ON ETHNOLOGICAL TESTS OF SENSATION AND PERCEPTION WITH SPECIAL REFERENCE TO TESTS OF COLOR VISION AND TACTILE DISCRIMINATION DESCRIBED IN THE REPORTS OF THE CAMBRIDGE ANTHROPOLOGICAL EXPEDITION TO TORRES STRAITS.

By E. B. TITCHENER.

(*Read April 7, 1916.*)

The work of the Cambridge Anthropological Expedition to Torres Straits was planned to include a psychological study of the native population. The "equipment of a small psychological laboratory" was taken out to Murray Island and set up in the disused missionary house; and the tests made, there and elsewhere, were conducted by Messrs. Rivers, Myers, McDougall and Seligmann. The leader of the expedition was already known to the natives, whose goodwill was thus assured.[1] On the whole, I suppose that field-work in psychology has never been done under better conditions: the apparatus had been considered and chosen beforehand, the experimenters were competent, the natives were amenable. Yet I think that no home-staying experimentalist can read the psychological part of the Report with satisfaction. At any rate, the impression left by my own repeated reading is that the tests were inadequate to their purpose.

In the present paper I criticize, as severely as I can, two samples of this field-work. I have chosen tests whose results have been widely quoted, and I have chosen them for their strength rather than for their weakness. It is true that the Report is fifteen years old; and it is true that, during the past fifteen years, experimental psychology has made great methodical advances, social anthropology has grown

[1] See Reports of the Cambridge Anthropological Expedition to Torres Straits, II., 1901, v. f., 1 ff. This volume is referred to, in later notes, by the letter R.

apace, and psychotechnics has come to the forefront of discussion. It does not follow, however, that my criticism is out of date. For field-psychology is still in its first beginnings. Something, no doubt, has been learned; and an expedition that we should organize today would hardly be content to repeat the programme laid out for Murray Island. The labor and ingenuity spent upon mental tests have not been simply thrown away. I doubt, nevertheless, whether the new expedition could be certain of improving, in any material way, upon the work of its predecessor; and if this doubt is well founded —nay, if there is any doubt in the matter at all—then a detailed criticism of the older tests, supported as mine is by collateral experiment, ought to be of service.

It is, perhaps, not out of place to add that I am criticizing experiments with which I have the greatest sympathy. My own lifelong interest in psychology came by way of anthropology; and if I chose the laboratory in preference to the field, that was only because I was convinced that the first necessity of experimental psychology, as a science, was the standardizing of instruments and procedures. I realize that time presses; the primitive stocks are fast changing or disappearing. I realize, too, the difficulties of work in the field: the unaccustomed mode of life, the lowering of health, the indifference and trickiness of the native, the frequent breakage or failure of apparatus and the impossibility of replacement or of proper repair. My criticisms are thus offered in the friendliest spirit; and my aim is to lead up to positive suggestion rather than to conclude with a merely negative result.

One further point, and these introductory remarks may be ended. The critic of tests performed in the field is at a serious though unavoidable disadvantage, in that he has nothing more to go upon than what the field-workers include in their report. When an experimental study is published by a laboratory, it is open to any other laboratory to repeat the work with observers of the same order of intelligence and training, by the same methods, with similar instruments. Tests made in Torres Straits cannot be thus controlled; the printed pages are all that we have. I have tried, nevertheless, to make parallel and illustrative experiments of my own. If the reader

finds them relevant to the issues discussed, the credit is largely due to the fullness with which the authors of the Report have written. They remind us, time and again, that human nature is much the same the world over. The old lady on Murray Island who, "by associations of some kind," gave her own and her friends' names when she was asked to name a set of colored papers, have we not met her—with all respect be it spoken—in our summer sessions? The man who was asked to arrange the colors in the order of his liking, and who handed them back in almost exactly the order in which they had been presented by the experimenter in an earlier test, is a fairly familiar figure in our laboratories. If, then, I presently compare the observations of my colleagues with those of a Papuan chief, I am not necessarily falling into absurdity.

I. The Delicacy of Tactile Discrimination.

I begin with McDougall's experiment upon the limen of dual impression upon the skin. His conclusion is as follows: "These figures indicate that in the skin areas tested the Murray Islanders have a threshold of tactile discrimination of which the value, in terms of distance of two points touched, is just about one half that of Englishmen, or we may say in other words, that their power of tactile discrimination is about double that of Englishmen. And . . . we may assume that this result is true for all or most parts of the skin."[2] I shall try to show that the conclusion does not follow from the experimental data.

McDougall used "a small pair of carpenter's dividers with blunt metal points."

> "These two points were applied to the skin simultaneously with light pressure lasting about one second. The subject was told to keep his eyes shut, and the area of the skin operated on was further guarded from his view. . . . He was told to say 'one' or 'two' according as he judged that one or both points touched his skin. . . . The threshold that I sought was . . . not that distance at which two points can be distinctly felt, but a slightly lower one, that distance at which they yield a sensation perceptibly different from that yielded by a single point.
>
> "One point was applied in every experiment about as frequently as the two points. . . . In order to keep the attention and interest of the subject as keen as possible, I found it necessary to tell him after each answer whether

[2] R, 192, 195.

he was right or wrong, for in default of this precaution many of the subjects soon contented themselves with random answers."[3]

It is a minor but still a relevant detail of criticism that the metal points should have been replaced by hard wood or hard rubber. The metal points appeal to the temperature senses as well as to the sense of pressure. I pass, however, to more important things. It will be noticed that although, on general principles, "the procedure should be, as far as practicable, without knowledge on the part of the subject,"[4] the method is in fact full of suggestion. The subject knows what is being done to him; he is to judge whether "one or two *points* touched his skin." The stimulus-error, with all that it brings in its train, is thus not only admitted but even welcomed. Secondly, the one-point stimuli are "about as frequent" as the two-point. The subject thus has recurring opportunity to check or control his dual judgments. And since he is told "after each answer whether he was right or wrong," the control is continually renewed and the difference between single and dual stimulation is continually emphasized.

The results obtained may be illustrated from the forearm-limens;[5]

MURRAY ISLANDERS.

Average of 50 men ..... 19.8 mm. (median, 20; extremes, 40 and 2)
Average of 25 boys ..... 14.0 mm. (median, 15; extremes, 25 and 2)

ENGLISHMEN.

Average of 23 men ..... 44.6 mm. (median, 40; extremes, 90 and 10)

It is clear that the Englishmen have the higher average limen. It is also clear, however, that the range of the results is excessive. In the case of an elementary perceptive discrimination, we do not expect values that range between 40 and 2, or between 90 and 10 units of measurement. The results suggest, then, that different subjects may have been doing different things. Fortunately we have collateral evidence which not only confirms the suggestion but also indicates what the different things aimed at and done may have been.

The limen sought by McDougall was, it will be remembered, "not

[3] R, 190 f.
[4] R, 189.
[5] R, 191 f.

that distance at which two points can be distinctly felt, but . . . that distance at which they yield a sensation perceptibly different from that yielded by a single point." It has long been known, however, that between the limits "point" and "two points" there are a number of distinguishable perceptive forms. Gates, working recently in my laboratory, distinguished "circle," "line" and "dumb-bell"; but there were slight differences within these categories; and there is no doubt that more such forms could be made out.[6] Suppose, then, that an observer is set or disposed for "dumb-bell"; his limen will be relatively high. Suppose, on the other hand, that he is set for "circle"; his limen will be relatively low. In either event he will be reporting "a sensation perceptibly different from that yielded by a single point" and lying on the hither side of "two points distinctly felt"; but the limens will be the limens of two different perceptive forms, and therefore will not be comparable the one with the other. Here, I take it, is the principal key to the wide range of McDougall's results. Some of his observers judged "two" as soon as the impression of the points differed in the slightest degree from "one"; others judged "two" only when the points were on the brink of falling apart or had actually done so. If this inference is correct, the comparison of the Murray Islanders with the Englishmen is null and void.

