J. clin. Path. (1948), 1, 68.

# THE ERROR OF COLOUR MATCHING WITH HALDANE'S HAEMOGLOBINOMETER

BY

JOHN L. D'SILVA AND E. C. TURTON
*Department of Physiology, St. Bartholomew's Hospital, London*

(RECEIVED FOR PUBLICATION, SEPTEMBER, 1947)

Haldane and Smith (1899–1900), in the description of their well-known technique for the determination of haemoglobin in blood, realized that the comparison of the tints of solutions of carboxyhaemoglobin contained in tubes may be inaccurate. They wrote as follows (p. 334): ". . . when the tints are really equal the tubes may look unequal, the apparent inequality being, however, in the other direction when they are transposed. This commonly affects not merely the depth, but also the quality of the tints. Errors of such a kind are avoided by repeatedly transposing the tubes during the comparison." In the same communication (p. 333), they stated: "The points at which there was just too little, and just too much, water were noted, and the mean taken as the correct result."

This advice, to judge from a recent publication, is often overlooked. Macfarlane (1945, p. 65) inquired of sixty "competent observers" engaged in haemoglobin surveys for the Medical Research Council as to the precise manner in which they used the instrument. Thirty-nine stated that they transposed the standard and unknown tubes as advised by Haldane, whereas seventeen did not transpose the tubes, and four did not reply to the inquiry. It appeared further that, of these sixty observers, only five adopted Haldane's advice in its entirety.

We have investigated the accuracy with which two observers (J.L.D. and E.C.T.) could match the tints of solutions of carboxyhaemoglobin under different experimental conditions. The results of these experiments are in the course of publication elsewhere. They may be stated briefly as follows:

(*a*) When two identical, ungraduated, haemoglobinometer tubes containing solutions of carboxyhaemoglobin are viewed side by side, and touching against a background of an illuminated ground-glass screen, the observer may judge the solution on the right to be "too dark" (that is, strong) when, in fact, it is the weaker of the two.

(*b*) The error alters as the distance between the observer's eyes and the comparison tubes alters.

(*c*) At any one distance the error changes if the intensity of illumination is changed, or if a daylight bulb is substituted for a "clear" bulb as the source of light.

In the present series of experiments, the errors of twenty observers (six pathologists, four senior technicians, and ten medical students) were investigated.

## Plan of Research

The following groups of experiments were carried out.

(*a*) Each observer made a series of comparisons in daylight.

(*b*) Each of the ten students made two series of comparisons, (i) with the eyes at 40 cm., and (ii) with the eyes at 125 cm. from the comparison tubes, the source of light being a "clear" electric bulb set up as described below.

(*c*) Conditions were found in which the student observers could compare the tints of solutions of carboxyhaemoglobin with considerable accuracy.

## The Experimental Method

A series of nine dilutions of carboxyhaemoglobin was made up from a 2 per cent stock solution of blood whose haemoglobin content was 100 per cent (Haldane). The tint of the solution labelled "1" approximately matched a Haldane standard. The remaining eight solutions differed from one another and from the solution "1" by steps of 2 per cent. The strongest solution was labelled 1.08, and the weakest 0.92.

The same two ungraduated haemoglobinometer tubes of average internal diameter 6.60 and 6.63 mm. were used in all these experiments, and in every comparison of tints the same tube always occupied the left-hand position. The tubes were not transposed. The tubes were filled by an experimenter with any pair of the nine available solutions of carboxyhaemoglobin. The observer's part in the experiment was solely to report whether, in his opinion, the solution in the right-hand tube was "too strong" or "too weak" relative to that in the left-hand tube. If he could not distinguish between them, he reported a "match." The tubes were set up side by side and touching in a cork holder. The solutions were presented for comparison in a random order, and the number of comparisons which made up an experiment was left to the discretion of the experimenter, who was guided by the observer's replies and continued the experiment until he was sure that a given difference in concentration between the solutions in the two tubes could be recognized correctly, and with certainty, by the observer.

