64.0	7	14	87.5	84.8	D
	G	46	9	32	60.0	12	48.0	5	8	50.0	68.8	N
	G	48	9	34	59.3	14	56.0	7	14	87.5	P
	P	88	6	91	41.3	8	32.0	7	14	87.5	79.2	G
	G	41	8	50	54.0	12	38.0	4	8	50.0	77.4	P
	P	75	7	76	46.0	11	44.0	7	14	87.5	G
	G	29	9	27	61.7	11	44.0	7	10	62.5	63.3	N
		P	93	6	94	40.0	9	36.0	5	8	50.0	68.8
A '17	D	8	11	7	76.0	16	64.0	7	11	68.7	P
	P	71	7	73	47.0	11	44.0	4	7	43.7	P
	G	52	8	55	52.3	11	44.0	7	12	75.0	P
	F	124	3	125	21.0	4	16.0	7	9	56.2	P
	P	67	7	69	48.3	10	40.0	7	16	100.0	G
	G	53	8	56	52.0	11	44.0	7	10	62.5	P
	P	90	6	92	41.0	10	40.0	6	8	50.0	G
	N	118	4	118	24.5	5	20.0	7	14	87.5	P
	G	17	10	14	68.0	13	52.0	7	14	87.5	P
	P	73	8	62	50.4	10	40.0	7	13	81.2	P
	P	55	8	58	51.3	10	40.0	7	14	87.5	G
	G	16	10	13	69.0	16	64.0	7	12	75.0	P
	D	5	11	5	78.0	18	72.0	7	13	81.2	P
	P	72	7	74	46.7	10	40.0	5	8	50.0	N.S
	P	64	9	37	58.3	14	56.0	7	15	93.7	G
	N	123	4	123	22.0	5	20.0	7	11	68.7	P
	N	109	4	112	27.5	9	36.0	6	10	62.5	G
	P	65	9	37	57.7	11	44.0	7	13	81.2	P
	P	98	6	96	39.3	11	44.0	7	15	93.7	P
	P	70	7	72	47.3	11	44.0	7	15	93.7	N
N	120	4	120	23.5	7	28.0	6	10	62.5	F	
	F	125	4	124	21.5	7	28.0	4	9	56.2	N
A '18	N	112	4	113	27.0	7	28.0	5	7	43.7	P
	G	25	10	15	67.0	14	56.0	5	9	56.2	P
	G	20	9	20	64.0	15	60.0	6	8	50.0	G
	N	103	5	103	36.0	9	36.0	7	14	87.5	G
	N	114	5	110	29.0	8	32.0	6	11	68.7	P
	N	111	5	109	30.0	9	36.0	6	10	62.5	G
		N	110	5	108	31.0	7	28.0	7	10	62.5
B '17	P	68	7	70	48.0	11	44.0	7	14	87.5	P
	D	9	10	9	73.0	16	64.0	7	13	81.2	G

TABLE II (Continued)

CLASS	TEST A							TEST B			Term Mark Psych. 2	
	Grade	Rank order	Points in 14	2d rank order	Per- centage for 2d rank order	Points in 25	Per cent	Points in 7	Points in 16	Per cent		
E '15	G	49	9	35	59.0	13	52.0	7	12	75.0	D	
E '17	G	34	8	46	55.3	12	48.0	7	16	100.0	P	
	P	66	8	61	50.6	12	48.0	5	11	68.7	N	
	P	92	6	94	40.3	9	36.0	6	9	56.2	P	
	G	35	8	45	55.7	12	48.0	7	16	100.0	G	
	P	57	8	59	51.0	12	48.0	7	14	87.5	G	
	F	127	3	127	20.3	6	24.0	6	9	56.2	P	
	D	3	11	3	80.0	18	72.0	7	15	93.7	G	
	G	24	8	40	57.3	12	48.0	7	14	87.5	G	
	P	91	6	93	40.7	11	44.0	7	14	87.5	P	
	P	97	6	101	37.7	10	40.0	7	14	87.5	P	
	G	31	9	29	61.0	12	48.0	6	11	68.7	G	
	P	100	6	98	38.7	11	44.0	7	16	100.0	P	
	P	76	6	85	43.3	10	40.0	7	16	100.0	G	
	F	128	3	128	20.0	5	20.0	7	14	87.5	P	
	D	7	11	8	75.0	15	60.0	7	14	87.5	G	
	G	19	9	22	63.3	14	56.0	5	13	81.2	G	
	P	69	7	71	47.6	10	40.0	6	12	75.0	G	
E '18	F	126	3	126	20.7	5	20.0	5	8	50.0	P	
	N	115	4	115	26.0	8	32.0	7	12	75.0	G	
	N	105	5	105	34.0	9	36.0	7	13	81.2	P	
	P	74	7	75	46.3	10	40.0	6	10	62.5	P	
	G	36	8	44	56.0	14	56.0	7	14	87.5	P	
	G	22	9	24	62.7	12	48.0	7	14	87.5	G	
	N	108	5	107	32.0	6	24.0	7	15	93.6	P	
	D	2	11	2	81.0	20	80.0	7	14	87.5	P	
	G	38	8	48	54.7	11	44.0	6	11	68.7	P	
	D	4	11	4	79.0	18	72.0	7	15	93.7	G	
	P	59	7	65	49.7	11	44.0	5	8	50.0	G	
	D	6	11	6	77.0	17	68.0	7	16	100.0	P	
	Grad.	P	56	7	64	50.0	11	44.0	
		G	45	9	31	60.3	13	52.0	
P		94	6	95	39.7	7	28.0		
G		39	8	43	56.3	13	52.0		
N		119	4	119	24.0	5	20.0	P	
N		117	4	117	25.0	7	28.0		
G		51	8	54	52.7	11	44.0		
P		81	6	87	42.7	7	28.0		
P		60	7	66	49.3	10	40.0		
P		63	8	60	51.8	12	48.0		
G		54	8	57	51.7	12	48.0		
P		58	7	36	58.7	11	44.0	7	11	68.7	P	
N		107	5	111	28.0	7	28.0	P	
D		11	10	11	71.0	16	64.0		
D		13	9	17	65.0	15	60.0		
G		14	9	19	64.3	11	44.0	7	11	68.7	P	
G		43	8	52	53.3	10	40.0	7	12	75.0		