I made, to test it, a rather venturesome experiment. With McDougall's instrument "it was not possible satisfactorily to apply the two points at an interval less than 2 mm."[7] The æsthesiometers regularly used in my laboratory have hard-rubber points (diameter 1 mm.) which may be directly apposed. With our instruments, therefore, it is possible to employ a dual stimulation with a separation of 0 mm.; in other words it is possible to compare the impression of a single point with that of two apposed points. It occurred to me that I might be able, under the suggestive influences of McDougall's method, to discriminate these two impressions. I asked

[6] E. J. Gates, "The Determination of the Limens of Single and Dual Impression by the Method of Constant Stimuli," *Amer. Journ. Psych.*, XXVI., 1915, 152 ff. M. Foucault ("L'Illusion paradoxale et le seuil de Weber," 1910, 124 f.) distinguishes six intermediate perceptive forms.

[7] R, 191.

the first available experimenter to make up a haphazard series of 32 stimuli, 16 dual and 16 single, and to perform the test upon my forearm in McDougall's way. The experimenter was wholly unpractised, so that I worked under unfavorable conditions: the time of impression varied, its intensity varied, the instrument was set down sometimes slowly and sometimes with impact, the ready-signal was often omitted, and so forth. As a preliminary I was given four stimulations, two single and two dual; by these few experiences I tried to fix in mind the difference of the two perceptive forms. Then came the series; and it turned out that I was able to judge the two-point impression correctly in 11 out of its 16 appearances. That is a correctness of almost 70 per cent.

The result was encouraging—so much so that I determined to repeat the trial under better conditions. A few days later, then, I worked through the same series with change of order (not that I should have recognized the original order!) and with proper regulation of times and intensities and warning-signals. I now judged the two-point impression correctly in 15 out of 16 cases. Since McDougall's liminal value requires only 13 correct judgments,[8] my present limen of dual impression, calculated in his way, is something less than zero. I have at least equalled the performance of Meiti and Tapau and Biskak.[9]

I give the series in full, since they are instructive to those familiar with the æsthesiometric experiment.

TRIAL I. PRELIMINARY STIMULATIONS, 4.

Stim.	Jdgt.	Stim.	Jdgt.	Stim.	Jdgt.	Stim.	Jdgt.
2	2	1	1	2	1	1	2
1	1	2	2	1	2	2	2
1	2	1	1	2	2	2	1
2	2	2	1	2	2	2	1
2	2	1	1	1	1	2	1
2	2	1	2	1	2	1	1
1	2	2	2	2	2	2	2
1	2	1	1	1	2	1	1

Summarizing we have:

2 judged as 2.....11 times
2 judged as 1..... 5 times
1 judged as 1.....8 times
1 judged as 2.....8 times

[8] R, 190.

[9] R, 203 f.

TRIAL II. PRELIMINARY STIMULATIONS, 6.

Stim.	Jdgt.	Stim.	Jdgt.	Stim.	Jdgt.	Stim.	Jdgt.
2	2	2	2	1	1	1	1
2	2	2	2	2	2	2	2
1	1	2	1	2	2	1	1
1	1	1	1	2	2	1	1
2	2	2	2	1	1	2	2
1	1	1	2	1	1	1	1
1	2	2	2	1	1	2	2
2	2	1	1	2	2	1	1

Summarizing we have:

2 judged as 2.....15 times	1 judged as 1.....14 times
2 judged as 1..... 1 time	1 judged as 2..... 2 times

I need not say that there was no hint of duality in the perceptive form which underlay the judgment "two"; I yielded, without compunction, to the stimulus-error. Moreover, since I was interpreting and not describing, I can say very little of the perceptive forms themselves. The one-point impressions were solidly homogeneous, and often had a trace of sting in them; the two-point impressions were duller and coarser, and at times gave a hint of something like granulation. I did not as a rule think of the two-point impressions as larger, more diffuse than the others, though I recall that the one-points with sting were rather definitely small.

The range of McDougall's results might thus be accounted for. We have still to discuss the fact that the average limen of the Murray Islanders is smaller than that of the Englishmen.

Rivers, writing of these same Murray Islanders, lays great stress upon "the over-development of the sensory side of mental life"[10] in the savage. Myers, dealing with a like theme, points out the interpretative character of their sensory interests. "Probably the mode of life led by primitive peoples and their general mental status combine to make them more aware of and attentive to the majority of external stimuli than we ourselves are. . . . A faint odor may be simultaneously perceptible to the civilized and to the uncivilized individual. To the latter it will be full of meaning and so will at once engage his attention; for the opposite reason it is apt to escape the notice of the former." Let us look at McDougall's results in the light of these general statements. The task set to the Murray

[10] R, 44 f., 64.
[11] R, 181 f.

Islanders was a task of sensory interpretation; they were not to get a distinct perception of two separate points, but only "a sensation perceptibly different from that yielded by a single point." I have shown that they might have chosen any one of several perceptive forms, and I have argued that different subjects did, as a matter of fact, read different meanings into the instruction given. Their general tendency toward a low limen I ascribe to that sensory interest, that ingrained habit of interpretation of external stimuli, which Rivers and Myers attest. They were to find a "sensation perceptibly different" from another, and they carried the difference—some of them—as low in the scale of separation as 2 mm.

The Englishmen also read different meanings into the instruction given. They were, however, as a group, less interested in the minutiæ of external stimuli than the savages; their power of sensory interpretation was less; they paid, we may suppose, more attention to the particular instrument used, and to its probable effect upon the skin. They looked for a sensibly dual impression; and though they did not all confine the judgment "two" to cases in which the stimuli fell apart for perception, yet they naturally tended in that direction. It is very significant that "among the Englishmen were five of the educated class, and these gave a rather higher threshold than the rest, who were all of the lower class."[12] The farther we go from the savage's sensory interest and power of sensory interpretation, the larger do our limens become! Not, of course, that the educated Englishmen were necessarily less sensitive than the rest, but simply that they took "duality of impression" in a stricter sense.

I conclude, then, that we have no right to say "of these Murray men that their sense of touch is twice as delicate as that of Englishmen." That may or may not be the case; but, in any event, the conclusion does not follow from McDougall's experiments. So far as relative "delicacy of tactile discrimination" is concerned, these experiments leave us precisely where we were.

I have tried to find a reasonable explanation of McDougall's results taken at their face value. I have now to express a complete distrust of his formal procedure. The "test" of the Murray Is-

[12] R, 192.

landers—McDougall describes it as "a combination of the method of minimal changes with the method of right and wrong answers"[13]—bears all the earmarks of an incomplete psychophysical method. My criticism, in brief, is that fragments or sections of established methods cannot, without very rigorous trial and definite proof of reliability, be combined to form a new method.

