

## VOCABULARY AND WORD-BUILDING TESTS.

BY PROFESSOR GUY MONTROSE WHIPPLE,  
*Teachers' College, University of Missouri.*

At the meeting of the American Psychological Association in December, 1905, Professor Kirkpatrick gave a brief account of a simple vocabulary test, which he has since described more fully.<sup>1</sup> This test I have applied with some modifications to two classes of sophomores and juniors at Cornell University; the discussion of the results of this test forms the first section of this paper. In the second section there is described another test, that of word-building, which exhibits certain points of similarity to the vocabulary test. Attention is also paid to the correlation between these two tests, and between them and school standing.

### A. KIRKPATRICK'S VOCABULARY TEST.

This test consists, in principle, in the examination of a limited number of words taken by chance from the pages of a dictionary, and the determination of the percentage of these words whose meaning is known, unknown, or doubtful. On the basis of this examination, the 'understanding vocabulary' is estimated at that percentage of the total number of words in the language that the test has indicated. Naturally, the longer the list examined, the more accurate is the index secured. Kirkpatrick is, however, convinced that 100 words form a sufficiently lengthy list upon which to compute the index in question, at any rate, sufficiently well to be representative of one's vocabulary for comparative purposes.

*Purpose of the Test.* — My object in using the test was: (a) Primarily to examine its reliability, *i. e.*, to see whether college students would be apt to overestimate or underestimate

<sup>1</sup> *Pop. Sci. Mo.*, LXX., February, 1907, 157-164. The examination of the test with which this paper deals was nearly completed when the originator's paper appeared, and although he has anticipated therein certain points of criticism, I have published my results in their first form. The reader may also consult E. H. Babbitt's article in the same volume, p. 378.

their vocabulary in the absence of a strict definition-check; (b) to determine the average vocabulary of college students; (c) to determine the correlation of the corrected vocabulary with ability to construct words in the word-building test; (d) the correlation with standing in college classes, and (e) the correlation with sex.

*Nature and Method of the Test.*—(1) The following list of words, the same used by Kirkpatrick, was distributed in the class, with these instructions: Place a plus sign (+) before all the words you know, a minus sign (−) before those you do not know, and a question mark (?) before those of whose meaning you are in doubt. When through, count, and fill out these blanks. Number known . . . Number unknown . . . Number doubtful . . .<sup>1</sup>

abductor	escheat	paid
abeam	escort	pail
abed	eschalot	publication
abet	gourd	pudding
baron	gout	puddle
baroscope	govern	puddy
barouche	gown	scroll
barque	intercede	scrub
channel	interdict	scruff
chant	interest	scrunch
chanticleer	interim	subcutaneous
chaos	matting	sub-let
decemvirate	mattock	subdue
decency	mattress	tycoon
decide	maturate	tymbal
deception	page	type
eschar	pagoda	
amalgamation	filing	photo-lithograph
amanuensis	fill	rejoice
amaranth	hodman	rejoin
bottle-holder	hoe	rejoinder
bottom	hoecake	rejuvenate
bottomry	hog	skylight
boudoir	lanuginose	skyrocket

<sup>1</sup> The request to write the opposite of a series of ten words, to define ten words, five of which were in the list, and to indicate the magazines and books read within a specified time, which formed a part of the Kirkpatrick test, was omitted, in part to save time, in part because it was desired to see whether the test could be used with safety without these devices for instilling an attitude of cautiousness in the student.

concatenate	lanuginous	skysail
concatenation	lanugo	skyward
concave	lanyard	tenderloin
conceal	muff	tendinous
disentomb	muffin	tendon
disentrance	muffle	tendril
disepalous	mufti	virago
disestablish	photograph	virescent
filiform	photographer	virgin
filigree	photography	

(2) To arrive at a correct vocabulary-index and to detect the presence of a tendency either to overestimate or underestimate, I required the students to write the definitions of forty words selected from the above list *after* they had completed the first test. This definition test was, of course, entirely unexpected by the students. It was intended to require the definition of every word that might possibly be unknown or doubtful, but, as events showed, to accomplish this result the list would have to be extended to embrace sixteen additional words.