REPRODUCTION OF SHORT PROSE PASSAGES. 197

TABLE II (Continued)

CLASS	TEST A							TEST B		
	Grade	Rank order	Points in 14	2d rank order	Percentage for 2d rank order	Points in 25	Per cent	Points in 7	Points in 16	Per cent
Grad. (cont.)	G	21	9	23	63.0	11	44.0	6	10	62.5
	G	47	9	33	59.7	14	56.0	6	12	75.0
	P	86	6	89	42.0	9	36.0	6	11	68.7
	G	30	8	42	56.7	13	52.0	4	6	37.5
	N	116	4	116	25.5	5	20.0	4	7	43.7
C. C. T.	P	96	7	84	43.7	11	44.0	6	8	50.0
	P	84	7	81	44.3	12	48.0	6	11	68.7
	P	102	6	100	38.0	10	40.0	7	10	62.5
	G	32	9	38	58.0	10	44.0	3	5	31.2
	P	95	7	82	44.0	11	44.0	6	11	68.7
	G	28	10	16	66.0	13	52.0	5	8	50.0
	P	85	6	88	42.3	9	36.0	5	6	37.5
	P	87	6	90	41.7	12	48.0	4	5	31.2
	P	82	7	79	45.0	10	40.0	5	8	50.0
	D	10	10	10	72.0	16	64.0	7	12	75.0
	P	83	7	80	44.7	10	40.0	7	13	81.2
	P	62	8	63	50.2	14	56.0	6	9	56.2
	N	104	5	104	35.0	7	28.0
	N	121	4	121	23.0	7	28.0	1	1	6.2
	P	99	6	97	39.0	9	36.0	7	12	75.0
P	89	5	102	37.0	8	32.0	6	11	68.7	
N	106	5	106	33.0	6	24.0	4	9	56.2	
S. S.	P	80	7	78	45.3	10	40.0	3	4	25.0
	P	61	7	67	49.0	13	52.0	7	13	81.2
	G	37	8	47	55.0	13	52.0	5	11	68.7
	P	78	7	77	45.7	11	44.0	5	7	43.7
	G	26	9	26	62.0	13	52.0	7	14	87.5
	G	40	9	28	61.3	14	56.0	7	16	100.0
	P	79	8	68	48.7	11	44.0	7	10	81.2
	G	44	9	30	60.7	13	52.0	6	12	75.0
	G	50	8	53	53.0	10	40.0	7	11	68.7
	P	77	6	86	43.0	9	36.0	7	15	81.2
	D	1	12	1	86.0	22	88.0	6	15	93.6
	G	18	9	21	63.7	15	60.0	6	12	75.0