Subjects.	Number of Preliminary Trials.	Separation of Points (Mm.) in Test-series.	Correct Dual Judgments.	Test-limen (Mm.).
1	29	22	5	
		24	4	25?
2	21	40	5	
		48	3	
		60	9	
		56	10	
		52	10	
		48	7	
		50	4	
		52	8	52
3	16	30	4	
		40	7	
		50	9	
		45	8	45
4	12	20	6	
		30 (discarded)	8	
		30	7	
		25	8	25
5	6	30	6	
		40	8	40
6	12	30	8	
		25	7	
		35	7	
		40	10	30
7	56	—	—	?
8	11	30	2	
		40	4	
		50	8	50
9	23	36	9	
		34	8	34
10	21	49	9	48
11	10	35 (discarded)	8	
		35 (discarded)	3	
		40	4	
		50	8	
		60	8	
		80	10	
		70	9	50

The aim of McDougall's procedure, on the quantitative side, was to get eight correct dual judgments out of a possible ten. I have before me the data of a careful performance of Whipple's test (1910)

[13] R, 191.

on the "discrimination of dual cutaneous impressions."[14] They have not been published; but by the kindness of my colleague, Professor Kennedy-Fraser, I am allowed to quote from them in this place. The aim of this test also is to "seek a distance such that about 8 correct judgments in 10 are made, *i. e.*, such that double contact is reported as double in 8 of 10 trials." The condensed results from eleven subjects, obtained after the prescribed preliminary practice, are given in the table on page 212.

The subjects of this test received no instruction regarding perceptive form, and the range of the limens (25 to 52 mm.) may be partly due to that omission. If we consider the results solely from the quantitative side, we may draw the following inferences:

1. It is dangerous to repeat a test-series (Subject 2); to repeat is to give the method an opportunity to contradict itself. Conversely, the result comes out most neatly with the use of few test-series and fairly wide steps.

2. Inversions are not uncommon. Subjects 1, 2, 4 and 6 furnish instances in which a lesser separation of the compass-points yields a greater number of correct judgments.

3. The weight to be attached to the preliminary trials is left to the discretion of the experimenter (Subjects 1, 10). The test-procedure thus contains an element of uncertainty.

4. The test as prescribed may break down altogether (Subject 7). The failure in this particular instance is due, not to the intercurrence of paradoxical judgments,[15] but to irregularity of the normal judgments; the subject evidently changed his standard as the trials went on.

[14] G. M. Whipple, "Manual of Mental and Physical Tests," 1910, 207 ff.

[15] This possibility was foreseen by Whipple (op. cit., 211). The discarding of series by the experimenter (Subjects 4 and 11) was done for cause; but it too introduces an element of uncertainty.

To meet the variety of perceptive form, Whipple recommends in his second edition (I., 1914, 247 ff.) a method of contrast. "The threshold may be taken as the distance at which two errors are first made with the ten double points, unless subsequent better records with lesser separations show that these errors were due to a temporary lapse of attention." I am afraid that repetition would still be dangerous; not only attention but also basis of judgment might shift from series to series.

All of these defects are characteristic of an imperfect, incomplete, partial method. It may be said, I think, without qualification that numerical results of the same kind as those obtained by McDougall and by the performers of Whipple's 1910-test may be got by fractionating the results of the method of constant stimuli. Here is an illustration from my own laboratory.

DETERMINATION OF THE TWO-POINT LIMEN BY THE METHOD OF CONSTANT STIMULI.

80 series, 400 observations. Stimuli 0, 6, 12, 18, 24 mm.

(A) Results Arranged in Groups of 10 Series.

Stimuli.	Percentage of Dual Judgments.								Percentage Limits.	Difference.
0	30	10	20	10	0	0	20	20	0–30	30
6	30	80	20	30	30	50	40	20	20–50	30
12	50	*50*	70	60	40	70	80	50	40–80	40
18	*30*	90	90	70	80	90	90	50	50–90	40
24	100	100	*80*	90	100	100	100	80	80–100	20 Av. 32

(B) Results Arranged in Groups of 20 Series.

0	20	15	0	20	0–20	20
6	55	25	40	30	25–55	30
12	*50*	65	55	65	50–65	15
18	60	80	85	70	60–85	25
24	100	85	100	90	85–100	15 Av. 21

(C) Results Arranged in Groups of 40 Series.

0	17	10	10–17	7
6	40	35	35–40	5
12	57	60	57–60	3
18	70	77	70–77	7
24	92	95	92–95	3 Av. 5

(D) Result of 80 Series.

0	14
6	37
12	59
18	74
24	94

The cases of inversion (4) are printed in italics.

We notice that, as the size of the group increases, the range decreases and the distribution of judgments grows increasingly constant; that is what we should expect. We notice also, however,

that the 8 series of (*A*) are of the order of test-series. If we ascribe them hypothetically to 8 different subjects, we may fix the 8-in-10 limen of the last five as 21, 18, 15, 12 and 24 mm. respectively (range 12–24); no limen can be calculated for the first three. Even, then, if the basis of the test is enlarged, and we take 50 judgments from every one of our supposed 8 subjects; and even if the basis is regularized, and we test all the subjects by the same stimuli; even so, the limen shows a wide range and proves in certain cases to be indeterminable. It is plainly not enough, in this test, to secure eight out of ten right answers by a rapid procedure applied "under the same conditions" to a number of unpractised subjects. For the conditions during such a test are only by chance the same for different persons; the probability is that they are diverse; and the repetition of the test on the same subject may change the original finding. From this point of view, also, McDougall's results are open to grave suspicion.

It is fatally easy for the field-worker and the laboratory-worker to misunderstand each other. So I had better say at once, and emphatically, that I do not want to see the refinements of the home-laboratory carried into the field. When Galton suggested that "in testing the delicacy of the various senses . . . we should do wrong if we pursued the strict methods appropriate to psychophysical investigations,"[16] I take him to have been heartily in the right. I criticize McDougall's combination-method on the formal ground that it is not a method, whether fine or coarse; it is, so far as I can see, essentially the rough equivalent of a section or fragment of the method of constant stimuli; and as a mere fragment of a method it can lead us nowhere. We are fortunate in having the method of constant stimuli in the background, to give us numerical norms; but we cannot use a piece of a method as if it were the whole.

II. Color Vision

I have chosen, as a second example, Rivers's work upon color vision. This is a very painstaking bit of research, and its conclusion

[16] F. Galton, "On the Anthropometrical Laboratory," etc., *Journ. Anthrop. Inst.*, February, 1885, p. 4 of offprint.

is stated with becoming caution. The conclusion is "that the color vision of the Papuan is characterized by a certain degree of insensitiveness to blue (and probably green) as compared with that of Europeans."[17] I do not think that the observations made warrant the inference drawn from them; and I therefore take up the cudgels on behalf of the Papuan. To avoid overmuch detail I confine myself strictly to Rivers' report; I deal with the major question of blue, and leave green out of account; and I consider only the observations taken on Murray Island.

The quantitative work was done with Lovibond's tintometer. "The essential part of the instrument consists of three series of colored glasses, red, yellow and blue, very delicately graded so that each forms a series by means of which one passes from a color so faint as to be indistinguishable from colorless glass up to a glass of a high degree of saturation."[18] The results, stated in terms of the unit of the instrument, are as follows:[19]

	R.	Y.	B.
17 Murray Islanders.....	17.6± 7.66	26.5±9.71	60.0±16.5
18 Englishmen.....	31.7±22.5	20.5±8.11	36.4±15.13

Rivers has excluded from his Murray averages the results found with one boy who probably suffered from "a slight degree of photophobia."[20] He includes in the English averages the results of four observers who were "exceptionally insensitive to red" and of two who were "insensitive to blue."[21] If we ourselves exclude these cases, we get the revised figures:

Englishmen ..... (14) 18.0 ± 5.5 (18) 20.5 ± 8.11 (16) 34.1 ± 13.66

The exclusion of the defective subjects brings the averages into closer accordance with Rivers' theory; and for that very reason he

[17] R, 94.

