

AN ANALYSIS OF SEVERAL WELL-KNOWN TESTS

H. E. GARRETT AND V. W. LEMMON

Columbia University

This paper is an attempt to analyze, through the technique of partial correlation, two well known cancellation tests—the A-test and the a-t test—and the equally well known Color Naming test. Whipple in his “Manual of Mental and Physical Tests” says in the introduction¹ that “What we need is not new tests, though they are welcome enough, but an exhaustive investigation of a selected group of tests that have already been proposed.” And even a superficial examination of the literature of tests will convince one that very often indeed, workers with tests have used them with no clear understanding of what the tests measure, or—what is worse—with the idea that a test actually measures say, attention, or perception, or reasoning, because it is so labeled. Indications that even if we do not know exactly what a test measures, at least we realize it, are seen in the tendency of present-day workers to call a test by some descriptive title rather than in terms of some more or less vaguely defined mental function which it may be thought of as measuring; and so we have cancellation tests, opposites tests, completion tests, etc., instead of tests of perception, association, or reasoning.

It seems to the writers that the next step is to separate out, so far as this can be done, the various factors which make for efficiency of performance in a given test. The better these factors can be identified, the better able we are to say what our test measures. Rosenow² has shown in a very complete

¹ Manual of Mental and Physical Tests, Simpler Processes, 1914, p. 4.

² The Analysis of Mental Functions, Psych. Mono., 1917, vol. xxiv, no. 106.

fashion how the method of partial correlation may be used in analyzing mental tests and the purpose of this paper is to show how the method may be employed to find some, at least, of the factors which condition the performance of tests, which, though generally considered simple, are often very differently interpreted.

I. CANCELLATION TESTS

Though generally admitted to be among the simplest tests that we have, tests of cancellation have presumably measured different functions for different investigators. Generally,³ such tests have been known as tests of "discrimination" or "attention"; with variations, they have been called tests of "speed of perception," "efficiency of perception," "speed and accuracy of perceptual discrimination," "time of discrimination, association, and movement." More specific information can be got by studying the correlations of cancellation tests with other tests. Cancellation correlates uniformly low with tests of general intelligence, and very seldom higher (and often much lower) than 0.40 with such tests as analogies, word-building, completion etc. While these correlations tell us something about cancelling ability, they leave much to be desired in the way of positive information.

If we look at cancellation from the standpoint of one taking the test, it would seem that we are able to discover at least two factors upon which efficiency in the tests depends. First, there is certainly present the factor of perception or recognition of the symbol or symbols to be marked; and secondly there is always present the factor of motor movement—the simple crossing out of the symbol after we have found it. Vogt⁴ has attempted to separate out the motor movement factor by comparing the amount covered in cancellation of the usual kind, and the amount covered when the symbols are simply recog-

³ Whipple, *Manual of Mental and Physical Tests, Simpler Processes*, p. 305 ff.

⁴ Whipple, *Manual of Mental and Physical Tests, Simpler Processes*, p. 320

nized but not marked. He reported that the marking movement accounted for from 15 to 40 per cent of the time required depending on the practise of the subject. This method would seem to be the logical way of getting at the marking factor; provided the test, except for the absence of marking, remains in all other respects the same. That this is not true, however, anyone taking the test can easily discover for himself. Interferences, difficulty in holding the place, or marking progress, arise to make the test a *new* test, and not simply cancellation minus the marking movement. It was to avoid this "fallacy of subtraction" that partial correlation was used as a means of analysis.

We began with the idea of finding tests which would measure the two factors, (1) movement involved in marking, and (2) simple recognition involved in finding the symbol to be cancelled. With this end in view, we gave the A-test and the a-t test to a class of 54 undergraduates in Columbia College. The task in the first test is to mark out every A (there are 100 A's) in a paragraph of pied capital letters. In the second test, the subject is required to mark out every word which contains both an *a* and a *t* from a page of Spanish text. Following these, several other tests were given:

1. *A test of Rapid Motor Coördination.* This test, devised and named by C. K. Taylor, requires the drawing of four short vertical lines crossed by a horizontal line (in the form of a "gate"). The measure of performance used by us was the time taken to make 100 little "gates" on a sheet ruled in ten columns and ten rows. This test is a measure of simple motor activity, and compares with other tests of movement such as making dots, dealing cards, etc.⁵ It seemed to us that this test should measure, with a minimum of other factors involved, the simple marking employed in cancellation.

2. *A test of Rapid Recognition.* (Also devised and named by Taylor.) In this test, the subject is given a page of jumbled numbers running from 1 to 50, which he is to join by straight lines, going from each number to the one which follows it directly.