TABLE III—SELECTION A. SUBJECTS ARRANGED IN RANK ORDER

Rank	Sex	Class	Grade	Points in 14	Second rank order	Per cent	Points in 25	Per cent	Term mark Psych. 2
1	F	S. S.	D	12	1	86.0	22	88.0	..
2	F	E '18	D	11	2	81.0	20	80.0	P
3	F	E '17	D	11	3	80.0	18	72.0	G
4	F	E '18	D	11	4	79.0	18	72.0	G
5	M	A '17	D	11	5	78.0	18	72.0	P
6	F	E '18	D	11	6	77.0	17	63.0	P
7	F	E '17	D	11	8	75.0	15	60.0	G
8	M	A '17	D	11	7	76.0	16	64.0	P
9	F	B '17	D	10	9	73.0	16	64.0	G
10	F	C. C. T.	D	10	10	72.0	16	64.0	..
11	F	Grad.	D	10	11	71.0	16	64.0	..
12	M	A '16	D	10	12	70.0	16	64.0	P
13	F	Grad.	D	9	17	65.0	15	60.0	..
14	F	Grad.	G	9	19	64.3	11	44.0	..
15	M	A '15	G	9	18	64.7	14	56.0	P
16	M	A '17	G	10	13	69.0	15	64.0	P
17	M	A '17	G	10	14	68.0	13	52.0	P
18	F	S. S.	G	9	21	63.7	15	60.0	..
19	F	E '17	G	9	22	63.3	14	56.0	G
20	M	A '18	G	9	20	64.0	15	60.0	G
21	M	Grad.	G	9	23	63.0	11	44.0	..
22	F	E '18	G	9	24	62.7	12	48.0	G
23	M	A '16	G	9	25	62.3	16	64.0	D
24	F	E '17	G	8	40	57.3	12	48.0	G
25	M	A '18	G	10	15	67.0	14	56.0	P
26	F	S. S.	G	9	26	62.0	13	52.0	..
27	M	A '15	G	8	41	57.0	12	48.0	D
28	F	C. C. T.	G	10	16	66.0	13	52.0	..
29	M	A '16	G	9	27	61.7	11	44.0	N
30	F	Grad.	G	8	42	56.7	13	52.0	..
31	F	E '17	G	9	29	61.0	12	48.0	G
32	F	C. C. T.	G	9	38	58.0	10	44.0	..
33	M	A '15	G	8	49	54.3	10	40.0	G
34	F	E '17	G	8	46	55.3	12	48.0	P
35	F	E '17	G	8	45	55.7	12	48.0	G
36	F	E '18	G	8	44	56.0	14	56.0	P
37	M	S. S.	G	8	47	55.0	13	52.0	..
38	M	E '18	G	8	48	54.7	11	44.0	P
39	M	Grad.	G	8	43	56.3	13	52.0	..
40	F	S. S.	G	9	28	61.3	14	56.0	..
41	M	A '16	G	8	50	54.0	12	48.0	P
42	M	A '16	G	8	51	53.7	12	48.0	N
43	M	Grad.	G	8	52	53.3	10	40.0	..
44	F	S. S.	G	9	30	60.7	13	52.0	..
45	M	Grad.	G	9	31	60.3	13	52.0	..
46	M	A '16	G	9	32	60.0	12	48.0	N
47	M	Grad.	G	9	33	59.7	14	56.0	..
48	M	A '16	G	9	34	59.3	14	56.0	P
49	F	E '15	G	9	35	59.0	13	52.0	D
50	F	S. S.	G	8	53	53.0	10	40.0	..
51	M	Grad.	G	8	54	52.7	11	44.0	..
52	M	A '17	G	8	55	52.3	11	44.0	P
53	M	A '17	G	8	56	52.0	11	44.0	P
54	F	Grad.	G	8	57	51.7	12	48.0	..
55	M	A '17	P	8	58	51.3	10	40.0	G
56	M	Grad.	P	7	64	50.0	11	44.0	..
57	F	E '17	P	8	59	51.0	12	48.0	G
58	F	Grad.	P	7	36	58.7	11	44.0	..
59	F	E '18	P	7	65	49.7	11	44.0	G
60	M	Grad.	P	7	66	49.3	10	40.0	..
61	M	S. S.	P	7	67	49.0	13	52.0	..
62	M	C. C. T.	P	8	63	50.2	14	56.0	..
63	F	Grad.	P	8	60	51.8	12	48.0	..
64	M	A '17	P	9	37	58.3	14	56.0	G

TABLE III (Continued)