The definition list, for which twenty minutes was allowed, was as follows: abductor, abet, baroscope, chanticleer, chaos, decemvirate, eschar, escheat, eschalot, gourd, interdict, interim, mattock, maturate, pudgy, scruff, scrunch, subcutaneous, tycoon, tymbal, amalgamation, amanuensis, amaranth, bottomry, concatenate, disentrance, disepalous, disestablish, filiform, hoecake, lanugo, lanyard, mufti, photo-lithograph, rejoinder, skysail, tendinous, tendril, virago, virescent. Concatenation, lanuginose and lanuginous can, of course, be checked off by the definitions given for concatenate and lanugo.

The words which experience showed should have been added were: abeam, abed, barque, barouche, boudoir, disentomb, filigree, hodman, pagoda, rejuvenate, scroll, sub-let, tenderloin, — a rather amazing list for college students!

*Results.* — In the vocabulary test proper, 70 students, aged 16 to 25 years, 16 men and 54 women, on the average, marked 77.6 per cent. known, 17.2 per cent. unknown, and 5.2 per cent. doubtful. The highest known was 92 per cent., the lowest 64 per cent.

In the second part of the test, the actual definitions of the 40 selected words, the results were treated in the following manner. A record was made for each student of (1) the number of words not defined, and (2) the number of words wrongly defined. (3) These were added to give the total number of words unknown in the list of 40. (4) By reference to the

vocabulary test the number of other words unknown beside those in the definition test was ascertained (this was from two to ten). (5) There remained in some instances (33) from one to three words each (averaging for all those tested 0.6) that had originally been marked doubtful, and which still remained doubtful because they were not on the definition list: these scattering doubtful cases were regarded as unknown. (6) By combination of all the unknown cases, it was then possible very easily to compute the final corrected vocabulary-index of each student, and (7) by comparison with the original markings, to determine the amount of over- or under-estimation.

The results, based on 70 cases, are as follows:

Average number of words in list of 40 not defined.....	15.50
Average number of words in list of 40 wrongly defined.....	6.64
Average number of words in list of 40 not known.....	22.14
Average number of words outside list of 40 not known.....	4.20
Average number of words left doubtful.....	0.60
Average corrected vocabulary-index.....	73.26 (m. v. 5.5)
Average amount of overestimation.....	4.37
Number of students overestimating.....	59
Number of students underestimating.....	10
Number of students neither over- nor underestimating.....	1
Largest overestimation.....	18
Largest underestimation.....	4

The decided nature of the tendency towards overestimation is more clearly evident from its distribution as given in Table I.

TABLE I.

OVERESTIMATION OF THE VOCABULARY-INDEX (70 COLLEGE STUDENTS).

Per cent. Overestimated.	Number.	Per cent. Overestimated.	Number.
18	1	5	7
15	2	4	6
14	3	3	7
13	1	2	10
12	3	1	9
11	1	0	1
9	2	-1	6
8	2	-2	2
7	1	-4	2
6	4		

It will be seen that, of 70 students, 11 overestimate by more than 10 per cent., and that 20, or more than one quarter of the

students, overestimate by 5 per cent. or more. From these figures it is evident that, without a somewhat elaborate definition-check, the value of the vocabulary-test is distinctly lessened.<sup>1</sup>

The corrected vocabulary-index is of sufficient interest to justify a brief presentation of its distribution (Table II.).

TABLE II.

DISTRIBUTION OF THE CORRECTED VOCABULARY-INDEX  
(70 COLLEGE STUDENTS).

Index.	Number of Cases.
85-89	5
80-84	5
75-79	19
70-74	22
65-69	13
60-64	6
55-59	1
Highest index,	89.0
Average index,	73.26
Lowest index,	58.0

Sex differences in vocabulary cannot with certainty be established from the small number of cases here considered, but the sixteen men average 75.8 and the 54 women 72.6 for the corrected index.

On the basis of Webster's Academic dictionary, circa 28,000 words, Kirkpatrick estimates the average vocabulary of college students at 20,120. From the present test, the students' own estimate figures 21,728, while their corrected vocabulary would be 20,512. The largest vocabulary would be 24,920, the smallest 16,240, or approximately that assigned by Kirkpatrick to the average second-year high-school student. Thus the wide individual variation in size of vocabulary to which Kirkpatrick calls attention is still further confirmed and emphasized.

