OIL & FAT INDUSTRIES

VOLUME V MARCH, 1928 NUMBER 3

Color Sensitivity

Tests of Color Sense of A.O.C.S. Members and Data on Sensibility to Change in Lovibond Red*

BY IRWIN G. PRIEST†

Abstract

THIS paper reports primarily upon tests of the color sense of about thirty men who have had large experience in the color grading of vegetable oils in terms of Lovibond glasses, nearly all of these men being members of the American Oil Chemists' Society. Similar tests on five other observers are also included for comparison.

All of the oil chemists were first tested by the Stilling Chart test, but the results by this test (although reported) are not the principal matter of interest. This test was regarded merely as preliminary, serving to discover any gross abnormality of color sense.

The tests of specific interest relate to the observer's ability to report correctly very small differences in Lovibond red at 35 yellow 7.6 red on the Lovibond scale, brightness differences being eliminated so that the judgment depends, in effect, entirely upon the observer's sensibility to difference in dominant wave length at equal brightness. The subject's ability in this regard was tested by the well known psycho-physical method of "right and wrong answers." The results for each observer are expressed so as to show the probability of his perceiving correctly given small differences in Lovibond red, under certain specified conditions.

The chief results may be summarized as follows:

1. A difference of 0.1 red at 35 yellow 7.6 red is perceived *with certainty* by only very few exceptional observers. However, this difference is perceptible in the sense that its presence does affect the observer's judgment *in the average,* although he is very doubtful of the reality of such a difference.

2. The perceptibility of a difference amounting to 0.27 red is notably higher; but even this difference is not perceived *with certainty* by all normal observers.

3. About two-thirds of the total number of oil chemists tested perceive a difference of 0.3 red at 35 yellow 7.6 red with practical certainty.

4. Although they have not been tested at the small difference of 0.1 red, it appears likely that about one-third of the chemists already tested could detect this difference with practical certainty if put to the test.

5. Four of the chemists tested appear to be unable to detect with certainty differences as great as 0.5 red. One proved to be very uncertain about differences as great as 10 units.

* Publication approved by the Director, Bureau of Standards, U. S. Department of Commerce. This report was presented orally at the meeting of the American Oil Chemists' Society in New York, October 28, 1927. Other aspects of these same data have been considered in a paper, "Preliminary Data on the Least Perceptible Difference in Dominant Wave Length . . . ", presented at the Twelfth Annual Meeting of the Optical Society of America, Schenectady, New York, October 20, 1927. See Proceedings in J. O. S. A. & R. S. I., vol. 16, p. 117; Feb., 1928.

† *Chief, Colorimetry Section, U. S. Bureau of Standards.*

The author wishes to acknowledge the assistance rendered by his associate, Dr. Deane B Judd, in the preparation of this paper.

It is emphasized that these results are to be regarded as preliminary, and the following recommendations are made:

1. The tests should be repeated in a more systematic way and under standard constant artificial daylight illumination, including the present observers and all other members of the Society who are engaged in the color grading of oils. For each observer the tests should be made for each of the following red differences in the order given: 0.5, 0.3, 0.2, 0.1.

2. When the data so obtained are available, the Society should, in the light of these data, set a standard of performance for those to be charged with the duty of color grading of oil.

Incidentally, these tests have shown that the normal observer's sensibility to difference in wave length is much greater than has been believed heretofore. Good observers can detect with certainty a difference of wave length as small as 0.1 millimicron for yellow light. This is of the order of one-fifth or one-tenth of what has been given as the "least perceptible difference" by previous authorities. This aspect of these data has been discussed in a paper communicated to the Optical Society of America at its meeting, October, 1927 (abstract to appear in the "Journal of the Optical Society of America," vol. 16, p. 117; February, 1928).

Introduction

CHEMISTS dealing with vegetable oils are called upon to make the nicest discrimination of color in grading the oils. It is well known that many men in the general population are "color-blind," and that many others who may not be properly called "color-blind" do not have a strictly normal color sense nor the ability to perceive small differences in color. It seems obvious that the color grading of oils, which is a matter of paramount commercial importance, ought to be conducted by observers of proven ability in color discrimination. It developed, however, that very few, if any, of the men who have been engaged in this business for years had ever been adequately tested in this regard.