⁵ Franz, Handbook of Mental Examination Methods, 1919.

3. *A test of more complex recognition.* In this test the task was to locate and mark all of the three-letter words found in a paragraph of unspaced and pied material; and directly afterwards to locate and mark all of the four-letter words in a second paragraph. There were twenty-four words in each paragraph. The total time taken on the two paragraphs was recorded and a correction added for each word omitted. Three months after the first giving of this test, it was repeated on the same group, this time with the marking movement omitted. The students were instructed to note each word but not to mark it, and as a check were required to write down the number of words found at the end of the test.

TABLE 1
(All records are in seconds)

	TESTS:					
	A-test	a-t test	R M C	Rapid recognition (not used)	Recognition of words with marking	Recognition of words without marking (not used)
Average	101 28	209 48	122 82	239 72	179 37	314 59
SD	20 07	48 02	14 40	65 20	36 97	67 48

The averages and the SD's for each of these tests are given in table 1. The average of 101.28 seconds on the A-test is closely in line with the results of Whitley⁶ and other investigators who have used this test with college students. In table 2 are given the correlations of the A-test and the a-t test with each of the other tests. None of the correlations are very high (though all are positive), the correlation of the A and the a-t tests—e.g. 0.62—being the highest. Apparently none of our tests overlap very completely either of the two cancellation tests, though the recognition of words seems to be more closely related to the cancellation tests than the rapid motor test.

⁶ Tests for Individual Differences, Archives of Psych., no. 19, 1911,

The next step was to work out a regression equation which would give the relative importance in cancellation of the marking movement factor and the recognition factor. It seemed to us fair to take the test of Rapid Motor Coordination (known hereafter as r.m.c.) as representative of the marking activity which is present in cancellation, but in selecting a test which would represent the recognition element, the task was not so easy. The so-called Rapid Recognition test was first discarded because it did not seem to be exactly the same sort of recognition which we find in cancelling, and further because it

TABLE 2

	A-TEST	B-T TEST	R M C.	RAPID RECOGNITION	RECOGNITION OF WORDS WITH MARKING	RECOGNITION OF WORDS WITHOUT MARKING
A-test		0 62	0 37	0 61	0 43	0 02
a-t test	0 62		0 22	0 38	0 53	0 38
R.M.C.	0 37	0 22		0 31	0 16	0 18
Rapid Recognition	0 61	0 38	0 31		0 27	*
Recognition of Words with marking	0 43	0 53	0 16	0 27		0 42
Recognition of Words without marking	0 02	0 38	0 18	*	0 42	

* No correlations found.

is complicated by the drawing of lines, and hence, to some degree at least, is dependent on movement. It would look as though the Recognition of Words, without marking, would be the best test of the recognition of certain symbols, minus the marking movement. Further consideration, however, caused us to decide against this form of the test in favor of the test, as first given, in which the subject was required to mark the words when found. In the first place, the marking of the words, as reported by the subjects and checked by observations of the experimenter, served rather as a means of keeping the place

than as a distraction or as a kind of motor movement. That it actually facilitated recognition is shown by the fact that the total time taken to mark a blank was, on the average, nearly twice as great for the test without marking as for the test with marking; and further by the evident difficulty which the men had in keeping the place when no marks were made. Several students used their finger or followed the lines with a pencil in order to keep track of the words as they were recognized. All reported that the test was far less difficult when the marking was permitted. The relatively small part which speed of marking, and the relatively great part which ability to find the words in the complex material, play in determining one's score seemed to recommend this test as a fair measure of that recognition, whose influence in cancellation we wished to find. From the statistical standpoint we believed that the two tests selected, r.m.c., and Recognition of Words were satisfactory, in that they are highly enough correlated with the two cancellation tests to indicate the presence of common factors, and yet they are not so highly related to them as to be merely measures of the same function.

The first regression equation was worked out from the following variables:

1. A-test.
2. Rapid Motor Coördination test.
3. Recognition of Words (with marking).

The partials and the regression equation are given below:

$$r_{12} = 0.34$$

$$r_{13} = 0.40$$

$$r_{23} = 0.00$$

$$x_1 (\text{A-test}) = .31x_2 (\text{Rapid Motor Coordination}) + 0.38x_3 (\text{Recognition of Words})$$

In this equation the SD's of all of the distributions have been taken equal⁷ in order to show the relative contribution of the two tests to the A-test, irrespective of the unit of measurement used in either test. The regression equation shows that the two factors are nearly of equal importance. We are able to

⁷ T. L. Kelley, Bulletin University of Texas, 1916, 27, p. 8.

say, therefore, that any given score on the A-test is contributed to equally by the marking-movement and the recognition of the symbols, in so far as these factors enter into the A-test. Presumably, many other factors are present and unaccounted for by the two tests above. Several facts should be considered in the light of these conclusions.