The results of the definition-test are worth considering in

<sup>1</sup>This result may be compared with Kirkpatrick's conclusion that very young children are apt to underestimate because the isolated words of the list fail to arouse associations such as they would if they had a context. Again, when Kirkpatrick defined the words of the list to normal school students, he found that the errors of over- and under-estimation tended to cancel one another, while when college classes defined 20 words, 114 of 246 students (about 46 per cent.) correctly defined the same proportion that they had marked as known, and only 7 per cent. erred by as much as 3 in 20.

more detail. No word of the forty was correctly defined by every student, and, as we have noted, there were 16 other words that were unknown or doubtful. It follows, therefore, that *only 44 of the 100 test words were certainly known by every one of 70 college students.*

It is at times rather difficult to decide from the definitions whether the student does or does not know the meaning of a word with sufficient exactness to be credited with knowledge of the term in question.

To a slight extent, therefore, the ranking is arbitrary and might not agree with that of another experimenter who was using the same test. In general, however, I have erred on the side of leniency in consideration of the difficulty of accurate definition and of the short time (about twenty minutes) that could be expended upon the forty definitions. To illustrate, the following were accepted: 'disestablish — to overthrow,' 'decemvirate — a body of ten,' 'mattock — a garden tool,' 'amaranth — a flower'; while the following were disallowed: 'lanyard — one of the spars of a ship,' 'decemvirate — Roman civil officer,' 'gourd — a hollow vessel from which to eat and drink,' 'concatenate — to argue,' 'baroscope — an instrument for measuring something.'

Of interest, in this connection, are the erroneous definitions. These, as Kirkpatrick has pointed out, are most likely to result from the confusion of the given word with some word resembling it in sound or spelling. Other definitions are evidently sheer guesses from the fancied etymology of the word, *e. g.*, disepalous — without a head (from *dis* + *cephalous*). In the following list of typical errors of definition, the assumed source of confusion is indicated by the terms in parentheses after the definitions:

- amanuensis — poet laureate, lovingness (amativeness).
- amaranth — a precious stone (amethyst).
- abet — although (albeit), a wager (a + bet), diminish (abate).
- bottomry — the art of bottoming chairs, deceit, bottom of anything.
- chanticleer — one who sings a loud song, one who leads a chant.
- decemvirate — composed of five, count out by tens, formerly a group of ten men, but any number now.
- disentrance — failure to enter.
- disepalous — apart from the head, without shoulders.
- gourd — reward (guerdon), to slash or whip (goad), morning glory.
- interim — time between two reigns (interregnum).
- lanugo — a kind of language.
- lanyard — yard where leather is tanned (tanyard), yard about the lane.
- mattock — a lock of hair (matted locks?), a kind of bird, a sort of rug, a kind of robe (cassock).

maturate — to ripen (mature), to matriculate.

sky-sail — a sail in the sky, a kite.

tycoon — a violent wind (typhoon), an animal, a silk-worm, a natural phenomena (sic).

tendril — a membrane connecting two bones (tendon).

tendinous — capable of endurance (tenacious?).

scrunch — a good for nothing person, (scrug?).

virago — a kind of bird (!) (vireo), a disease, giddiness (vertigo).

virescent — sparkling (iridescent), of or pertaining to a man (!) (virile).

## B. A WORD-BUILDING TEST.

The word-building tests were suggested to me by the familiar game of anagrams as well as by the advertisements often seen in magazines in which a prize is offered to the person who can make the most words from a given word or series of letters. This test is easily administered and evaluated: it is one that calls for ingenuity and active attention: it might fairly be said to demand that ability to combine isolated fragments into a whole, which Ebbinghaus has declared to be the essence of intelligence and for the measurement of which he devised his well known 'combination method';<sup>1</sup> and, finally, its execution is conditioned to a certain extent by the richness and readiness of the examiner's word-vocabulary. One may expect, therefore, to find a correlation between this test and the vocabulary test and possibly between it and school standing or general intelligence.

*Nature of the Test.* — After some preliminary trials, two word-building tests were arranged and distributed to the classes above mentioned in the form of mimeographed blanks which read as follows:

### WORD-BUILDING TEST, NO. I.

Make as many words as you can from the six letters given below. You may use any number of letters from one to six, but no letter may be used twice in the same word, and no other letters than these six are to be used. You will have five minutes.

a e o

b m t

<sup>1</sup>H. Ebbinghaus, 'Ueber eine neue Methode zur Prüfung geistiger Fähigkeiten und ihre Anwendung bei Schulkindern,' *Zeits. f. Psych. u. Physiol.*, XIII., April, 1897, 401-459.

Then followed numbered spaces for 30 words, and blanks for name, date, etc. The second test was exactly similar save that the six letters given were: *e, a, i* and *r, l, p*.