From the observations, the ratio of the concentrations of the solutions in the right- and left-hand tubes which gave the observer an equal chance of recognizing a difference in tint correctly was calculated. This was called the "50 per cent ratio." Thus, if the 50 per cent ratio is 1, it means that the observer did not bias his observations in favour of either tube. On the other hand if it is 0.95, it means that the observer judged the solution in the right-hand tube to be "strong" when, in fact, it was "weak." The value of the 50 per cent ratio, therefore, is a measure of an observer's error in tint discrimination.

The method of calculation is based on that used for determining the LD50 for a drug (*cf.* Gaddum, 1933), Behrens's (1929) calculation to allow for the smallness of the sample being employed. When the normal equivalent deviation (N.E.D.) of the percentage of correct observations at each ratio tested is plotted against the ratio concerned, the points lie on or near a straight line. The ratio corresponding to zero N.E.D. represents the ratio at which the observer has an equal chance of discriminating between the tints of the two solutions. The standard deviation (S.D.) of the 50 per cent ratio was calculated from formula 13 on p. 27 of Gaddum's (1933) paper. The application of the method is described in detail in another communication (D'Silva and Turton).

## Comparisons in Daylight

Each of the twenty observers undertook a series of comparisons as described above, in a laboratory lighted by daylight. The observer stood and held the tubes in the cork holder against a ground-glass screen placed in front of the window at eye level. The intensity of illumination varied, but in no case would it have been deemed too low for haemoglobinometry in a pathological laboratory.

Having permitted the observer to choose the distance between the tubes and his eyes, this distance was determined by the experimenter as follows: two cardboard scales marked by vertical lines at intervals of 5 cm., the one being slightly wider than the other, were arranged parallel with the observer's line of sight. One was about 30 cm. and the other 40 cm. from that line. The "zero distance" on each scale was some 2 cm. nearer to the observer than the ground-glass screen, and coincided with the usual position of the comparison tubes when they were held near the screen. Unknown to the observer, the experimenter observed the distance at which the tubes were held from the eyes for each comparison. The arrangement of the graduated scales eliminated any error due to parallax.

Two sets of results were obtained from this series of experiments: (*a*) the error of comparison of each observer; and (*b*) the least, the mean, and the greatest distance between the tubes and the observer's eyes for each set of comparisons.

The error of each observer is recorded in Table I. It will be seen that in each group some observers are fairly accurate and some are inaccurate. The errors of both groups are similar in spite of the fact that one group of observers was trained to compare these tints and the other was untrained. Some of the student observers had used the Haldane method as a class experiment, so that to achieve the degree of accuracy shown by them required little or no training.

TABLE I

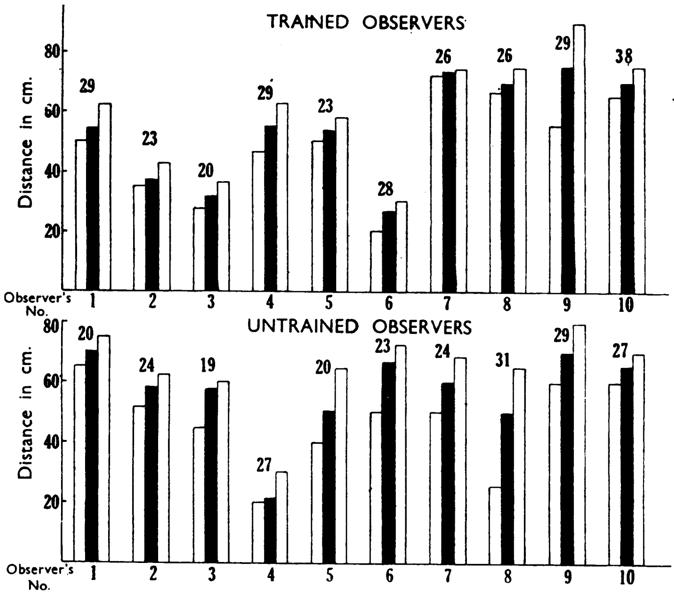
ERRORS MADE BY TWENTY OBSERVERS IN COMPARING THE TINTS OF SOLUTIONS OF CARBOXYHAEMOGLOBIN IN DAYLIGHT