At the request of officers of the Society, I undertook certain tests of the color sense of members of the American Oil Chemists' Society at the annual meeting in Memphis, Tenn., May, 1927. This is the report upon these tests.

The purposes of this report are:

1. To give the members of the American Oil Chemists' Society an idea of the variations found among those who were tested, in order to afford a definite basis for considering the need of testing the color sense of those who are depended upon to grade the color of oil.

2. To give definite data upon the sensibility to change in red on the Lovibond scale[1] at 35 yellow 7.6 red.

In most cases the conclusions are to be regarded as indicative and tentative rather than definitive or final. In the case of most of the adverse findings, I do not believe that the tests at Memphis are *conclusive*. There are several reasons for this, e.g.:—(1) The circumstances of a convention are not the most favorable for a good score in such tests; (2) A low score on one day *may* be due to a merely temporary indisposition rather than to chronic abnormality; (3) The time available at Memphis was not adequate for an entirely satisfactory performance of these tests; (4) The test ought to be repeated with a constant standard artificial daylight illumination, after making certain improvements in apparatus and method.

Description of Tests

The tests made were of two kinds, viz:

1. Ability to read the Stilling Charts.[2]

[1] Cf. B. S. Sci. Pap. 547, particularly p.11. and bibliography relating to the Lovibond system of color specification on p. 21.

[2] J. Stilling. Pseudo-isochromatische Tafeln. 14th Edition, Leipzig, 1913.

2. Ability to perceive small differences in red on the Lovibond scale at 35 yellow 7.6 red, differences in brilliance being eliminated so that the judgment was concerned only with differences in hue.[3]

Concerning the nature of the Stilling Chart test, the reader is referred to the publication just cited.

The tests of ability to perceive small differences in red at 35 yellow 7.6 red on the Lovibond scale were made with apparatus arranged as follows: A Martens photometer[4] was set with its axis horizontal so that the two halves of the field were uniformly illuminated by light from the north sky transmitted by a small plate of milk glass held in a vertical plane before the photometer. A 35 yellow glass and a 7.6 red glass were mounted together between the photometer and the observer's eye so that the observer looking into the instrument saw both halves of the field in the color due to this combination of glasses. Between the milk glass and the photometer was a holder in which could be inserted a Lovibond glass (hereafter known as the "difference glass") so as to cover either half of the photometer field as desired.[5]

The tests were conducted in a well lighted room in daylight so that the observer's eyes were light adapted; however, the photometer was provided with an eye shade to protect both of the eyes from disturbing side lights.

The tests were conducted substantially according to the following procedure:—The observer looked into the photometer and viewed the two-part field. He was instructed to match the brightness by adjustment of the photometer and observe that the two halves of the field were then exactly alike in color. A Lovibond red glass (the difference glass) adequate to produce a small but readily noticeable difference in hue was then introduced in front of the photometer so as to affect only one-half of the field, while a non-selective (neutral or so-called "colorless") glass was inserted on the other side. The observer again matched the brightness and observed the difference in hue. He was advised that, with this manner of observation, but with smaller differences in red, the test would consist in his reporting on which side the difference glass was inserted in each of ten separate trials in which he would be ignorant of its true position except in so far as he could tell from viewing the field. In other words, he knew that one side of the field was 35 yellow 7.6 red and the other, 35 yellow (7.6 + n) red, where n is some small fraction; and he was only required to indicate (by saying "right" or "left") in which side the "more red" appeared after he had matched the brightness. After each trial, while the observer looked away, the brightness match was disturbed by adjustment of the photometer (whether the difference glass was changed or not) so that he could not infer the position of the difference glass from the difference in brightness in the next trial. He was instructed to report his *slightest inclination* to one side

[3] The observers were accustomed in the general practice of oil color grading to the presence of a combination of both difference factors, *i. e.*, of hue and brilliance.

[4] Martens. Physikalische Zeitschrift, 1, pp. 299-303; 1900.

[5] This holder consisted of a small black box with slots to hold the Lovibond glass. On one side of this box, a clamp served to attach it to the photometer, while the other side carried the above mentioned plate of milk glass. (See Fig. 1).

TABLE 1

GENERAL SUMMARY OF TESTS OF COLOR SENSE MADE FOR A.O.C.S.

AND COMPARISON WITH DATA FOR OTHER OBSERVERS.

(Small Letters in Parentheses in this Table refer to Notes on the Following Page.)