[18] R, 70 f. Various forms and uses of the tintometer are discussed by J. W. Lovibond, "Measurement of Light and Color Sensations," 1893. I am not familiar with the instrument, and Lovibond's description is, unfortunately, not always full. The yellow glasses "have a distinctly greenish tinge" and the red glasses are "distinctly bluish" (R, 71, 74 f.).

[19] R, 72.

[20] R, 73.

[21] R, 73.

may have thought it fairer to leave them in. The Murray Islanders and the Englishmen now agree as regards red and yellow; both the liminal values and their mean variations are of the same order; but they differ markedly as regards blue.

On this point I have two remarks to offer. (1) In experiments which I have made with Hering's colored glasses, my observers have reported that it is difficult to determine the chromatic limen of R and B (I have had no such report for Y), because the faintly colored glass shows, at the moment of exposure, a flush or wave of color which immediately disappears through adaptation. It is possible that the four English observers who were "exceptionally insensitive to red" failed to notice this momentary flush in the neighborhood of the limen, while the other Englishmen and the Murray Islanders (whom we know to be keenly interested in red) noticed it and placed the limen accordingly. The range of results for the Murray Islanders is 40 to 5; the range for the 14 Englishmen is 40 to 10: a remarkably close agreement. In the case of blue, it is possible that the 16 English observers noticed the flush and reported it as color, while the Murray Islanders (who are definitely uninterested in blue, since they have in their surroundings no blue object of reverence or utility) disregarded it. The range for the Murray Islanders is 100 to 30, and that for the 16 Englishmen 80 to 15; the ranges are thus of the same order of magnitude, are very large, and show a wide overlap. My distinction, therefore, must not be pressed; the possibility is rather that the Murray men *tended* to overlook the flush, the Englishmen to report it as blue. In the case of yellow (for which, as it happens, I have no report of the flush) the ranges are 50 to 10 for the Murray Islanders and 60 to 4 for the 18 Englishmen. The agreement is less striking than for red, but is still fairly close.

There is, then, a possibility that the difference in the case of blue may be due, in part at least, to regard or disregard of the supraliminal flush of color to which I have called attention. I have myself had student-observers who disregarded the flush, both of R and of B, because they thought the field should look colored during the

whole period of observation.[22] There is, however, a second possibility, to which I now turn.

(2) Geissler, working upon the chromatic limen with colored papers, found that increase of the general illumination markedly decreased the limen of B, while it had no such marked effect upon the limens of R and Y. I am not quite sure of the figures which should be quoted for comparison; but it appears that, for two observers, increased illumination (natural daylight "several times as bright" as the artificial daylight otherwise used) lowered the limen of B, in degrees of the color disc,

from 4.12 or 4.64 to 1.83,

and

from 5.28 or 5.80 to 2.96.

Whichever pair of larger figures we take, as a basis of comparison, "the striking fact is that the limen for blue in natural light was lowered."[23] I can parallel this result by observations on colored glasses. In 1890 I visited with Mr. (later Sir Francis) Galton the anthropometrical laboratory which he had equipped at South Kensington; and among other things I worked awhile with Galton's "instrument for testing the perception of differences of tint."[24] My notes tell me that the laboratory attendant and I "made a great hash" of our trials with blue, and that Galton remarked on the gloom of the laboratory as unfavorable to color work. I cannot remember that we worked

[22] The tintometer limens for the four red-insensitives are 50, 70, 80, 120. Rivers thinks that the subjects with the two highest limens "had probably some degree of weakness of the red-green sense" (R, 73), and they may belong to Nagel's anomalous or atypical trichromates (W. Nagel, "Ueber typische und atypische Farbensinnsstörungen," *Zts. f. Sinnesphysiol.*, XLIII., 1908, 299 ff.). The two blue-insensitives have limens of 50 and 60; other subjects, not characterized as insensitive to blue, give 45, 60, 60, 80. These high limens may be due simply to lack of practice; experience in the laboratory shows that undergraduates are far more likely to give the name "gray" to a slightly bluish gray than to a red-gray or a yellow-gray of the same chroma; blue is undoubtedly *like* gray. These considerations modify my argument in detail; they do not affect its principle.

[23] L. R. Geissler, "Experiments on Color Saturation," *Amer. Journ. Psych.*, XXIV., 1913, 177 f.

[24] F. Galton, "Exhibition of Instruments," etc., *Journ. Anthrop. Inst.*, August, 1889, 27 f.

with glasses of other colors; perhaps we did; the thing that struck me, at any rate, was the difficulty of distinguishing blues in a darkish room. Recent work with Hering's colored glasses shows that a shift of the instrument employed, on a gray snowy day, from the window to the middle of a large gray-painted room produces the following changes in liminal values (three observers; method of limits; conventional units):

Obs.	Light.			Dark.		
	B.	Y.	R.	B.	Y.	R.
I....	36.0	48.5	59.75	68.5	54.5	83.5
II....	26.75	32.0	45.75	41.25	42.5	51.0
III....	42.2	47.0	43.0	59.2	68.6	78.4

I give no further details, since I attach little importance to the numerical values; however carefully the work is done, it is full of errors. I notice only that B is the one color that suffers consistently in the dark, and that it suffers on the average much more than Y and about as much as R. The figures are:

Increase of Limen in Dark.

Obs.	B.	Y.	R.
I.....	32.5	6.0	23.75
II.....	14.5	10.0	5.25
III.....	19.5	21.6	35.4
Average.....	22.1	12.5	21.5

If, now, the tintometer in Torres Straits was set up in a room of "the disused missionary house," and in England in a well-lighted laboratory room—we are not informed as to the English conditions—then we may be pretty sure that the Murray Islanders were at a disadvantage on the score of blue. That conclusion follows both from Geissler's results and from my own. Such a disadvantage, whether acting alone or combined with a tendency to disregard the supraliminal flush, might account for the difference in the average limens of Murray men and Englishmen. It may be that the Murray Island limen for Y (slightly higher than the English, despite its somewhat

closer range) is due to the same difference of general illumination. We are then, it is true, left at loose ends in the matter of R. Here, if comparson is at all permissible, Geissler's results are in agreement with Rivers'; my own are in disagreement. But we do not know whether the two R glasses were alike. Above all, I cannot tell whether the tintometer-series of colored glasses are so combined with grays as to show the same "brightness" throughout, or whether during an experiment they brighten with decrease of chroma; I cannot tell, either, whether the colorless field, with which the colored field is compared, itself varies with the "brightness" of the colored glasses or remains the same for all.[25] In my own experiments "brightness" varied with chroma, but the colored and colorless fields were of approximately the same "brightness" in every observation. Here are possible differences of procedure that might affect the results.

I said just now that we are not informed of the conditions under which the English tests were made. It is worth noting, however, that Rivers found excessively high limens "in testing Europeans in too strong a light."[26] Since too strong a light (daylight) could hardly be obtained save in a specially lighted laboratory room, it seems probable that my supposition as regards the placing of the tintometer is correct.

All of these criticisms are offered, of course, with the greatest reserve; Rivers may be able to meet them point for point. Taking his report as it stands, I think they are sufficient to cast serious doubt upon the conclusion which he draws from the tintometer-results. The report, however, is incomplete; we lack details of experimental procedure and conditions; and the English observers, who

[25] Lovibond speaks of neutral-tint glasses, standards, units; he also has diffusive glasses, thin slips of ground glass (*op. cit.*, 21, 31 ff., 48, 101 f.). It would therefore seem possible to keep hue and "brightness" of the colored glasses constant while chroma varied, and to equalize the "brightness" of the colored and colorless fields. Rivers does not tell us whether this was done. In setting up the instrument for differential determinations, he found that the "difference in brightness" of the glasses rendered the results inconclusive (R, 74). If brightness affected these results, must it not have affected the others? and if it was compensated in the other experiments, might it not have been compensated in these?