1. The raw correlation between the A-test (1) and r.m.c. (2), e.g., $r_{12} = 0.37$, is only slightly affected by ruling out the Recognition of Words test (3), e.g., $r_{12.3} = 0.34$. The raw correlation between the A-test and the Recognition test, e.g., $r_{13} = 0.43$, is also only slightly affected by ruling out r.m.c., e.g., $r_{13.2} = 0.40$. This would indicate that both (2) and (3) actually do measure something which is present in the cancelling test independently of each other, while their partial correlation, $r_{23.1} = 0.00$, would indicate that they measure different aspects or phases of the total process.

2. The relative importance of the two factors as we have given them holds only for rapid marking and recognition as defined by our tests. The multiple correlation coefficient $R_{1(23)} = 0.52$ serves to dissipate any delusions that we may have that we have measured *everything* that there is in the A-test. In so far as our tests do measure speed of marking and ability to find certain designated symbols, however, it seems reasonable to conclude that these two factors play an equal rôle in determining the score on the A-test.

3. To one who has taken or given the A-test many times, we do not believe that it will seem exceptional that the marking movement should be so important. The finding of the A's is a simple sort of recognition, done with little hesitation or search except the first time the test is performed. Increase of speed with practise would then (to hazard a guess) be due to the increasing mechanization of the "finding process" and the consequent changing of the test into a simple marking exercise. Our equation holds only for the *first* performance; as practise goes on, the relative weight of the two factors would probably shift, the marking becoming more important, and the recognition less important.

The procedure which we followed with the A-test was repeated with the a-t test. The three variables are

1. a-t test.
2. Rapid motor coordination test.
3. Recognition of words (with marking.)

The partial correlations and the regression equation are as follows:

$$r_{12.3} = 0.17 \quad r_{13.2} = 0.51 \quad r_{23.1} = 0.05$$

$$x_1 (\text{a-t test}) = 0.15x_2 (\text{r.m.c.}) + 0.50x_3 (\text{Recognition of Words})$$

The SD's have been taken as equal for the reasons given previously. This equation indicates that the two factors are differently related in the a-t test than in the A-test. In a-t, recognition accounts for three times as much of the score as the marking movement. This result is certainly in line with common sense experience with the test. Finding the words to be cancelled in the a-t test is much more difficult than finding the A's to be marked in the A-test. The low raw correlation of a-t and r.m.c. and the slightly lower partial correlation with recognition out, gives little weight to the simple marking factor; while the practically zero correlation of r.m.c. and recognition with a-t out, $r_{23.1} = 0.05$ is consistent with the results found for the A-test, and serves to substantiate our earlier statement that the two tests are independent. On the other hand, ruling out the effect of r.m.c. ($r_{13} = 0.53$ and $r_{13.2} = 0.51$), has a negligible effect on the correlation of a-t and recognition. Evidently, ability to find the words containing *a* and *t* is much more important in determining one's score than the ability to mark the word when found. The multiple coefficient $R_{1(23)} = 0.55$ shows that there are other factors involved in the a-t test than those we have found; chance errors, interferences, slips of attention, etc. are probably some of these.

The score on the a-t test is influenced then in the ratio 3:1 by the tests of recognition and the marking movement. It is highly important that the word *score* be included in this statement. If we were able to say that Cancelling Ability is made up

in a certain way of the two factors, recognition and marking movement, we could then say that the score on a cancellation test should be influenced in exactly the same way by the two factors. It would, however, be disastrous if we argued in the reverse direction; from test to hypothetical ability.

II. THE COLOR NAMING TEST

One of the most striking features of the Color Naming test is the inhibition or interference which the subject experiences during its performance. He starts out confidently, naming the colors rapidly, but before he has gone very far he begins to hesitate and make mistakes; sometimes he will even stare at a color for several seconds before he can give its name. There are great individual differences in the amount of interference experienced; some subjects feel practically none, some experience a little, but by making an effort overcome it, while others are bothered throughout the entire performance. Interference was therefore selected as one of the factors to be studied for its effect on Color Naming.

Next the Color Naming test certainly involves perception of the colors—or probably recognition of the colors—and it seems reasonable to assume that individual differences in speed of recognition would influence performance on the test.

Finally, it is possible that an individual's performance might be influenced to some extent by the speed with which he is able to speak the names of the colors.