| Trained observers | | Student observers | |
|---|---|---|---|
| Initials and no. | 50% ratio and 2S.D. | Initials and no. | 50% ratio and 2S.D. |
| 1. A.E.K.-S. | 0.997 ± 0.017 | 1. R.Z. | 0.970 ± 0.009 |
| 2. B.C.C. | 0.969 ± 0.008 | 2. P.J.L. | 0.969 ± 0.012 |
| 3. M.B.McI. | 0.969 ± 0.008 | 3. G.C.H.C. | 0.968 ± 0.014 |
| 4. A.McG. | 1.033 ± 0.010 | 4. C.C.M. | 0.964 ± 0.014 |
| 5. R.J.S. | 0.960 ± 0.008 | 5. P.H.S. | 0.958 ± 0.013 |
| 6. d'A.K. | 0.957 ± 0.012 | 6. A.J.W. | 0.957 ± 0.015 |
| 7. G.D. | 0.956 ± 0.014 | 7. R.H.G.B. | 0.945 ± 0.012 |
| 8. H.I.C. | 0.955 ± 0.011 | 8. D.M. | 0.937 ± 0.016 |
| 9. D.A.L. | 0.941 ± 0.016 | 9. D.E.W. | 0.936 ± 0.026 |
| 10. G.E.M. | 0.873 ± 0.018 | 10. W.R.C.B. | 0.921 ± 0.009 |
| Average 0.961 | | Average 0.953 | |

The diagram shows the mean distance (black area), the least distance, and the greatest distance at which comparisons of tint were made. In the student group, 50 cm. or above was the usual distance, as it was in the group of trained observers, except for one observer in the former group and three in the latter who made their comparisons at distances nearer than 50 cm. It was not possible to correlate the observer's error with the distance at which he made the comparisons of tint. The distance of comparison usually varied within fairly narrow limits except in the case of D.E.W. (student group) and D.A.L. (trained group), whose comparison distances were very variable.

## Comparisons in Artificial Light

The source of light was a 60-watt "clear" bulb mounted inside a black box, the front of which was covered except for a circular aperture 4 in. in diameter placed opposite the bulb. A ground-glass screen was set up




The mean distance (black rectangle), the least, and the greatest distance at which observers made their comparisons in daylight. The observers' numbers correspond to those in Table I. The number of observations is given at the head of each group of three rectangles.

3.5 ft. from the bulb (illumination of 10 f.c. approximately) and the comparisons of tint were made against it. The room was darkened apart from the light referred to above. The comparisons were made as described previously.

Each of the student observers carried out a set of comparisons with the eyes at (*a*) 40 cm. and (*b*) 125 cm. from the tubes. The results are in Table II.

In every subject, except P.J.L., R.Z., and W.R.C.B., the 50 per cent ratio changed when the distance between the eyes and the comparison tubes was altered from 40 to 125 cm. The change in distance can affect the error of comparison in three ways: (*a*) the 50 per cent ratio may be unchanged; (*b*) it may become less; or (*c*) it may become greater. Of the observations made at 40 cm., only three observers (R.H.G.B., C.C.M., and P.H.S.) had 50 per cent ratios not significantly different from 1. All the observers (except R.Z. at 40 cm.; *cf.* Table II) were consistent in their observations, though only three were accurate. Their results showed that the precise difference in concentration between the solutions in the two tubes which led to difficulty in distinguishing between the tints was different for different observers.

### Conditions under which Accurate Comparisons were made by some Observers

In this series of experiments (results in Table III) we tried to find conditions under which the observers' 50 per cent ratios would not differ significantly from 1. Three of them (R.H.G.B., C.C.M., and P.H.S.) made accurate tint comparisons at 10 f.c with the eyes 40 cm. from the tubes (*cf.* Table II). The results of the other seven observers show that when the distance of comparison was altered from 40 to 125 cm., in three (D.M., A.J.W., and G.C.H.C.) the 50 per cent ratio increased, in one (D.E.W.) it diminished, and in three (P.J.L., R.Z., and W.R.C.B.) it was unchanged.