[26] R, 73.

might have been thoroughly tested, are characterized only in round terms. The instrument employed may have been the best available; it embodies a vast deal of patient labor; but it has not made its way into our laboratories, and it has not (so far as I know) received a thorough trial in any laboratory. We have no proof that it is adequate to a comparative testing of color vision.

I pass to the vexed question of color nomenclature. On Murray Island "there was great definiteness and unanimity in the nomenclature for red, rather less so for orange and yellow, less so for green, and very great indefiniteness for blue and violet."[27] "In Murray Island there is no proper native term used for blue. Some of the natives, especially the older men, use *golegole,* which means black [cuttle fish], but the great majority use a term [*bulubulu*] borrowed from English and modified so as to resemble the other members of their color vocabulary [the color names are formed by reduplication from the names of various natural objects]. Another word, *suseri-suseri* [rainbow], is used occasionally for blue and also for green, and in the absence of the borrowed word this might have been used more often."[28]

The absence of a word for blue, if the fact stood alone, is no argument against sensitivity to blue. For the savage names only what interests him, and we have seen that his interest is directed upon the interpretation of sensory stimuli. But there is in Murray Island no such sensory stimulus, no object of daily use or interest—no pigment, for instance—of a blue color. I think that Rivers would not dispute this position, if the fact stood alone. It is only because of other facts, and because the character and distribution of the other color-terms in the vocabulary arouse suspicion, that he would argue—always tentatively—from "defective color language" to "defective color sense." We must therefore get a conspectus of the vocabulary at large.

I have arranged the data for Murray Island in the form ot a table.[29] Above stand the names of the objects from which the color

[27] R, 54 f.
[28] R, 66 f.
[29] From R, 53 ff.

names are derived. The colors are indicated by their initials (R, red; I, indigo; etc.). A capital letter means that the color was designated by the name of the object at the head of its column in the majority of cases, or (for Y, I and V) by considerable groups of the subjects; a small letter means that the color was so designated less often, sometimes by a single subject. Rivers has, unfortunately, not stated these results in numerical form. A + or — sign after a letter means that the color name was modified by "big" or "little"; thus, o— means that orange was called "little blood-blood" by a few subjects; p+ means that purple was called "big blood-blood" by (as it happens) one man.

Blood.	Red Ochre.	Turmeric.	Yellow Ochre.	*Zom* Flower.	Bright.	Turmeric + Cuttle-fish.	Sea + Pus.	Turtle Gall-bladder.	Newborn Child.	Pus.	Leaf.	Rainbow.	English "Blue."	Cuttle-fish.	Sea + Cuttle-fish.	Dull.	Blue + Cuttle-fish.	Rainbow + Cuttle-fish.	Dark.	Secretion of a Mollusc.	A Mollusc + Blood.	*Somer* Fruit + Blood.
R, R+, r—																						
o, o—	o	O	o						y													
		Y	Y	y	y																	
						yg	yg	YG	yg	yg	yg	yg										
								G	g	g	g	g										
			g—					bg	bg	bg	bg	BG										
			bg—					b	b	b	b	b	B	b		b						
												i	I	I	i							
v—												v		V		v	V	v	v	v		
P—, p+																				p	p	p

I have included in the table all but two of the color names employed. BG was called "dirty yellow-ochre" as well as "little yellow-ochre"; I have not thought it worth while to make a separate column for this compound form. V was called blood-white, if *kake-kakek* means white; I return to this case presently.

Let us now look at the distribution of the color-names, taking the column of the table as unit. Then

R has 1 name
O has 4 names
Y has 5
YG has 7
G has 6
BG has 6
B has 8 (or 7 if "blue" is excluded)
I has 4 (or 3 if "blue" is excluded)
V has 8 (or 9 if "blood-white" is included)
P has 4

Rivers' phrases are:

"great definiteness and unanimity for R"	(1 name)
"rather less so for O and Y"	(4 and 5 names)
"less so for G"	(6 names)
"very great indefiniteness for B and V"	(8 and 8 or 9 names)

The rise from 5 to 8 thus changes a "rather less great definiteness and unanimity" to "very great indefiniteness." Yet the B names group, clearly enough, with the YG, G and BG names. We need, as I said above, the actual numbers of subjects who used particular terms; then we could weight the data of the table.

Rivers' further findings[30] are as follows:

White paper is called	*KAKEKAKEK, zazerzazer,* lime.
Deep black paper	CUTTLE, BIG CUTTLE, *kukikuki.*
Dull black paper	CUTTLE, LITTLE CUTTLE, ashes-cuttle, ashes, gray clay, New Guinea earth, dirty.
Gray (162° W + 198° Bk)[31]	*kakekakek,* little *kakekakek,* ashes-*kakekakek,* ashes-*zazerzazer.*
Gray (49° W + 311° Bk)[31]	ashes, gray clay, cuttle.
Holmgren's pale green test wool	*kakekakek,* ashes, little turmeric, *zazerzazer, zom*-flower.
Violet test wool	ashes, *kakekakek, zazerzazer,* little blood, cuttle.
Brown wools are called	little blood, little turmeric, ashes, cuttle, dull, "according to their prevailing tone and shade."
Slightly saturated colors and dull black are called	*dudu.*
Color of native skin	cuttle.
Dark colors	*kupekupe.*
Dark	*kupekupe, kukikuki, kakerikakeri.*
Bright, glittering	*zoromzorom.*

The cruces of the vocabulary are *kakekakek* and cuttle. Rivers translates *kakekakek* roundly by white; he does not know the derivation of the word.[32] We find it applied to white paper, light gray paper, Holmgren's apple-green test wool, the violet test wool, and

[30] R, 54.

[31] We are not told whether these grays were mixed from the dull black or the deep black paper.

[32] R, 49, 56.

(in the form blood-*kakekakek*) to violet paper.[33] I suggest that it is the equivalent, on the side of light, to *dudu* on the side of dark; that is, that it means "light and inconspicuously colored." If I were guessing at its derivation I should expect to find that the object *kakek* is something of use and interest which is not easily found by reason of its lightish, faded or washed-out color. The term "blood-white" for violet would then be practically the same as the "little-blood" already entered in my table.

Cuttle is applied to black paper, dark gray paper, brown wool, native skin, blue paper, indigo paper, and violet paper and test wool. Rivers translates the word by black; he notes that the older men use it for blue, and remarks on the fact that "intelligent natives should regard it as perfectly natural to apply the same name to the brilliant blue of sky and sea which they give to the deepest black."[34] The thing is strange to us: but we must consider the native. In the first place, we do not know whether the derivation of the word is present to the native's mind. Rivers thinks that "newborn child color" may simply mean "light";[35] I suppose that the specific reference to the skin-color of the newly born has dropped away. Cuttle-color, in the same way—especially since the word is an old one—may have a general and not a specific meaning. What, then, in the second place, may this meaning be? Rivers seems to derive it from the inky secretion of the animal. But the word *gole* means, not cuttle ink, but cuttlefish; and it is characteristic of these animals that they change color, chameleonwise, to suit the color of their surroundings.[36] May it not be that the thought in the native's mind, when he uses the word *gole,* is "can't find him," "can't see him"? (The chief of Muralug called black by a word which another native translated as "No, can't see him."[37]) And if this is the case, is it not natural

[33] The "violet test-wool" is apparently the relatively unsaturated violet that Rivers used in his experiments on color-matching (R, 49).

[34] R, 55, 56, 66, 94.

[35] R, 55, 56.