Observer's Identification No. (a)	Class on Stilling Chart Test (b)	Perceptibility of Small Differences in Red at Lovibond 35Y 7.6R (c)							
		Red Diff. (Lovib. Units)	0.093	0.16	0.27	0.41	0.53	1.00	1.22
		Δ Λ (d) (millimicrons)	0.1	0.2	0.35	0.5	0.65	1.2	1.5
OIL CHEMISTS		Eye							
1	N+		(Not tested)						
2	N		(Record lost)						
3	N+	right		4	10				
4	N	left			9				
5	N	left		7	10				
6	Q	left			7.5(f)	10			
7	N	right		5	10				
8	N-	right			9	10			
9	N	right			7	10			
10	N	right		9	10				
11	N	right		9	9				
12	D	right			8(g)	4	8		
13	N+	right		9	9				
14	Q	right	(May 25 9		7		8		10
14		right	(May 26)	(h)	8				
15	N	left		10	10				
16	N+	right		9	9				
17	N	left			8				
18	N	right		8	9				
19	N	right		7	10				
20	N	left			6	10			
20		left		7	9				
21	N	right		10	9				
22	N	right		4	10				
23	N-	right			8				(i)
24	N	right			4		8(j)		
25	N	right		6	9				
26	N	left		7	9				
27	N	left		8	9				
27		right	(Without glasses)		9				
27		right	(With glasses)		8				
28	N+	right		9	10				
29	D?(e)	both						4(q) 9(r)	(k)
25									
30	N		(Not tested)						
31	N (a)		(Not tested)						
32	N (e)		(Not tested)						
Mean of all				7.5	8.5				
Mean of 17 also tested at red diff. 0.16					9.5				
Total number of observers				17	26				
Number scoring 10 (maximum possible)				2	8				
" " 9 or more				7	18				
" " less than 9				10	8				
" " more than 5				14	25				
Minimum score				4	4				
Number making only minimum score				2	1				
OTHER OBSERVERS (Approximately Normal Trichromats)									
D. B. Judd (l)		right	10	10					
K. S. Gibson (m)		left	7.4	8.2					
C. K. Walker (n)		right	8	10					
J. O. Riley (o)		right	7						
M. E. Brown (p)		left	8	9.5	9.5				

The values given in this part of the table, for observers other than the oil chemists, are estimated from more extensive data which are shown in detail in Table 2.

or the other even though he was very doubtful of the correctness of the answer. It was emphasized to him that a small (even if hardly perceptible) difference was actually present, and it was up to him to make his "best bet" on it. He was further advised that if he insisted upon saying "matched" (no difference) instead of "right" or "left," his answers so made would be arbitrarily divided into equal numbers of "correct" and "incorrect" in computing his score. Thus, the result would depend entirely upon chance quite unassisted by his vision. The purpose of this rule was to force the observer to disclose the

NOTES TO TABLE 1

(As indicated by Letters in Table)

(a) Oil chemists, men who are experienced in color grading oils in terms of Lovibond glasses. (Nearly all are members of A. O. C. S.). Except as noted, tests were made at the Memphis Convention, May, 1927.

(b) The symbols in this column have the following meanings:

N means normal. (N + means that the observer read the charts with unusual ease and accuracy. N — means that the observer had some difficulty in reading the charts, as shown by his hesitation.)

Q means that, although the observer can not be definitely classed as decidedly abnormal, he experienced so much difficulty in reading that one does not feel justified in classing him without qualification with those who read with much greater ease.

D means definitely defective in color discrimination.

D! means undoubtedly abnormal to a marked degree.

(c) The numbers in the columns headed by differences in red (0.093, 0.16, 0.27, etc.) are the number of times that the observer reported the difference *correctly* in 10 trials. Blank spaces indicate that no test was made for that difference.

(d) $\Delta \wedge$ is the difference in dominant wavelength (in millimicrons) corresponding to the difference in Lovibond red given in the column heading.

(e) Tests made at the Bureau of Standards, Washington.

(f) Saw "no difference" 5 times out of 10.

(g) After practice on the larger differences, 0.41 and 0.53.

(h) These tests were made rather late in the afternoon (about 5 P. M.). The illumination was not as high as desirable. This observer might do better in the morning or with a higher illumination.

(i) Incomplete test. Observer answered correctly 4 times in 4 trials for difference of 1.22 R.