The possibilities of combination are larger than one might suppose: from the first at least 70, from the second at least 105 words may be formed. Proper names were allowed, but words in foreign languages were not allowed.

Test No. 1 was given to two groups of college students and likewise to a group of fifty boys from the seventh and eighth grades of the Ithaca public schools.<sup>1</sup> Test No. 2 was given to the college students only.

*Results.* — In Table III. the chief results of these tests are summarized. It will be seen that Test No. 2 offers more possible combinations, that more words are made with it and a greater

TABLE III.

Test.	Date.	Grade.	Male.	Female.	Largest.	Smallest.	Average	Mean Variation.	Total No. Different Words.	Possible Words.
1	2/19,06	College	7	15	25	10	18.6	3.54	41	70
1	3/19,07	"	9	27	26	10	18.6	3.30	41	70
1	4/10,06	Grammar	50	0	21	6	12.4	2.37	38	70
2	3/ 6,06	College	7	26	32	14	21.9	4.00	57	105
2	3/19,07	"	9	27	33	15	23.5	3.66	61	105

number of different words. It will also be noted that the individual differences in rank are large; thus it happens that not a few grammar school boys make more words than some of the college students: to be more explicit, it may be stated that, with Test No. 1, 10 grammar pupils make 15 words or over, while 13 college students make fewer than 15 words.

By examining the papers in detail, and tabulating the total number of words formed and the number of times each of these words is given, one may discern something of the principles which govern the operation of the test. The following are the data thus secured:

TEST NO. 1. 58 COLLEGE STUDENTS. (45 DIFFERENT WORDS.)

Over 50 times — bat, mat, bet.

40-49 times — eat, met, Tom, at, boat.

<sup>1</sup>This was in connection with an extensive investigation upon the correlation of various physical and mental tests, the results of which I hope to publish soon.



30-39 times — meat, to, tea, beat, team, tab, ate, am, moat, mob, me, beam, toe.  
 20-29 times — tame, oat, be, mate.

10-19 times — boa, mote, bate, abet, tomb, tome, tam.

5- 9 times — Mab, Abe, Mae, ma, atom, a.

1- 4 times — bot, mot, o, Moab, beta, bema, tabe.

Not given<sup>1</sup> — ab, ambe, ambo, amt, atmo, ba, bam, bo, bom, boma, bote, ea, eam, eb, em, eta, mao, meta, mo, moa, moe, ob, obe, om, ta, tambo, tema,

TEST NO. 1. 50 GRAMMAR GRADE BOYS. (38 DIFFERENT WORDS.)

Over 40 times — mat, bat.

30-39 times — bet, at, met.

20-29 times — to, eat, Tom, beat, tea, meat, be, am, boat.

10-19 times — toe, mob, beam, me, ate, team, tab, boa, oat.

5- 9 times — ma, bate, a, moat, mot, tame, mate, bot.

1- 4 times — tam, tomb, Abe, mote, Moab, Mae, o.

Not given — those not given by college students, plus abet, atom, bema, beta, Mab, tabe, tome.

TEST NO. 2. 69 COLLEGE STUDENTS. (66 DIFFERENT WORDS.)

Over 60 times — lip, lap.

50-59 times — rip, rap, pear, ear, real, pie, leap, rail, pale, reap.

40-49 times — reap, pail, pile, ale, pair, arc, ape, lie, pea, peal.

30-39 times — pare, earl, pearl, air, par, lair, ripe, liar.

20-29 times — ail, Lear, rape, ire, pal.

10-19 times — lea, pa, rile, pire, era, pier.

5- 9 times — per, a, alp, Eli, plea.

1- 4 times — I, paler, peril, lira, rep, rale, ile, lare, ra, pil, piler, ril, April, Ira, la, pareil, pi, pilar, Rea, Rae.

Not given<sup>1</sup> — ai, aiel, aile, aire, al, apl, ar, Ariel, aril, ea, el, ela, epi, er, eria, il, irp, le, lep, lepra, lerp, li, lier, lire, lirr, paie, pali, parel, parl, pela, pel, pia, piel, pila, plie, plier, prial, prie, re, rei, rial, ril, ripa.