[36] R. Lydekker, "The New Natural History," VI., 327. The inky discharge of the cuttle fish is also regarded as a "defensive reaction"; and "cuttle ink" as well as "cuttle-fish" might suggest the thought "hard to find."

[37] R, 59.

that the adjective *golegole* should be applied to any large expanse within which no discriminable features can be made out? The dark of night, the skin of the body, the expanse of sea and sky, these are precisely the things to which a term meaning "uniform," "even," "undifferentiated," is suitable. Then, of course, when the native is asked to characterize black, blue, brown, violet papers or wools,—a wholly novel task,—he gives them the name that he has ascribed to the color-expanses, the large even color-fields; he calls them all *golegole*. On this hypothesis, we may pair *golegole* with *warowar*,[38] just as we paired *kakekakek* with *dudu;* for *warowar* is used of marked, patterned, particolored fields, as I have assumed *golegole* to be used of undifferentiated expanses.

Guesswork! the objector will reply. Guesswork, no doubt; but a guess that is suggested directly by the reading of Rivers' text. The Murray Islanders have, for instance—my table shows it—a color name derived from the secretion of the purple-yielding mollusc, and another derived from the name of the mollusc itself. The Western Tribe has, apparently, only one word, derived from the name of the mollusc; the same tribe has a term for dark brown derived from *saingui,* ink of cuttlefish, but no color name derived from the cuttlefish itself.[39] So far, then, as my data go, it is not fanciful to argue that *gole,* meaning cuttlefish, does not necessarily carry the meaning of inky black.

Even, however, if this particular guess is wrong, the argument on behalf of the Murray men is still not at an end. I see no reason why they should be interested in the "brilliant blue of sky and sea"; for the brilliant blue means fine weather and calm. One or two individuals of the Western Tribe called orange by a word meaning sunrise, and violet by a word meaning a cloud which is black on the one side and *kiaur* [violet?] on the other; there is no reference to blue sky.[40] The chief of Muralug called YG sea-color, G "like another kind of sea, another wind," and B "sea with another kind of

[38] R, 55.
[39] R, 56, 60.
[40] R, 61.

wind, plenty blow";[41] the inference is plain. I cannot here take the space to analyze these other vocabularies; I have chosen that of Murray Island as the most favorable to Rivers' hypothesis of defective blue-vision. But I point out that the case against that hypothesis does not stand and fall with my interpretation of *golegole*.

The subject of color nomenclature may be approached from another angle. No one can read the Murray Island terms without wondering what a group of civilized persons would make of colored papers, if they were required to give them the names of concrete objects. I thought it worth while to make an experiment on the matter. I cut two-inch squares from three of the Milton Bradley series, the spectral colors, the tints No. 1, and the shades No. 1; 54 colors in all. These were mixed in haphazard order, and were shown to 5 observers, two women and three men. The instruction ran: "You are to name these colors by first impression as soon as shown. Use no abstract color names, but use always the name of some concrete object that the color suggests to you. You may also use the words, Dull, Dark, Bright, Light, if the stimuli impress you in that way, without seeking any specific color name." I included these four terms because the Murray Islanders used words of the same significance. The experiment went smoothly in a period of some 20 minutes; the only modification of procedure was that, if an observer gave "sky" for blue, I called for a second concrete name. "Sky," as I have shown, was foreign to the Murray Island vocabulary. I subjoin the results for all colors in which blue was involved.

I have italicized the terms Dull, Dark, Bright, Light. In the whole series of 270 namings, these words were used 64 times (I exclude the cases in which Sky was changed to Light, though I regard these as significant). In the 120 cases of what we may call the blue-range, quoted above, the words occur 43 times. The general percentage is thus 23.7, the percentage for the blue-range is 35.8. If we take simply the three blues (GB, B, VB), the percentage is 35.5 (or, if the "Sky-Lights" are included, 42.2). Were then my

[41] R, 59. "These instances," says Rivers, "illustrate very well the liking of these people for similes." They seem rather to show the direction of practical interest.

observers "insensitive to blue"? Not a bit of it: their color-vision, by the regular laboratory tests, is normal.

TINTS No. 1.

Obs.	BG.	GB.	B.	VB.	BV.	V.	RV.	VR.
1 (f.)	[sky] *light*	[sky] *light*	*bright*	violet	hyacinth	violet	*light*	*light*
2 (f.)	robin's egg	small girls [frocks]	[sky] skies badly painted	amethyst ring	handkerchief border	ribbon	a perfume	orchid
3 (m.)	robin's egg	water	a dress	*light*	*light*	a gown	*light*	violet
4 (m.)	faded grass	[sky] *light*	[sky] *light*	*dull*	*dull*	*dull*	*dull*	*light*
5 (m.)	robin's egg	robin's egg	a wagon	*dull*	*dull*	hat-ribbon	*light*	violet

SPECTRAL COLORS.

1 (f.)	turquoise	*bright*	*bright*	ink	violet	violet	violet	*dark*
2 (f.)	china	Japanese porcelain	*bright*	aniline dye	poodle	a smart furniture shop	a magazine	books
3 (m.)	grass	water	[sky] background of picture	a dress	a blacked eye	peacock	*light, dull*	*dark*
4 (m.)	*light*	wall-paper	[sky] Japanese stamp	[sky] Japanese stamp	a dress	*dark*	*dull*	photographic paper
5 (m.)	verdigris	*dull*	lake-water	ink	violet	violet	violet	*dull*

SHADES No. 1.

1 (f.)	*dull*	*dark*	a dress	*dark*	ink	violet	*dark*	*dark*
2 (f.)	?	Vantine's shop	linen hangings	blotting pad	marking pencil	?	?	Alice in Wonderland [cover]
3 (m.)	dark trees	bluejay	baseball	*dark*	necktie	*dull*	grape-juice	grapes
4 (m.)	wall-paper	*dark*	*dark*	*dark*	sunset	*dull*	*dark*	?
5 (m.)	curtains	*dull*	?	*dull*	*dull*	violet	*dull*	*dark*

The question-marks mean that no judgment was given in the sense of the instructions; the observer said "thick like velvet," or "pleasant," or "ugly," or "dignified," or "thought of color-theories."

We cannot hope, of course, to invert this experiment; we cannot expect, if we ask for names of R, Y, G and B objects, to find a greater scattering in the B-list than in the others. For blue flowers, blue articles of dress, blue hangings and blue china and other household gear are common enough; and experiment shows that blue sky and green grass are more often associated than are red and yellow to any object of their color.[42] We live in a world where blue has its acknowledged place. The Murray Islander does not. Blood he knows, and red and yellow ochre, and turmeric, and the brilliant deep-green gall-bladder of the turtle—all of them objects of the highest importance in the conduct of his life; but blue he has no dealings with. "Every detail of the behavior of the natives in connection with the naming of color was consistent with the idea," Rivers says, "that blue was to them a darker or a duller color than it is to us."[43] I submit that their behavior is equally consistent with the idea that blue did not interest them.

Not much need be said of Rivers' work on the matching of wools. "The natives," we are told, "understood what they were required to do very readily in most cases."[44] Rivers does not himself inform us what this requirement was; but it was evidently the matching of wools for hue (color tone). The lack of any explicit statement to this effect is, I think, significant. We are so accustomed to classify colored objects by their hue, their "color" proper, that the classification seems to us to be natural and normal. The Murray Islander, however, appears to classify by total impression; the wools appeal to him by their combined hue, tint and chroma; and

[42] I asked the 19 students who happened to be in my laboratory at the time (8 women and 11 men) to write down the names of the first five R, Y, G and B objects that occurred to them. The order was varied, so that any practice-effect might be roughly compensated. Green grass came 17 times, blue sky 15 times, and red blood only 8 times, out of the possible 19. The B objects fell into the same rough groups as the others (person, personal adornment; clothing; articles of personal use;—house, household furniture; —vegetation, flowers, fruit; beasts, birds, insects;—landscape and seascape).