(j) Observer very slow in making answers.

(k) This observer is an anomalcus trichromat (protanomaly), quite incompetent to make accurate color matches. He proved to be uncertain about differences as great as 10 units of Lovibond red!

(l) Dr. Judd is Associate Physicist in the Colorimetry Section of the Bureau of Standards. He has had a large experience in making color matches with great care.

(m) Dr. Gibson is Physicist in the Colorimetry Section of the Bureau of Standards. He has had much experience in colorimetry and photometry. N. B.—See footnote 6 in text.

(n) Miss Geraldine K. Walker is now Research Associate of the American Oil Chemists' Society, in the Colorimetry Section of the Bureau of Standards. These tests were made when she applied for this position. She had had previous experience in colorimetry at the Munsell Research Laboratory, but had never seen a Lovibond glass before these tests were made. The consistency of her current work in grading the glasses indicates that she would make a still better score in this test at the present time. (November, 1927).

(o) Mr. Riley is Aid in the Colorimentry Section of the Bureau of Standards. He had had some experience (not very great) in color matching before these tests were made.

(p) Miss Mabel E. Brown, as Assistant to Mr. Carl W. Keuffel in spectrophotometric work with Keuffel and Esser, had had a great deal of experience in photometry and spectrophotometry, but not so much in other color matching. She had never worked with the Lovibond glasses prior to this test.

(q) First trial.

(r) Second trial after practice with larger differences.

very smallest differences he suspected even if he was quite doubtful of their reality. It was made plain to him that he had "everything to gain and nothing to lose" by reporting his slightest suspicions as to the difference in color. Proceeding in this manner the conductor of the experiment placed the difference glass 5 times to affect the left field and 5 times to affect the right field, the observer being ignorant of the order in which this was done.[6]

[6] The usual order, with very few exceptions was: (1) left, (2) right, (3) right, (4) left, (5) left, (6) right, (7) right, (8) left, (9) left, (10) right.

The procedure, instructions, and advice to the observer were not always *literally* as just outlined. In practice it becomes almost impossible to treat every observer *exactly* alike.

Results of the Tests

The results of the tests are presented in Table 1 which shows the number of correct answers in ten trials for the several differences used in making the tests. In considering these results the following remarks should be noted:

1. Owing to lack of time the data are not as complete as could be desired. Each observer should be tested a number of times at each of several differences beginning with larger and working to smaller differences until the difference becomes so small that the observer's answers indicate his inability to perceive it.

2. The score 10 indicates that the difference in question was perceived with practical certainty. (The observer can report it correctly, *provided he trust his vision i.e., provided he answer in accord with his slightest suspicion of difference even if very doubtful.*)

3. Of course, even one incorrect answer in the 10 trials shows that the observer is somewhat less than "certain" about the difference. On the other hand we have no guarantee that he would answer correctly 100 times in 100 simply because he answered correctly 10 times in 10. It is, however, not likely that he would answer correctly much less than 90 times in 100. In comparing different individuals on the basis of only 10 trials we are hardly justified in distinguishing between those who score 9 and those who score 10. In default of more extensive data, we may say that a score of 9 or 10 indicates "practical certainty."

4. The score 5 indicates that he can perceive the difference just as well as a blind man. In other words, by flipping coins or throwing dice, he should get the correct answer about 5 times out of 10 without looking into the color comparator. (Cf. results for zero red difference in Table 2 and in footnote 6.)

It is believed, however, that the procedure actually followed in all cases was such as to lead to the same results as would be found by the procedure that has been described.

Only two observers (No. 6 at Memphis and Dr. Gibson in Washington) have insisted on reporting "matched" in some cases. On reading the MS of this paper, Dr. Gibson states that, according to his memory, the tests in his case (Tables 1 and 2) were *not* governed by the instructions outlined here. In particular, he says, his reports were *not* based on *a priori* knowledge that the difference glass was actually in all cases on one side or the other but were, in effect, based on the assumption that it might be on *neither* side, in which case "matched" would be a legitimate answer. It should also be recalled that nearly all observers actually are strongly inclined to report "matched" for differences less than 0, 3R; and would readily do so in many cases if not required to report "right" or "left." However, a careful repetition of the tests, December 27, 1927, without allowing the answer "matched" to be made, gives results for Dr. Gibson in complete accord with the "Fictional Fraction Correct" given for him in the last column of Table 2. The results of these new tests (December 27, 1927), the determination for each difference being based on 20 trials, are as follows:

Lovibond Red Difference	Fraction Answered Correctly	Fraction Answered "Left"	Clock Time: Start of Test	Clock Time: End of Test	Elapsed Time in Minutes
0.495	1.00	0.50	11:16 a.m.	11:20 a.m.	4
0.296	1.00	0.50	11:26	11:30	4
0.185	0.85	0.65	11:31	11:46	15
0.104	0.80	0.60	11:49	12:13 p.m.	24
0.000	0.35	0.45	12:16 p.m.	12:32	16

The following particulars concerning these tests of December 27 should also be noted:

(1) The order of presentation was *not* that given in the beginning of this footnote, but was in each case determined by chance, subject to the condition that the number of presentations on the left equal the number on the right.

(2) The observer says he felt no greater difficulty in making decisions for the zero difference than for the difference 0.104. He was not aware that the test was being made for zero difference but supposed that a very small difference (less than he could perceive with certainty) was being used in the test.

(3) The "elapsed time" in making the answers is significant as indicating the observer's greater hesitation in deciding for the smaller differences.

(This footnote will be better understood after reading the next section of the paper on "Results of the Tests.")

TABLE 2

SENSIBILITY TO CHANGE IN RED AT LOVIBOND 35 YELLOW 7.6 RED

SUMMARY OF ALL TESTS AT THE BUREAU OF STANDARDS, MAY 11 TO OCTOBER 8, 1927.

Observer	Date	Difference in Lovibond Red Units	Difference in Dominant Wave Length	Total No. of Trials in one set of Answers	Number of Correct Answers C	Number of Incorrect Answers I	Number called "Match" M	Fraction Correct	
								Literal $\frac{C}{C+I+M}$	Fictional $\frac{C+\frac{1}{2}M}{C+I+M}$
D.B.Judd	5-11-27	0.16	0.2mμ	20	20	0	0	1.00	1.00
	5-11-27	.093	.1	20	20	0	0	1.00	1.00
	9- 9-27	.093	.1	20	20	0	0	1.00	1.00
	5-13-27	ZERO *	0.00	20	9	11	0	0.45	0.45
K.S.Gibson (N.B.-See footnote 5 in text)	5-11-27	0.16	0.2	20	14	1	5	0.70	0.82
	5-11-27	.093	.1	20	10	2	8	.50	.70
	9- 9-27	.093	.1	20	14	3	3	.70	.78
G.K.Walker	8-26-27	0.17	0.2	10	10	0	0	1.0	1.0
	8-26-27	.104	.1	10	7	3	0	.7	.7
	8-26-27	.093	.1	10	9	1	0	.9	.9
M.E.Brown	10- 8-27	0.28	0.35	20	19	1	0	0.95	0.95
	10- 8-27	.15	.2	20	19	1	0	.95	.95
	10- 8-27	.104	.1	20	17	3	0	.85	.85
	10- 8-27	ZERO *	0.00	20	7	13	0	.35	.35
J.O.Riley	9- 9-27	0.093	0.1	20	14	6	0	.7	.7

* The meaning of "ZERO" here is as follows: The test being conducted just as in other cases, a neutral glass (non-selective crown) is surreptitiously substituted for the fractional Lovibond glass. The observer imagines he is responding to small differences as in the bona fide experiments. The most probable value for $(C+\frac{1}{2}M)/(C+I+M)$ in these circumstances is 0.5. Of course, it is not to be expected that this value will be obtained every time in such an experiment. Indeed, the probability of obtaining exactly 0.5 in 10 trials is only 0.24, and in 20 trials, only 0.18. However, the probability that $(C+\frac{1}{2}M)/(C+I+M)$ will lie between 0.3 and 0.7 inclusive, for 20 trials, is 0.96.