Rank	Sex	Class	Grade	Points in 14	Second rank order	Per cent	Points in 25	Per cent	Term mark Psych. 2
65	M	A '17	P	9	37	57.7	11	44.0	P
66	F	E '17	P	8	61	50.6	12	48.0	N
67	M	A '17	P	7	69	48.3	10	40.0	G
68	F	B '17	P	7	70	48.0	11	44.0	P
69	F	E '17	P	7	71	47.6	10	40.0	G
70	M	A '17	P	7	72	47.3	11	44.0	N
71	M	A '17	P	7	73	47.0	11	44.0	P
72	M	A '17	P	7	74	46.7	10	40.0	N. S.
73	M	A '17	P	8	62	50.4	10	40.0	P
74	F	E '18	P	7	75	46.3	10	40.0	P
75	M	A '16	P	7	76	46.0	11	44.0	G
76	F	E '17	P	6	85	43.3	10	40.0	G
77	F	S. S.	P	6	86	43.0	9	36.0	..
78	F	S. S.	P	7	77	45.7	11	44.0	..
79	F	S. S.	P	8	68	48.7	11	44.0	..
80	M	S. S.	P	7	78	45.3	10	40.0	..
81	M	Grad.	P	6	87	42.7	7	28.0	..
82	F	C. C. T.	P	7	79	45.0	10	40.0	..
83	F	C. C. T.	P	7	80	44.7	10	40.0	..
84	F	C. C. T.	P	7	81	44.3	12	48.0	..
85	F	C. C. T.	P	6	88	42.3	9	36.0	..
86	F	Grad.	P	6	89	42.0	9	36.0	..
87	F	C. C. T.	P	6	90	41.7	12	48.0	..
88	M	A '16	P	6	91	41.3	8	32.0	G
89	M	C. C. T.	P	5	102	37.0	8	32.0	..
90	M	A '17	P	6	92	41.0	10	40.0	G
91	F	E '17	P	6	93	40.7	11	44.0	P
92	F	E '17	P	6	94	40.3	9	36.0	P
93	M	A '16	P	6	94	40.0	9	36.0	P
94	M	Grad.	P	6	95	39.7	7	28.0	..
95	F	C. C. T.	P	7	82	44.0	11	44.0	..
96	F	C. C. T.	P	7	84	43.7	11	44.0	..
97	F	E '17	P	6	101	37.7	10	40.0	P
98	M	A '17	P	6	96	39.0	11	44.0	P
99	M	C. C. T.	P	6	97	39.0	9	36.0	..
100	F	E '17	P	6	98	38.7	11	44.0	P
101	M	A '15	P	6	99	38.3	7	28.0	P
102	F	C. C. T.	P	6	100	38.0	10	40.0	..
103	M	A '18	N	5	103	36.0	9	36.0	G
104	M	C. C. T.	N	5	104	35.0	7	28.0	..
105	F	E '18	N	5	105	34.0	9	36.0	P
106	M	C. C. T.	N	5	106	33.0	6	24.0	..
107	F	Grad.	N	5	111	28.0	7	28.0	..
108	F	E '18	N	5	107	32.0	6	24.0	P
109	M	A '17	N	4	112	27.5	9	36.0	G
110	M	A '18	N	5	108	31.0	7	28.0	G
111	M	A '18	N	5	109	30.0	9	36.0	G
112	M	A '18	N	4	113	27.0	7	28.0	P
113	M	A '16	N	4	114	26.5	7	28.0	N
114	M	A '18	N	5	110	29.0	8	32.0	P
115	M	E '18	N	4	115	26.0	8	32.0	G
116	F	Grad.	N	4	116	25.5	5	20.0	..
117	M	Grad.	N	4	117	25.0	7	28.0	..
118	M	A '17	N	4	118	24.5	5	20.0	P
119	M	Grad.	N	4	119	24.0	5	20.0	..
120	M	A '17	N	4	120	23.5	7	28.0	F
121	M	C. C. T.	N	4	121	23.0	7	28.0	..
122	M	A '15	N	4	122	22.5	10	40.0	P
123	M	A '17	N	4	123	22.0	5	20.0	P
124	M	A '17	F	3	125	21.0	4	16.0	P
125	M	A '17	F	4	124	21.5	7	28.0	N
126	M	E '18	F	3	126	20.7	5	20.0	P
127	F	E '17	F	3	127	20.3	6	24.0	P
128	F	E '17	F	3	128	20.0	5	20.0	P

TABLE IV
DISTRIBUTION OF 128 STUDENTS BY POINTS REPRODUCED
SELECTION A—14 POINT ANALYSIS

	Men	Women	Total	
12 points in 14.....	..	1	1	
11 "	2	5	7	
10 "	4	4	8	
9 "	11	11	22	
				38 scored 66 % or over
8 "	14	11	25	
7 "	9	11	20	
6 "	8	10	18	
5 "	7	3	10	
4 "	12	1	13	
3 "	2	2	4	
				90 scored less than 66 %
				Total 128

SELECTION A—25 POINT ANALYSIS

22 points in 25.....	..	1	1	
20 "	1	1	
18 "	1	2	3	
17 "	1	1	
16 "	4	3	7	
15 "	1	3	4	
14 "	6	3	9	
13 "	5	5	10	
				36 scored over 50 %
12 "	4	11	15	
11 "	12	10	22	
10 "	10	9	19	
9 "	5	5	10	
8 "	4	..	4	
7 "	11	1	12	
6 "	1	2	3	
5 "	4	2	6	
4 "	1	..	1	
				92 scored less than 50 %
				Total 128