The three factors chosen for investigating color naming were, therefore,

1. Interference or inhibition.
2. Speed of recognition.
3. Speed of speech.

1. Interference

A test which would measure the interference factor directly would be very difficult, if not impossible, to devise. This factor could, however, presumably, be measured indirectly by a test identical with the Color Naming test in all particulars

except the interference-causing feature. This feature, according to Wordworth and Wells,⁸ is very likely the equal readiness of all five colors from immediately preceding use. A test involving 100 *different* colors, all just as familiar to the subject as the five used in the standard test, would probably be free from interference; hence a measure or index of the amount of interference experienced could be obtained by comparing the time for the standard test with the time for the 100-color test. As this sort of test was obviously impossible, it was thought that the next best thing would be to have the subject perform the standard Color Naming test in a modified manner, naming all of the squares of a given color in succession, instead of in a mixed order. That is, he would start with the first color, say Red, and sweep his eyes across each line from left to right, calling out Red each time he perceived a red square. Immediately upon saying the last Red, he would go back to the beginning and name all of the Blacks, followed in turn by the Yellows, Greens, Blues, using the first five colors of the top line as a guide to the order. The subjects experienced practically no interference in this test, which is referred to hereafter as the Color Finding test. The "interference index" of the subject was then taken to be the ratio of his time on the Color Naming test, standard method, to his time on the Color Finding test. Although this index is a ratio of two scores, and not a directly measured quantity, it is believed that it does give information concerning the amount of inhibition experienced by the subject, which cannot be obtained from either of the component scores alone. It is therefore to be expected that an individual showing signs of considerable interference in the Color Naming test would give a high interference index, and vice versa. Incidental observation showed that this was usually the case. It is evident, however, that the "interference index" may not depend entirely on the interference in color naming, but may also be influenced to some extent by ability in the particular "hunt and find" type of activity which is characteristic of the Color Finding test. The effect of this factor can, however, be cared for by statistical treatment.

⁸ Association Tests, Psych. Mono, 1911, no. 57.

2. Speed of recognition

"Speed of recognition" is a rather indefinite term, and it is probable that most individuals are not equally proficient in all types of "recognition activity." Recognition of colors is, however, the most important type from the standpoint of this study, and it was believed that the Color Finding test described above would serve as a good measure of such ability. For purposes of comparison it was decided to use also two "perception tests" of the conventional type; these were the a-t test and the word recognition test, previously described under cancellation.

3. Speed of speech

A test which would measure speed of speech without being influenced by speed of recognition might for example require the subject to recite, at maximum speed, some material which he had thoroughly memorized. Another type might require the subject to read material so easy and familiar that perception could take place much faster than speech could follow. Recitation of the alphabet is a good example of the first type. When this was tried, two defects at once became apparent. The performance took so little time that in order to obtain a reasonable range of scores it was necessary to use a multiple performance, which introduced a possibility of errors in counting the repetitions. Also, the subjects tended to mumble and slur the letters in spite of instructions to enunciate clearly. A second test considered was the reading of very simple prose. It was felt, however, that this would test a smooth, flowing type of enunciation rather than the more disconnected type used in saying the colors. The test finally decided upon was the reading of a series of two-digit numbers. The numbers were those of the Woodworth-Wells Constant Increment test, but instead of being arranged in columns, as in that test, they were typewritten in ten horizontal rows of ten numbers each following the spacing of the squares in the Color Naming test. All of the subjects reported that they could perceive the numbers much faster than they could say them; that is, the limiting factor was ability to enunciate rapidly rather than to perceive rapidly.

The tests were given as individual tests to 50 subjects, 25 men and 25 women. Nearly all of the subjects were graduate students in psychology. With the Color Naming, Color Finding, and Reading Numbers tests a double performance was taken in order to increase the range of scores. A stop watch, reading to $\frac{1}{5}$ second, was used for recording the time of performance.

The averages and the standard deviations of the various tests are given in table 3. All of the scores are in seconds except that for interference index, which is a pure number.

The most important correlations between the tests were calculated by the Product-Moment formula and are given in table 4. Since the Color Finding, Recognition, and a-t tests all

TABLE 3

TESTS	AVERAGES	SD
1. Color Naming	118 6	22 1
2. Interference Index	1 206	0 229
3. Color Finding	99 6	13 7
4. Reading Numbers . . .	105 9	15 5
5. Recognition of Words	298 0	90 6
6. a-t test	181 8	38 1

correlate to approximately the same extent with the Color Naming, it was decided to use only the Color Finding test as a measure of the speed of recognition. The partial correlation coefficients are to be found in table 5. These are arranged to show the effect upon the raw correlation between Color Naming and each of the other tests, of the elimination of the effect of the other tests taken separately and jointly.