Of the three observers for whom the 50 per cent ratio increased, in two it changed from less than 1 to more than 1. For these two (D.M. and A.J.W.) we found empirically that accurate tint discrimination

TABLE II

ERRORS OF COMPARISON OF TEN STUDENT OBSERVERS WITH AN ILLUMINATION OF ABOUT 10 F.C. PROVIDED BY A 60-WATT "CLEAR" BULB

| Observer | 50% ratio and 2 S.D. | |
|---|---|---|
| | 40 cm. | 125 c.m. |
| R.H.G.B. | 0.999 ± 0.012 | 1.044 ± 0.016 |
| C.C.M. | 1.005 ± 0.012 | 1.056 ± 0.017 |
| P.H.S. | 1.006 ± 0.015 | 0.950 ± 0.010 |
| P.J.L. | 0.980 ± 0.011 | 0.977 ± 0.013 |
| D.M. | 0.968 ± 0.009 | 1.027 ± 0.022 |
| A.J.W. | 0.967 ± 0.015 | 1.057 ± 0.014 |
| D.E.W. | 0.955 ± 0.012 | 0.912 ± 0.015 |
| R.Z. | 0.954 ± 0.040 | 0.940 ± 0.017 |
| W.R.C.B. | 0.945 ± 0.014 | 0.958 ± 0.013 |
| G.C.H.C. | 0.924 ± 0.022 | 0.960 ± 0.014 |

TABLE III

CONDITIONS OF ILLUMINATION AND DISTANCE WHICH ALLOWED SEVEN UNTRAINED OBSERVERS TO COMPARE THE TINTS ACCURATELY: C = "CLEAR" BULB; D = "DAYLIGHT" BULB

| Observer | Illumination in f.c. | Dist. in cm. of eyes from tubes | 50% ratio and 2S.D. |
|---|---|---|---|
| R.H.G.B. | C 10 | 40 | 0.999 ± 0.012 |
| C.C.M. | C 10 | 40 | 1.005 ± 0.012 |
| P.H.S. | C 10 | 40 | 1.006 ± 0.015 |
| D.M. | C 10 | 50 | 0.991 ± 0.009 |
| A.J.W. | C 10 | 60 | 0.995 ± 0.014 |
| D.E.W. | C 100 | 40 | 0.991 ± 0.009 |
| G.C.H.C. | D 10 | 85 | 1.007 ± 0.014 |

was achieved when the distance of comparison was 50 cm. and 60 cm., respectively. On the same basis, G.C.H.C. presumably needs a longer comparison distance than 125 cm. (*cf.* Table II), but comparisons made at distances greater than 125 cm. are less accurate. When illumination was provided by a daylight lamp, however, G.C.H.C. was accurate at a comparison distance of 85 cm. (Table III).

In the case of D.E.W., the 50 per cent ratio diminished as the comparison distance increased and was always less than 1 at 10 f.c. We found no distance at which this observer could make accurate comparisons at 10 f.c. At 100 f.c., however, he made accurate comparisons when his eyes were 40 cm. from the tubes.

The remaining three observers (P.J.L., R.Z., and W.R.C.B.), for whom a change in comparison distance was without significant effect on the 50 per cent ratio, were not dealt with further. P.J.L. had only a small error (2 per cent $\pm$ 1 per cent) in his 50 per cent ratio at 40 cm. and it was felt that any significant change in this in the direction of greater accuracy would be difficult to establish. W.R.C.B. had a large error, but we could not find any conditions under which his 50 per cent ratio was near 1. It was later discovered that he had uncorrected astigmatism which may have contributed towards his inaccuracy. R.Z. at 40 cm. (Table II) had a wide scatter in his observations (as is shown by a high S.D.), and though he was more consistent at 125 cm. he was just as inaccurate.

### Commentary

Three points arise from this work in connexion with the determination of haemoglobin.

(*a*) All but one of our trained observers made appreciable errors in comparing the tints of solutions of carboxyhaemoglobin. They were no more accurate than our untrained observers and, as we have shown previously, the inaccuracy is not due to instrumental errors. It seems probable that it is the result of the make-up of the individual's visual apparatus. None of our observers was aware, during the course of the experiment, of any inaccuracy of comparison, so for accurate work it would be advisable for workers who do not transpose the haemoglobinometer tubes during their tint-matching procedure to undertake experiments to determine the magnitude of their errors.