[43] R, 94. I am glad to find myself, on this point, in substantial accord with R. S. Woodworth ("The Puzzle of Color Vocabularies," *Psych. Bull.*, VII., 1910, 329 ff.).

[44] R, 49.

now one of these moments and now another may be the basis of his judgment. "One could often hear a native saying *kakekakek* to himself as he picked up a colorless wool to place with the green," *i. e.*, with Holmgren's apple-green test wool.[45] Rivers regards the "natural tendency to put together all the wools to which the same name was given" as a fallacy of nomenclature.[46] It looks rather as if the native varied the basis of his comparison from hue to tint and chroma. For "the same name" is not applied to different colors without a reason.

The actual "matches" are as follows:[47]

Test Wool.	Matched Wools.
Holmgren's red	reds, saturated pinks.
Bright green	greens, often blue-greens, occasionally almost pure blues.
Holmgren's purple	"very readily matched by all, though the majority refused to take pink wools if much less saturated."
Holmgren's apple green	"matched correctly by the majority," often "with bluish or violet wools of about the same saturation," by 7 with "neutral wools with a faint pinkish tinge."
Yellow	"matched correctly by nearly all," by 2 men and 5 children with reddish wools, by 1 man and 3 boys with blue wools.
Blue	matched by 27 (out of 107) "with violet as well as with blue or bluish green wools," by 1 man with an almost colorless wool, by 1 man with a brown and with a yellow wool.
Violet	"matched or compared by 12 with neutral wools and by 14 with distinctly reddish or pinkish wools," by 1 boy with a brown wool, and by the chief (who called all *kakekakek*) with a B and a G wool of about the same saturation.

The only facts that claim particular attention, after our foregoing study of the color vocabulary, are (1) the confusion of yellow with

[45] R, 49.

[46] R, 49 f., 93. Here is a chance for the obverse fallacy: my observer 2 (f.) gives *marking-pencil* for BV Shade 1 and for Spectral OR. We happen to know that examiners use differently-colored pencils; but suppose that the vocabulary was strange to us, and that we found the same object called BV and OR!

[47] R, 50.

blue by a man and three boys; and (2) the confusion of blue with yellow by one man. The boys are ruled out by Rivers on the ground of carelessness; the men remain. Rivers, however, gives us the clue to their behavior: "the yellow test wool used by me was a dull yellow," and blue is also regarded by the natives as a dull color. If the men muttered or thought *akòsakòs* as they matched or compared all is in order.[48] My observer 5(m.) gives the name "dull" to the "shades" of YO, Y, GB and VB.

I said just now that we are accustomed to classify colored objects by their hue; the following experiment shows that we can classify them, without trouble, by other means. I showed the 125 wools of the standard Holmgren set to my colleagues Drs. Boring, Foster and Weld, with the following request: "Arrange these wools into groups of similars, on the basis of first impression. Do not set out to judge by hue or tint or chroma; do not try to maintain any uniform basis of judgment; group the wools simply by the first impression they make upon you." I had recourse to highly experienced observers, because I thought that they would be less biased in favor of hue than undergraduate or graduate students; I thought also that they would be less afraid of making "foolish" matches. The result of the test is that Boring distinguishes 9 groups, Foster 18, Weld 4. Hue has a marked influence on choices: but Boring, who shows this influence most strongly, still throws together R, O, P, V, and in another group Y, G, BG, V; Foster groups with the Holmgren test-purple R, G, BG, B, P and V wools; and Weld groups with the same test wool R, O, Y, G, BG, P and V wools. This is a single test, to be sure, and the same observers would probably have made different groups had it been repeated. The fact remains, nevertheless, that if the prejudice in favor of hue is weakened or removed an expert observer will find likenesses of color-impression at least as wide-ranging as those of the Murray people.

It has not seemed necessary to enter in detail upon other tests than those conducted on Murray Island; and it does not seem necessary to discuss here the remaining tests of contrast, after-images, etc. I find in them nothing to invalidate what has been said above.

[48] R, 50 f., 55 f.

Let us now see what can be urged on behalf of the blue-vision of the islanders.

1. Within the period of a generation, "the great majority" of the Murray Islanders have learned to use the modified English word *bulubulu* for blue colors. In the tintometer experiments, "owing to the fact that *bulubulu* had become the general term for blue, there was no indefiniteness in the naming of this color"; and in the work on contrast the subjects "were all in the habit of calling blue *bulubulu* when talking to me."[49] This ready adoption of a foreign term seems to indicate that the natives could distinguish blues when once their interest and attention had been directed toward them.

2. In testing Europeans with the tintometer, Rivers found that oftentimes "the subjective contrast color was seen when the objective color failed to be recognized." On Murray Island this phenomenon was rare. "In many hundred observations, a color was only stated to be on the wrong side 15 times. No less than 6 of these occurred with a B glass when the opposite aperture was called *bam-bam* [turmeric]; in one case the opposite aperture was called R. The aperture opposite the R glass was called *bulubulu* 4 times and *giazgiaz* [newborn child, light] once; opposite a Y glass, B was seen twice and R once. Some of these were no doubt accidental, but it is interesting that the instance which occurred most often was when the objective color was B, to which they seemed so insensitive."[50] It is equally interesting that the Y glass gives a B contrast twice out of three times; and the contrast-R may be justified for a greenish yellow glass shown in poor illumination. Moreover, the R glass was a carmine, and the complementary of carmine is BG. The report of *bulubulu* four times out of five (though *giazgiaz* is itself an occasional word for blue) does not look like insensitivity.

3. In the test of negative after-images "R was most readily seen and B was doubtful."[51] But the stimuli were zigzags of colored

[49] R, vi., 2, 66, 71, 80 f.

[50] R, 81. The anomalous reds (if both are anomalous) may be compared with the anomalous light violet or purple (for me the color is definitely a purple) which workers in my laboratory have found even under achromatic adaptation; see L. M. Day, "The Effect of Illumination on Peripheral Vision," *Amer. Journ. Psych.*, XXIII., 1912, 573; and cf. the anomalous peripheral pinks or purples of R, 78.

[51] R, 82.

paper shown for 10 to 20 sec. on a gray ground, or of gray paper on a colored ground; and in what illumination? The method favors R; I can reproduce the results with observers of normal vision. There is nothing to show that the Murray men were not the victims of circumstances; and Rivers does not report corresponding tests of Europeans.

4. Rivers lays no great stress on his observations of color-preference. "Among saturated colors," however, "R easily had the first place, followed by B, while Y and G were distinctly less favored. . . . Complementary colors were commonly worn together, Y with B and G with R."[52] The first of these facts is a little surprising, in view of the practical importance of yellow and green;[53] taken together with the second, it may perhaps indicate that the natives, once their attention has been called to color as such, have a normal appreciation of blue.

Here, however, we are dealing with fairly large surfaces of color, surfaces that would be viewed in indirect as well as in direct vision. One of Rivers' strongest points against the native is that, peripherally, "the color blue was recognized readily, even more readily than other colors. The color of the patch used was saturated,"[54] but even so the results do not accord with those obtained in direct vision, so that the two sets of observations need to be reconciled. I have tried to show that the tests of direct vision are not convincing. It seems, also, that Rivers' argument from the sensitivity of the

[52] R, 84.