The figures in the first column may be taken to mean that the correlation between Color Naming and Interference Index is very probably genuine and high; ruling out the Color Finding serves only to boost the correlation, from 0.71 to 0.85. Speed of speech (the Reading Numbers test) is practically without effect upon the correlation between Color Naming and Interference Index, since the correlation of no. 1 and no. 2 is not

effected when no. 4 is "partialled" out. Going to the second column, it is seen that the correlation between Color Naming

TABLE 4
Showing the intercorrelations between the different tests

	COLOR NAMING	INTERFERENCE INDEX	COLOR FINDING	READING NUMBERS	RECOGNITION OF WORDS	a-t TEST
	1	2	3	4	5	6
1. Color Naming	—	0 71	0 37	0 19	0 26	0 36
2. Interference Index	0 71	—	-0 18	-0 04		-0 04
3. Color Finding	0 37	-0 18	—	0 29		0 53
4. Reading Numbers	0 19	-0 04	0 29	—		
5. Recognition of Words	0 26				—	0 67
6. a-t test	0 36	-0 04	0 53		0 67	—

		PE			PE			PE
r ₁₂	0 71	0 05	r ₂₃	-0 18	0 09	r ₃₄	0 29	0 08
r ₁₃	0 37	0 08	r ₂₄	-0 04	0 10	r ₃₆	0 53	0 07
r ₁₄	0 19	0 09	r ₂₆	-0 04	0 10			
r ₁₅	0 26	0 09				r ₅₆	0 67	0 05
r ₁₆	0 36	0 08						

TABLE 5

1. Color Naming
2. Interference Index
3. Color Finding
4. Reading Numbers

r ₁₂	0 71	r ₁₃	0 37	r ₁₄	0 19
r _{12.3}	0 85	r _{13.2}	0 72	r _{14.2}	0 31
r _{12.4}	0 73	r _{13.4}	0 34	r _{14.3}	0 09
r _{12.34}	0 85	r _{13.24}	0 69	r _{14.23}	0 16

R₁ (234) : 0.88, PE: 0.02.

and Color Finding is nearly doubled when the Interference Index is ruled out. This may be taken as lending support to the

supposition previously made, that the main difference between the Color Finding and the Color Naming tests lies in the presence of the Interference factor in the second test, and its absence in the first. Again speed of speech, no. 4, has little effect. The coefficients in the third column are of little significance; they are all small and unreliable.

The multiple correlation coefficient $R_{1(234)}$, or the maximum correlation which we can obtain between test 1, Color Naming, and the combined effects of the other three tests taken together, is found to be 0.88. Some would interpret this coefficient as meaning that the chances are about 9 in 10 that any particular score on Color Naming depends on the scores in the three tests described. As the value which this multiple coefficient might be expected to have by a chance combination of our variables is 0.25 the obtained coefficient 0.88 is, therefore, probably reliable.

The regression equation, which expresses Color Naming (1) in terms of the other three variables (2), (3), (4), is given below. The coefficients have all been multiplied by a convenient constant in order to make their sum equal to 100, and thus show more clearly the relative weight of the three tests in determining the score on Color Naming. x_1 (Color Naming = $0.58x_2$ (Interference Index) + $0.36x_3$ (Color Finding) + $0.06x_4$ (Reading Numbers)).

It is clearly evident from this equation that the Interference Index is of greatest importance in the Color Naming score; speed of color recognition ranks next, while speed of speech evidently plays a minor part.

In view of the importance of interference, the cause of this state of things becomes a matter of much interest. Warner Brown⁹ has reported an experiment in which a comparison was made of the time required to name a series of colors (not those of the standard test) and the time required to read the same color names when typewritten on a sheet. He found that the names could be read in approximately half the time required to name the colors. Both performances were improved by practise, the

⁹ Psychological Review, 1915, vol. 22, p 45.

per cent of improvement being about the same in each. This experiment was tried, in an abbreviated form, in the present study—the standard Color Naming test being used,—and the additional fact was noted that practically no interference occurred when reading the color names from the typewritten sheet, in spite of the fact that the factors of recency and frequency affect all of the color names to the same extent as in the Color Naming test. Hence it seems reasonable to say that interferences which arise in naming colors are due not so much to an equal readiness of the color names as to an equal readiness of the color cognitive processes. Another factor present in interference is very probably the present strength of the associations between colors and their names, already determined by past use. Practise usually reduces interference in this test, though oftentimes the improvement is slight.