Fig. 1. *Testing the Ability to Perceive Small Differences in Lovibond Red*

The observer, looking into the Martens photometer, sees a circular field divided by a vertical diameter. Both halves of the field are illuminated by daylight transmitted by the milk glass, the two halves being illuminated respectively by different parts of the milk glass. The Lovibond combination 35*Y* 7.6*R*, *being placed permanently between the observer's eye and the photometer, determines the color of both halves of the field. The difference glass* (*e.g.* 0.1*R*, 0.2*R*, *or* 0.3*R*) *is inserted between the milk glass and the photometer by the conductor of the experiment so as to affect the color of one half only, which may be either the right or the left at the option of the conductor of the experiment. In any one set of answers the difference glass is actually placed on the right as many times as on the left, but in an order unknown to the observer. The observer has no means of knowing on which side it has been placed except by his ability to perceive hue differences. He matches brilliance* (*by turning the circle, C, which rotates the nicol prism of the Martens photometer*) *and indicates the side on which he believes the difference glass to be, by saying "right" or "left." The conductor keeps a score of correct and incorrect answers.*

5. Any score above 5 *indicates* that *probably* the observer's color sense had *something* to do with determining his answers, but obviously a score of 6 is, in practice, not of great significance. With "good luck" a blind man *might* guess right 6 times out of 10. His most probable score is 5, but his chance of scoring 4 or 6 is almost as great. (See tabular values just below.) However, the blind man's chance to guess correctly 10 times out of 10 is less than 1 in 1000. The significance of the scores in general is best made clear by considering the "blind man's chance" of making them, i.e., the probability of obtaining these scores by pure chance without the aid of the color sense. These probabilities are *approximately* as follows:

Score	0	1	2	3	4	5	6	7	8	9	10
Probability of score	1/1000	1/100	4/100	12/100	20/100	25/100	20/100	12/100	4/100	1/100	1/1000

The probability that the "blind man's score" be 4, 5, or 6 is 66/100 and the probability that it be 3, 4, 5, 6, or 7 is 89/100. The probability that it be more than 5 is about 38/100. The probability that it be more than 6 is about 16/100. The probability that it be more than 7 is only about 5/100; more than 8, 1/100. The observer who has answered correctly 10 times in 10 trials has done something which can be expected by chance only about once in a thousand times. If he repeats this performance (i.e., answers correctly 20 times in a total of 20 trials) he has done something which can be expected by chance only about once in 1,000,000 times.

6. For purposes of comparison, there is subjoined to Table 1 data on other normal observers obtained at the Bureau of Standards; and these data are shown in detail in Table 2.

Before beginning the tests the observer was first coached by telling him the true position so as to fix his criterion for the difference to be observed. Also, coaching and practice intervened between successive trials as indicated in the table. The observer was very hesitant in deciding, but never answered "matched." He was most confident of the correctness of some of the wrong answers. He also expressed the opinion that his sensibility was changing during the experiment. It appears from these data that the difference in Lovibond red at 35 yellow 7.6 red must be more than 3 units in order to be perceived with certainty by this observer, when the brightness difference has been eliminated.

This observer is not an oil chemist and not engaged in the use of Lovibond glasses; but the data are important as showing the uncertainty of Lovibond matches by an

TABLE 3

Hour A.M.	Difference in Lovibond Red Units	Total Number of Trials	Number of Correct Answers	Number of Incorrect Answers	Number Called "Match"	Fraction Correct
9:15	1.2	20	14	6	0	0.70
(Coaching and practice intervened)						
9:30	1.2	20	16	4	0	0.80
(More coaching and practice)						
9:40	3.1	20	19	1	0	0.95
(More coaching and practice)						
10:00	1.2	20	15	5	0	0.75

7. Data on an anomalous trichromat of the extreme deuteranomalous type are shown in Table 3. These data were obtained at the Bureau of Standards, May 13, 1927 in the same way as the data shown in Table 2.

observer of this type. One of the oil chemists (No. 29) is also an anomalous trichromat, but of the protanomalous type.

Conclusions

The following conclusions are

drawn from the data presented above.

1. A difference of 0.1 R at 35Y 7.6R (about 0.1 millimicron in dominant wave length) is perceived *with certainty* by only very few exceptional observers. Of all our observers, only one (D. B. Judd, Bureau of Standards) so far has a perfect score in reporting this difference correctly. He has reported it correctly 20 times out of 20 in *each of two* sets of 20 trials. (The chance of such a performance by accident is less than one in a million million.) None of the oil chemists was tested at this small difference. However, of 17 who were tested at a difference of 0.16 R, only 2 made perfect scores; and, of 26 who were tested at a difference of 0.27 R, only 8 made perfect scores.