SELECTION B—7 POINT ANALYSIS

7 points in 7.....	31	31	62	
				scored 100 %
6 "	15	12	27	
5 "	6	8	14	
4 "	4	3	7	
				48 scored more than 50 % and less than 100 %
3 "	1	1	2	
1 "	1	..	1	3 scored less than 50 %
				Total 113

TABLE IV (Continued)
SELECTION B—16 POINT ANALYSIS

16 points in 16.....	1	6	7	
15 "	3	5	8	
14 "	8	11	19	
13 "	4	4	8	
12 "	8	5	13	
11 "	10	8	18	
10 "	8	3	11	
				84 scored 66 % or over
9 "	5	2	7	
8 "	8	4	12	
7 "	2	2	4	
6 "	2	2	
5 "	2	2	
4 "	1	..	1	
1 "	1	..	1	
				29 scored less than 66 %
				Total 113

DISCUSSION.

This test has been called a test of "memory span for ideas." It is purposely made so long that it exceeds the average adult memory span for connected words. One student reproduced the shorter passage word for word, but the fact that she failed to reproduce the slightly longer paragraph in the same way, makes it unnecessary to inquire whether she had ever heard it before, whether she was accustomed to repeating a paragraph verbatim upon one hearing, and for how long she could retain a passage learned in this way. The Admirable Crichton is said to have been able to repeat a long oration verbatim after hearing it only once. The mathematical prodigy Inaudi has a memory-span of forty-two digits. Mozart and other musicians have gone home and played from memory a long melody heard once at a concert.

No prodigies are to be found in the group of university students tested, but it is conceivable that most of them would have done much better had they been tested singly, under favorable clinical conditions, and if they had been asked to repeat *orally* what they had just heard. Required to give back the paragraphs in writing, their habit of condensing in note-taking seems to have operated as a restraint upon ample expression. Another factor worth considering is the modern adult's laziness when it comes to using longhand as a medium. These factors were distinctly observable in the results of the tests, but they operated in varying degrees not susceptible to close estimation. In a similar way it is evident that this test should reveal something about the range and concentration of attention, but here again we can do little more than affirm what

we knew beforehand, that the better papers show a greater power of controlled attention to auditory stimuli of this sort, and greater persistence in writing down as much as possible.

Perhaps it would be more satisfactory, that is, approximately nearer to a complete understanding of what might be done with the test, to say that it tests an adult's imageability for phrases arranged in a paragraph and conveying moderately abstract ideas.

For purposes of comparison, the college records were obtained from the Dean's office for all undergraduates who had been in college three years or more. Three seniors had taken their course in three years; the remaining two had records for the full four years. Of the juniors in Arts, seven had been in college for three years.

Marks are recorded in the Dean's office by the five grades, D, G, P, N, and F. These grades have slightly different values in different departments. For example the passing mark in English is 70, in mathematics 60. Nevertheless it is admitted to be no more difficult to make a passing mark in English than in mathematics. In reducing these grades to percentages and striking an average, the method was adopted which was used last year and in previous years by Phi Beta Kappa,—D equivalent to 100 per cent, G to 80 per cent, P to 60 per cent, N to 40 per cent, and F to 20 per cent. This method has been subjected to criticism by members of the Pennsylvania Phi Beta Kappa Committee, and next year a revised scale may be used, in which D will be counted as 93 per cent, G as 85 per cent, and P as 73 per cent, these values representing a fairer median for the marks.

A comparison of columns 6 and 12 of table II shows that there is no correlation between college standing and the results of these tests. The most brilliant senior, with a general average of 94.5 per cent, ranked number 27 on our list with a score of only 57 per cent; and the next highest senior with an average of 93, ranked 33 on our list with a score of 54.3 per cent. These are the first rankings; if the ranking by points is compared, the discrepancies are even greater, as these men ranked number 41 and 49 respectively by points. Two juniors with exactly the same college standing, 68.8 per cent, ranked number 32 and 94 respectively on the second ranking.

This comparison enables us to discount the fact that the women students did very much better in these tests than the men. In the first ten per cent for Selection A (13 students) there were 10 women and 3 men; in the last 10 per cent there were 10 men and 3 women. The general distribution of men and women will be seen in tables III and IV.

There was not even the slightest correlation between the rankings of these tests and the term marks given in Psychology 2, in which the tests formed a class exercise. Of the undergraduates tested, 81 had marks in Psychology 2, as follows: 3 D's, 26 G's, 43 P's, 7 N's, 1 N.S. (no standing) and 1 F. The students receiving D ranked 23, 27, and 49 on test A. Those receiving G ranked as follows: 3, 4, 7, 9, 19, 20, 22, 24, 31, 33, 35, 55, 57, 59, 64, 67, 69, 75, 76, 88, 90, 103, 109, 110, 111, 115. Those receiving N ranked: 29, 42, 46, 66, 70, 113, 125.