Inspection of these lists shows (1) that three-letter words are in every instance those most frequently formed, (2) that two-letter words and the one-letter words, which one might expect to be most frequent since most simple, stand relatively low, *e. g.*, *ma*, *be*, *am*, *pa*, *me*, *a*, *o*, *I*,<sup>2</sup> (3) that grammar school boys give all the words given by college students save a few rather unusual terms such as *atom* and *tome*, (4) that usage and ordinary speaking vocabulary condition the formation of words, in as much as the most ordinary words have the greatest fre-

<sup>1</sup> This list is based on the words actually given in the Standard dictionary, not including, however, Scotch terms.

<sup>2</sup> It appeared, upon inquiry, that some of the college students had omitted words like *pa*, *ma*, *a*, *o*, and *I* on the ground that they were 'not real words,' or 'didn't count,' but, oftener, they seem to have been passed over because the attention was concentrated upon the making of combinations.

quency, *e. g.*, *bat, mat, bet, cat, lip, lap*, whereas words that are less frequently used in every-day speech, although their meaning is doubtless perfectly well known, do not suggest themselves so readily under the conditions of the test, *e. g.*, *tomb, tome, era, plea, paler*, (5) that the words not given by any one are, with one or two exceptions, *e. g.*, *plier*, words of extremely rare usage or unusual form, alternative spellings, etc.

Sex differences may not with certainty be made out, yet, as Table IV. indicates, in every group tested, the men did slightly better than the women: though this difference is small—less than the mean variation, yet it is constant in direction in all four instances.

TABLE IV.  
SEX DIFFERENCES IN WORD BUILDING.

Test.	Date	Men.	Average	Women.	Average.
1	1906	7	18.7	15	18.6
1	1907	9	19.7	27	18.0
2	1906	7	23.6	26	21.4
2	1907	9	25.8	27	22.7

### C. CORRELATIONS.

In the case of the college students I have been able to test the correlation of vocabulary and word-building test with one another and of each test with the class standing in Educational Psychology. This class standing was based upon marks obtained in examination and upon class exercises, and was secured with suitable precaution, at least in the case of the examinations, to render it absolutely objective.

With the grammar-school boys, it was possible to compute correlations between word-building and numerous other tests. Save for class standing the latter are reserved for a later paper. All of the correlations were worked by the well-known Pearson formula, save that when the distribution was symmetrical, the value of sigma (the standard deviation) was computed from the average deviation by multiplying by the constant 1.2533—a procedure which cuts the work of computation by one half and gives values closely approximating those obtained by the usual method of computing the standard deviation.

TABLE V.  
CORRELATIONS.

Data Correlated.	Cases.	r.	p.e.
Vocabulary index and word building .....	58	+.53	.05
Vocabulary index and class standing .....	58	+.45	.06
Word building and class standing .....	58	+.13	.07
Word building and class standing .....	50	+.08	.08

#### D. SUMMARY.

1. The determination of the vocabulary-index, as proposed by Kirkpatrick, by means of a list of 100 words selected by chance is clearly influenced by a tendency to overestimation when no precautionary restrictions are employed. In tests of college classes, 15 per cent. of those tested had an overestimation error amounting to 10 per cent. or more, and 25 per cent. of those tested had an overestimation error of 5 per cent. or more.

2. By check definition-tests it is possible to eliminate or to measure this error of overestimation. When such a check was applied, the average vocabulary-index for college students was found to be 73 per cent., with a maximum of 89 per cent. and a minimum of 58 per cent. This minimal index is less than that assigned by Kirkpatrick for the average second-year high-school student, and illustrates forcibly the unevenness of the mental equipment of college students.

3. The definition-test reveals an unexpectedly large number of erroneous definitions. The source of these errors may frequently be traced to confusions with words of similar appearance or to fancied etymological derivations.

4. Simple word-building tests correlate fairly well with the vocabulary test. Marked individual differences are apparent here, as in the vocabulary test; so marked, indeed, that some grammar-school pupils excel some college students in the construction of these lists.

5. An inspection of the lists of words thus constructed reveals the psychological conditions which affect the process. In general those words are most frequently given that are in most common use in the daily speaking vocabulary, whereas

words that are equally well known but in less frequent use are apt to be omitted.

6. Sex differences cannot be clearly established in these two tests, but what differences appear are constant in direction and suggest the superiority of men and boys over women and girls.

7. The vocabulary test correlates fairly well with class standing, but the word-building test rather unexpectedly shows no such correlation.<sup>1</sup>

<sup>1</sup>The MS. of this article was received November 15, 1907. — ED.