[53] I have said nothing of the supposed relative insensitiveness of the islanders to G, partly for reasons of space, partly because the position of B is the more important. H. E. Houston and W. W. Washburn ("On the Naming of Colors," *Amer. Journ. Psych.*, XVIII., 1907, 523) found no overlap of B and Y or R and G, but a marked confusion of B with G and B with P. It is noteworthy that H. K. Wolfe ("On the Color-Vocabulary of Children," *University of Nebraska Studies,* I., 1890, 23) finds no such confusion of B and G. Many points of Wolfe's investigation are of interest in connection with the Murray Island results. Thus "the pupils seem loth to confess their ignorance; four fifths of them attempted to name orange, and only one fifth knew what it was" (24); and the expressions "dark white" and "light dark" were used in good faith and with meaning (28).

[54] R, 79. If the tests were made in the order in which they are reported (cf. R, 49, 53, 70) the name *bulubulu* may have become standardized simply by the progress of the tests themselves.

peripheral retina proves too much. For if, by the aid of indirect vision, large expanses are better seen and more readily named than small patches of color, how is it that the brilliant blue stretches of sea and sky are still called *golegole* by subjects whose attention has been called to their color? And how is it that the blue of the rainbow is called *golegole?*[55]

On behalf of the Papuan, then, let so much have been said. I am not in the least concerned, in the present paper, with his macular pigmentation; that is another story. I am concerned only with the adequacy of Rivers' tests to various problems of sense-psychology. The tests appear to me to be inadequate. If Rivers can meet my objections, he must at any rate go beyond the limits of his printed report, and I shall have done some service in bringing out further observations and further arguments.

III. General Remarks.

What, now, are the requirements of a field-test? It should set the subject a task which is both simple and definite; it should be capable of performance in a relatively short time and with apparatus that is strong, portable and relatively cheap; it should be laid out so simply that its conduct is easily mastered and so definitely that there can be no variation in its procedure; and it should yield results that are directly relevant to the object of the test, are expressible in numbers and thus are intercomparable. These, in general terms, are the requirements: how shall we go to work to meet them?

We must realize, first of all, that the test is not a laboratory experiment; we must set ourselves at a certain remove from the laboratory; and especially we must avoid misleading analogies drawn from laboratory technique. I have pointed out that McDougall offers his æsthesiometric method as a combination of the methods of minimal changes and right and wrong cases; and I daresay that this title gives the method a sort of *cachet* in the minds of many readers; it may have had a reassuring influence upon McDougall himself. Rivers makes, I think a like mistake in his introductory discussion.

[55] R, 69 f. The good observer who described the rainbow from memory as red, white and black probably used the terms *kakekakek* and *golegole,* so that the description might be interpreted as red, faded-looking, blue.

"I believe," he writes, "that the smallness of the mean variation in most of the quantitative investigations will convince those acquainted with the procedure of experimental psychology of the trustworthiness of the observations."[56] But what is the procedure of experimental psychology? A set of instructions, carefully formulated and intelligently grasped; an instrument of precision; a large number of observations, made in accordance with a prescribed method and sufficient for mathematical treatment; a variation of conditions to throw this or that aspect of the subject-matter into high relief; experiments distributed regularly over months or years. Under such circumstances, truly, the m. v. may be an index of steadiness of attention or of general attitude. Not by any means necessarily under other circumstances! For a small m. v. may mean that the subject has over-simplified his instructions, and is performing an easier task than the task set him; or that he has discovered some secondary criterion of judgment, some short cut to response, and is not performing his allotted task at all. Or a small m. v. may mean that the unit of the instrument is too large, and that the performance of the task is thus artificially regularized. In these cases uniformity of result would spell laziness, or perverted ingenuity, or too gross a graduation of stimuli; in other and more extreme cases it might be due to fatigue or to motor habit. We cannot argue directly from laboratory-experiment to field-test.

Secondly, however, we must make full use of the laboratory. I suppose that most laboratories possess records of practice-work done by undergraduate students according to the schemata of the principal psychophysical methods; I have quoted a record of this sort, a determination of the two-point limen by the method of constant stimuli. Such records are not worth publishing, but they are worth preserving.[57] They furnish norms of the performance of comparatively un-

[56] R, 4.

[57] It is important to preserve not only the numerical values of the limen and of the measure of precision but also the rough data of the whole experiment. If the students are supplied with two cross-section forms and a carbon paper the duplicate may easily be obtained.

The rough data of the field-tests should also be accessible; all through this paper I have felt the need of further detail. I should think that sales to laboratories could be assured beforehand, enough to cover the cost of mimeographing the complete records.

practised observers, and they thus provide a ready means of testing any abbreviated test-method that may be proposed. Is it enough to take 10 observations of a kind, and has the 8-in-10 limen any definite significance?[58] Is it enough to take 5 observations, and has a 4-in-5 limen any value? These are questions not of pure but of applied mathematics; they can be answered only in the light of comparative data; and the control-data are ready to hand in the laboratories. It is useless to make tests in the field, and to repeat them later upon civilized subjects, until we know whether the test-procedures are themselves methodically reliable. So the laboratory may help on the score of method. It may also help in other ways; our discussion of the perceptive forms in aesthesiometry applies, *mutatis mutandis,* to more than one of the Torres Straits experiments. The analyses of the laboratory show what the tests are really doing, what psychological level has been reached. The test of an optical illusion, for example, may tell us nothing of the relative magnitude of the illusion in the case of savage and civilized subjects, but may nevertheless bring out the psychologically important fact that savage and civilized approach the particular task set them in different ways, or come to it in different attitudes of mind.

I am not inviting the field-worker to fall between two stools; I am rather pleading for coöperation. The field-worker seeks to obtain psychological data which shall enable him to rank a primitive race in relation to the various strata of his own civilized community. He knows, in a rough way, what can be done with a primitive population; the home-staying psychologist does not. The laboratory-worker, on his side, knows that a good many of the tests commonly employed are scientifically worthless; yet he cannot be continually playing the critic. Is it not a clear case for coöperation? So far as I know, we have to-day no approved æsthesiometric test that can be carried into the field. No: but if we settled, plainly and positively, what it is that we want the test to tell us; and if the field-worker kept guard over complexity of technique and the laboratory-worker over sources of error; then a test would be forthcoming.

[58] As regards the æsthesiometric test I have answered this question in the negative. My own first series with the apposed compass-points gives runs of 10 consecutive dual impressions with 80 per cent. correct judgments. Unfortunately the whole run was not 10 but 16.

In conclusion I offer a tentative suggestion as regards the general conduct of field-tests; it is always dangerous to be positive, but I take the risk. There is, I understand, a present tendency among those interested in mental tests to break away from tests of the "all or none" type and to substitute for them a set of tests which permits of fractional grading. The "all or none" tests, it is argued, cannot be applied to a long series of subjects, whereas tests which may be rated for part-performance have a practically unlimited range of application. I suppose that both kinds of test will be retained, each in what turns out to be its proper sphere; and I am disposed to think that, for anthropological purposes, the "all or none" type should, at first and on the whole, be preferred. Everyone who has worked with Hering's instruments must have been struck by the fact that they serve admirably for their one predestined experiment but that they can with great difficulty, if at all, be adapted to other uses. The demands of undergraduate teaching have led most of us, perhaps, consciously or unconsciously to favor instruments of a more flexible, more variously usable sort. Yet it may be that, for the primitive subject, tests of Hering's kind are, at least in the beginning, the more desirable. I wonder if a large number of testlets, each one sharply cut to a particular purpose, might not be better than tests which require serial or repeated observation; and if the single-value result might not lend itself to mathematical treatment better than the somewhat arbitrarily chosen "representative" value of a test-series.