There was likewise no correlation between the results of A and B. The prose passages used for the tests were very similar, in that both were trite, and badly phrased, and contained no ideas which could be expected to arouse keen interest in any member of a college class. They suggested no concrete sensory images; they stirred no emotion. To seize them in passing was a task comparable to the catching of the well-greased pig at the country fair. Furthermore, they were so nearly equal in length that the longer contained only six words more than the shorter. Yet in spite of all these similarities, the results do not correlate. The seven perfect scores on selection B were made respectively by numbers 6, 34, 35, 40, 67, 76, and 100 on the ranking list for selection A; and the one man who did worst in selection B, making a score of only 6.2 per cent, stood eighth from the last on selection A.

For testing students of university grade the passages should be chosen from the work of acknowledged masters of English prose. At least two passages and preferably three should be used, varying from abstract to concrete in ideas, and the results on the three tests averaged, if possible, for each individual tested. An experiment of this sort upon at least a hundred students would put us in a position to judge more fairly of the value of such tests, and to standardize them if it seemed worth while to retain them as tests of students' ability in college work.

As substitutes for the unsatisfactory passages used in the present study I would suggest the following three. The first is from Milton's "Eikonoclastes." It contains 80 words, and voices an abstract idea which should be within the comprehension and power of expression of anyone who has passed college entrance examinations in English. It should be read slowly, for it is so well organized that its relations are not immediately obvious. The papers could hardly be scored by points; but could be rated in five grades of excellence, and arranged in rank order; and could then be assigned percentages to indicate the ranking.

I.

To counterfeit the hand of God is the boldest of all forgery. And he who without warrant but his own fantastic surmise, takes upon him perpetually to unfold the secret and unsearchable mysteries of high providence, is likely for the most part to mistake and slander them; and approaches for the most part to the madness of those reprobate thoughts that would wrest the sword of justice out of God's hand, and employ it more justly in their own conceit.

The next is from Tyndall's "Fragments of Science," chapter 8, Radiation, and contains 85 words. It makes perhaps too much demand upon concrete visual imagery, but is well adapted to the abilities of college classes, and moreover would be easy to score as a whole or by points.

II.

Exposing his thermometers to the successive colors of the solar spectrum, Sir William Herschel determined the heating power of each, and also that of the region beyond the extreme red. Then drawing a straight line to represent the length of the spectrum, he erected, at various points, perpendiculars to represent the calorific intensity existing at those points. Uniting the ends of all his perpendiculars, he obtained a curve which showed at a glance the manner in which the heat was distributed in the solar spectrum.

The last is from Gibbon's "Decline and Fall of the Roman Empire," chapter 56, and contains 124 words. It is rich in concrete imagery, and contains so many ideas that it is easy to score by points.

III

The Normans are a cunning and revengeful people; eloquence and dissimulation appear to be their hereditary qualities: they can stoop to flatter; but unless they are curbed by the restraint of law, they indulge the licentiousness of nature and passion. Their princes affect the praise of popular munificence; the people observe the medium, or rather blend the extremes, of avarice and prodigality; and, in their eager thirst of wealth and dominion, they despise whatever they possess, and hope whatever they desire. Arms and horses, the luxury of dress, the exercises of hunting and hawking, are the delight of the Normans, but on pressing occasions they can endure with incredible patience the inclemency of every climate, and the toil and abstinence of a military life.

Tests suitable for children in the elementary schools are formulated by Supt. D. C. Bliss (3) of Montclair in an article in THE PSYCHOLOGICAL CLINIC for March, 1912. Mr. Bliss uses the five-point method of scoring, but applies it in such manner that it is really a three-point system. Number one group includes the papers which are unintelligible, or "yet so poor as to show that the child has not grasped the rudiments of English composition." The highest grade, number five, he says "may not be reached by more than one child in five hundred." He has standardized the method of giving the tests, and discusses the qualities upon which scoring is to be based, showing specimen papers for each grade.

IMAGINAL TYPES.

When this investigation was first begun, it was intended to include in it a study of the imaginal types of the persons tested, as shown in their reproduction of a short story, to be known as test C. This story was taken from Colvin and Myers' (4) study of imagination, and was planned in such a way as to arouse a nearly equal number of concrete visual, auditory, and kinesthetic sensory images, in addition to the verbal imagery involved in its recall. This investigation has been excluded from the present study for two reasons which seemed cogent. In the first place, as has been remarked, selections A and B contained nothing which could be expected to arouse concrete sensory images. In the second place, Colvin believes, "In the higher grades and the university . . . visual material seems to be better learned and retained by all three types. Here doubtless the auditory and motor types rapidly and with practically no loss translate the material from one sense department to another. This can readily be done, since concrete imagery has greatly fallen off and symbolic and verbal thinking has taken its place, a type of thinking to which visual imagery particularly lends itself."