Nevertheless, the difference of 0.1 R is perceptible in the sense that it *does* affect the result *in the average.* If this difference had no effect in determining the observer's answers, then in a large number of answers we should expect about an even division between correct and incorrect. In dealing with a number of observers, we should expect about as many to score below 50% correct as above 50% correct. Consider the data in Table 2 from this point of view. For the difference of *about* 0.1 R we have seven cases on the basis of 20 answers in one set.[7] Note that in each and every case the percentage of correct answers is not merely greater than 50 but is indeed greater than 60; and, with only one exception, greater than 70. The chance of the percentage of correct answers being greater than 60 in *one* such case "by accident" is about 13 in 100 (or, more accurately, 0.1316). The chance of it being greater than 60 in all of the seven cases is[8] less than 7 in 10,000,000. Since, in some of these cases, the percentage correct is still greater, the chance of the answers occurring by accident as they actually did occur is even less than these computations indicate. In the ordinary usage of language and in so far as it is usual or useful to make a distinction between doubt and certainty, this is equivalent to saying that it is *"certain"* that this distribution of answers is not due to chance. The trials with "zero" difference in red (Cf. Table 2 and particularly the note to that table) show that the percentage of correct answers is in fact a fair approximation to what would be expected to result from chance distribution according to the theory of probability assuming the answer "left" to be just as probable as the answer "right," in each trial. Consequently, it appears certain that the wide departure from chance distribution which occurs when the 0.1 R glass is substituted for the zero glass (all other conditions remaining the same) must be due to the color difference occasioned by the 0.1 R glass.

2. For a difference of 0.27 R (about 0.33 millimicron in dominant wave length), we find the perceptibility notably higher. Eight out of 26 oil chemists tested answered correctly 10 times in 10 trials (See Column 6 of Table 1). Eighteen scored 9 or 10; that is, they perceived the difference with practical certainty. Only one scored less than 6. In a total of 30 tests (on 26 observers) only one score was less than 6. The significance of these results as indicating the

[7] Two sets each for Judd and Gibson and one set each for Riley, Brown, and Walker, considering Miss Walker's two sets of ten as one set of 20 for convenience in computing.

[8] Computed as $(0.1316)^7$.

perceptibility of this difference is emphasized by considering the following probabilities. The chance of each and every one of 8 separate scores being 10 by accident is much less than 1 in 10^{24}. The chance of each and every one of 18 separate scores being 9 or 10 by accident is less than 4 in 10^{36}. The chance of each and every one of 29 separate scores being more than 5 is but little more than 5 in 10^{13}. The evidence that this difference (0.27R) is perceptible in the sense that its presence affects the answers to a notable extent is overwhelming. On the other hand, it must be admitted that there are many individuals to whom it is not perceptible *with certainty;* moreover we would not be justified in classing these individuals as abnormal.

3. A selected group of about 19 of the men tested at Memphis perceive a difference of 0.3 R with practical certainty.[9]

4. Although they have not actually been tested at a difference of 0.1 R, it appears likely (from the other tests) that about one third of the oil chemists already tested could detect with practical certainty[10] a red difference as small as 0.1 R at 35 yellow 7.6 red.

5. Four of the oil chemists tested appear to be unable to distinguish with certainty differences as great as 0.5 R at 35 yellow 7.6 red, and one (No. 29) is uncertain about differences as great as 10 units.[11]

[9] That is, their scores are 9 or 10 in this test.

[10] That is, by the test here described, would score 9 or 10.

[11] This observer states that he is not actively engaged in color grading oils, and has not been concerned with this work for some years past.

Individual Reports

For obvious reasons individuals among the oil chemists have been designated only by number in this report.

Each individual tested will be privately advised of *his own* score in the tests reported in this paper, but obviously *the author is not willing to communicate the results for named individuals to others than the person to whom they refer. All persons are requested to refrain from asking for the scores of others than themselves.*

Recommendations

In the light of the experience and preliminary data gained in these experiments the following recommendations are made:

1. The tests should be repeated in a more systematic way and under standard constant artificial daylight illumination, including the present observers and all other members of the Society who are engaged in color grading oils. For each observer, the tests should be made for each of the following red differences in the order given: 0.5, 0.3, 0.2, 0.1.

2. When the data so obtained are available, it is suggested that the Society should, in the light of these data, set a standard of performance for those to be charged with the duty of color grading oil.

Bureau of Standards, Washington, May-December 1927.

Revised, March 5, 1928.