(*b*) Macfarlane's recent communication (1945, p. 67) shows that six different methods of haemoglobinometry were in use by 56 observers, namely: (i) standard always kept on the left, (ii) standard always kept on the right, and (iii) standard first on one side and then on the other. All the observers using these three methods took the first observation of "match" to be the true match point. The observers (iv), (v), and (vi) used the methods (i), (ii), and (iii), respectively, in their tint-matching procedure, but took the mean of the last observation of "strong" and of the first observation of "weak" as the true match point.

We can illustrate the possible errors which can be made by an observer by considering one with a 50 per cent ratio of 0.90, that is, one who judges the right-hand tube to be 10 per cent stronger than it is. Let us assume that he uses the experimental procedures (i) to (vi) (above) in turn, and further that the true match should be at 100 on the Haldane scale.

If he keeps the standard always on the left, he will match at 110. If he keeps the standard always on the right, he will match at 90 for the same reason. If he uses (iv) or (v) his match points are more likely to be 112 and 92, respectively, because the difference between his last observation of "strong" and his first observation of "weak" will be 4 per cent or 6 per cent on the Haldane scale. In each of these four experiments he will have a large error solely because of his inaccuracy in tint matching.

Now consider the situation when the same observer examines the unknown solution first on the right of and then on the left of the standard. Assume that in the course of the experiment, dilution has reached the "90-mark." When the standard is to the left of the unknown, the latter will be recorded as strong. When the positions of the tubes are reversed, however, a match may be recorded because of the observer's 10 per cent error. The observer will be faced with a dilemma in which the same solution appears to be strong when it is viewed on one side of the standard, and a match when it is viewed on the other side. Two courses only are open to him. Either (*a*) he can record his judgment as a non-committal "match" and continue his observations, finally taking the mean of the readings at which he judged the unknown to be just too strong and just too weak as the true match point (*cf.* procedure (vi), above); or (*b*) he can judge the dilution at which the unknown appears as strong, when it is viewed on one side of the standard, as it appears weak when it is viewed on the other, and call this the true match point. The latter procedure introduces uninvestigated factors concerning judgment and memory for differences of tint and so, at present, must be regarded as unreliable. It is likely, therefore, that the observer will obtain the most accurate results when he examines the unknown first on the right of and then on the left of the standard and takes the mean of the last observation when the unknown appeared strong whether it was viewed on the right or the left of the standard, and the first observation when the unknown similarly appeared weak as the correct match point. This is the procedure which was advised by Haldane.

(*c*) To obviate the above inaccuracies, an observer should discover standard conditions of illumination and distance which enable him to compare the tints accurately, and use these conditions whenever he is determining the amount of haemoglobin in blood.

## Summary

1. There was no difference in the accuracy with which ten trained and ten untrained observers matched the tints of solutions of carboxyhaemoglobin in daylight. The distance at which they made their comparisons was usually over 50 cm. and varied only within narrow limits in eighteen out of twenty observers.

2. The accuracy of seven out of ten untrained observers changed as the distance between their eyes and the comparison tubes changed.

3. Conditions of illumination and distance of comparison were found for seven out of ten untrained observers which enabled them to make accurate comparisons of tint.

4. The application of these observations to haemoglobinometry is discussed.

We are indebted to Professor Hartridge for his advice in the preparation of this paper, and to our observers for their co-operation.

### References

Behrens, B. (1929). *Arch. exp. Path. Pharmakol.*, **140**, 237.
D'Silva, J. L., and Turton, E. C. (In course of publication.)
Gaddum, J. H. (1933). Med. Res. Cncl. Sp. Rep. Ser., No. 183, London.
Haldane, J., and Smith, J. L. (1899–1900). *J. Physiol.*, **25**, 331.
Macfarlane, R. G. (1945). Med. Res. Cncl. Sp. Rep. Ser., No. 252. p. 59, London.