Meumann (8) does not quite agree with Colvin as to the absence of loss in the translation from one kind of imagery to another. He cites Polhmann's observations which go to show that an observer of visual type is hampered by the rapidity of an oral presentation. But in general his opinion is the same as Colvin's. He says: "We must not fail to bear in mind that the method of presentation does not by any means determine the method which will be adopted by the observer in his perception and apprehension of the material. This latter is a joint product of the mode of presentation and of the ideational type of the observer. If we present any material orally to a number of pupils, only those who belong to the pure auditory type will apprehend it in terms of auditory images; the visualizers will, so far as possible, transform it into visual images, and the observers who belong to the auditory-motor type will pronounce the words to themselves during the presentation."

The graduate students noted on the margin of their papers the imaginal type to which each believed he belonged. This was arrived at, in most cases, by "easy chair" introspection, and would need to be checked up by some controlled introspection under strict laboratory conditions. Only once did the memorandum of imaginal type seem to throw any light on the score of selection A, and that was in the case of a man who made a very low score, ranking number 119, and who remarked, "I think I am visual in imagery—little of it in this."

TRANSPOSITIONS AND OMISSIONS.

Transpositions, which are so common in the recall of series of digits used as a test of memory span, were very rarely met with in these prose passages. Three men and seven women transposed ideas, at the same time preserving the meaning.

The sentence most frequently omitted was the second,—“It is important for science to learn how people differ and on what factors these differences depend.” Thirty-five men (50 per cent) and 26 women (44 per cent) omitted it. The concluding sentence,—“We may thus in some cases correct defects and develop abilities which we might otherwise neglect,” was the next most frequently omitted, being missing from the papers of 29 men and 19 women. The first sentence of selection A was omitted by no one, although the first sentence of selection B was the one most frequently omitted in that test.

One misunderstanding was so common as to be worth mention. “For the information of the person who is tested,” seemed to confuse some of the students, who took it to mean information *about* or *concerning* the person tested. This mistake was made in various forms by five men and seven women.

As might be expected when the originals were in such feeble English, the reproductions were even weaker. Errors in syntax, however, were not frequent. One or two foreign-born students used a singular noun with a plural pronoun and verb, and one graduate of Vassar split an infinitive. Another student wrote, “the purpose tested,” which was obviously a slip of the pen for the *person* tested. Misspellings were not numerous: *heridity*, *benefical*, *develope*, *envoironment*, *memorary*, *desition*, and *medocrety*.

One candid undergraduate wrote at the bottom of his paper, “I did not hear the whole dictation.” Judging from the way he rendered what he did hear, his failure to hear the rest was probably due to lack of attention. Another undergraduate, a sophomore in arts, coined this charming phrase,—“So life is neither unjust or too goo but a medium of mediocrity.”

Four women, three of them graduate students and one a student in the college courses for teachers, were present on two of the occasions when the tests were given, and therefore were represented by two papers. In order to put them on a footing of equality with the rest of the subjects, only their first papers were ranked and included in the results in tables II to IV. All four of them made distinctly better marks on the second trial. One raised her score from 7 points to 9, one from 7 to 11, one from 9 to 10, and one from 5 to 9. This would seem to show a certain amount of practice

effect, or assistance from the recollection and correction of omissions made in the first testing. It is remarkable that the woman who made the most improvement, rising from 5 points to 9 points, was already quite familiar with the substance of the passage. In a note to her second paper she says, "I have given this paragraph in the Kuhlmann tests a number of times. This time I tried to remember it verbatim, since I had given it in 62 yesterday. Pretty poor repetition, I should say." On her first paper she had confidently written, "Familiarity with test makes it a pretty simple visual memory (!), verbal too. There is no auditory but some motor elements."

SPECIMEN PAPERS.

The fullest paper in selection A was the following, which had the additional merits of being neatly written, correctly spelled and well punctuated:

"Tests of this nature are of value both for science and for the information of the individual to whom the test is applied. It is of importance for science to observe the differences in individuals and to find out which factors they are due to. If it can separate what is due to environment from what is due to heredity, it can in some cases correct defects which would be otherwise neglected and in others prevent the development of peculiarities which might otherwise hinder the progress of the individual."

This is somewhat transposed, but the meaning is not distorted and only two out of the fourteen ideas are omitted.

The following paper is typical of a large group. It ranked number 26, received a G, and contained nine out of the possible fourteen points:

"These tests are of value not only for the advancement of science but for the information of the individual. We may find the various factors influencing development. If we could separate the influences of environment from those of heredity we might gain certain facts concerning the behavior of the individual, which we might otherwise neglect."

Here the transpositions modify the meaning to a considerable extent, but do not contradict it. When we come down to some papers in the lower range of the series, this number 26 quite shines by comparison.

One man who caught very little of selection A wrote:

"The tests which we are now making are not only a test of the student's ability but a test made for the advancement of science." But compare with this the same man's paper for selection B, which contained all seven of the essential ideas:

"There are many opinions on life. Some call it good and some call it bad. Rather it should be called mediocre, for life is never as good as we would have it, nor never as bad as it might be. It is this mediocrity that makes life just rather than unjust."

A paper which seems to betray a state of mental confusion is the following:

"Tests are of value both for the scientific side and informational. That is to know how much information one has, and for the advancement of science. If we can find out how one acts for the tests, we can test their ability. If we test their parents' ability we can get hereditary information. This aids in the advancement of science." This paper received credit for three ideas in fourteen.

But compare with it this paper, which, although it contains only one more of the fourteen ideas, must be regarded as the work of a very much abler mind:

"Tests such as we have been performing are of value (1) for the information of the individual, and (2) for the advancement of science."

And finally, here is the remarkable effusion of the lowest ranking student in selection B:

"Things are never really what they seem or what we would wish them to be, or think they should be, it is here where the seemingly unjustness of the world comes in, but to live, to work sore, and not to receive in return just compensation for energy expended in an honest task is one of the evils of modern progress, under such progress mediocrity may become great, and greatness mediocre."

CONCLUSIONS.

The conclusions are negative, with unimportant exceptions:

(1) The results of these two very similar tests, given to the same students under the same conditions, do not correlate with each other.

(2) They do not correlate with college standing, as shown in the records of the Dean's office in the few cases where these were adequate for comparison. They do not even correlate with the term marks in the subject (Psychology 2) in which they were given as a class exercise.

(3) Women do better in these tests than men. This may point to the fact that women use verbal imagery of various sorts to a greater extent than do men. It probably means that women are better able to comprehend and transmit a spoken message.

(4) These tests are not satisfactory as diagnostic tests, when made under group conditions. They may well be stricken out of the Binet tests for fifteen years and over, unless given under standard Binet conditions. They might yield highly misleading results when not applied—as they are certainly not being applied

by many Binet testers—with the skill and discrimination of the experienced clinical psychologist.

(5) It seems probable that the most favorable results in the reproduction of selection A are represented by the papers scoring from six to ten points in a possible fourteen on the first trial. The better students, both graduate and undergraduate, made scores within this range. A lower score than six probably indicates on the part of the individual student some general incapacity for college work, or some temporary disability, as a severe headache, or an unusual noise outside the class-room, which interfered with attention. A student who makes too high a score in these tests, is very probably lacking in those habits of selection and condensation which are requisite to the taking of lecture notes. The better notes a student has trained himself to take, the less of this diffuse and insignificant material he will be able to hold in mind, down to a limit which seems to be represented by about six points for selection A. The test needs to be supplemented by other clinical observations of a student's behavior, before it can be determined whether the omissions are due to trained habits of selection, or merely to fluctuations of the untrained attention.

BIBLIOGRAPHY

1. BINET, A., and SIMON, TH. A Method of measuring the development of the intelligence of young children. Trans. by Clara Harrison Town. Lincoln, Ill., 1912.
2. BINET, A., et HENRI, V. La mémoire des phrases. *L'Année Psychologique*, I, 1894, 24-59.
3. BLISS, D. C. Some results of standard tests. *THE PSYCHOLOGICAL CLINIC*, vol. vi, no. 1, March 15, 1912, pp. 1-12.
4. COLVIN, STEPHEN SHELDON, and MYERS, E. J. The Development of imagination and the relation between ideational types and the retentivity of material appealing to various sense departments. *Psychol. Monog.*, vol. xi, no. 1, pp. 85-126 incl. Nov., 1909. (Bibliography.)
5. FINKELSTEIN, I. E. The Marking system in theory and practice. Baltimore: Warwick and York, 1913. *Educational Psychol. Monog.*, no. 10. (Bibliography.)
6. JEVONS, W. STANLEY. Elementary lessons in logic. N. Y.: Macmillan, 1902.
7. KUHLMANN, F. A Revision of the Binet-Simon system for measuring the intelligence of children. *Journ. of Psycho-Asthenics*, Monog. Supp., vol. 1, no. 1, 1912.
8. MEUMANN, ERNST. The Psychology of learning. Trans. by J. W. Baird. N. Y.: Appleton, 1913. Pp. xix+393. (Bibliography.)
9. TERMAN, LEWIS M., and CHILDS, H. G. A Tentative revision and extension of the Binet-Simon measuring scale of intelligence. *Journ. of Educa. Psychol.*, March, April, and May, 1912.
10. Condensed guide for the use of the Stanford revision and extension of the Binet-Simon measuring scale of intelligence. Privately printed, 1915.