Absolute Treshold Measurements with the Diastereo Test

Dean H. W. Hofstetter, Ph.D., O.D., F.A.A.O., U.S.A.

Chairman: T. Standish Mason, F.B.O.A.

Dr. Hofstetter, Dean of the Division of Optometry in Indiana University, a distinguished American optometrist and educational administrator, is well known in Europe for his numerous articles and works published in optometric journals.

Introduction

Recently Pardon[1] described an unusually simple test for screening out persons who do not demonstrate binocular stereopsis. He was able to demonstrate a virtually absolute validity and reliability of separation of subjects with and without stereopsis. For the test distance of 5 to 6 feet, corresponding to a stereopsis angle range of 36 to 51 seconds, all subjects with stereopsis made 100% correct responses whereas all subjects without binocular vision failed to do so. Because the criterion for passing was 100% correct responses, the conventional or "50% accuracy" thresholds were not determined.

More recently Koetting and Mueller,[2] and later Reismann,[3] essentially duplicated Pardon's results on slightly modified versions of the test, in the sense that they were able to demonstrate complete response dichotomies separating those with binocular stereopsis from those without. This feature of the test prompted its identification as the "diastereo test".

The present study differs from the above in that an attempt is made to use the same type of test to explore the absolute threshold values among persons who have binocular stereopsis. The absolute threshold, also called the "50% accuracy" threshold, is the value at which the correct and incorrect response probabilities are equal. In the case of the diastereo test only one out of three possible responses is correct, when the "50% accuracy" threshold corresponds to "$66\frac{2}{3}\%$ correct responses". This relationship can be represented by the formula

$$3y = 2x + 100 \quad \dots \quad \dots \quad (1)$$

where $y = \%$ correct responses, and $x = \%$ accuracy.

Procedure

The diastereo test described by Pardon was further simplified for this investigation. The features are shown in Fig. 1. An ordinary Ray-O-Vac 2-cell flashlight was used, one that had a shield protruding forward from the edge of the transilluminated face. This shield served to prevent shadows from laterally located ambient light sources and it also provided protection for the protruding discs mounted on the transilluminated face. Two aluminium discs 0.5 mm. thick and 10 mm. in diameter were cemented in direct contact with the translucent plastic disc serving as the transilluminated face. A third aluminium disc of the same size was cemented on one end of a transparent plastic rod 9 mm. long and 6 mm. in diameter, the other end of which was cemented to the plastic translucent face. The three aluminium discs were arranged equidistant from the centre of the face and equidistant from each other, as shown in Fig. 1. Though the discs were in fact the grey colour of aluminium, they appeared black by reason of contrast when the flashlight was turned on.

To further diffuse the transilluminating light a sheet of thin white copy paper was placed behind the flat glass lens which, in turn, was directly behind the tranlucent plastic layer.

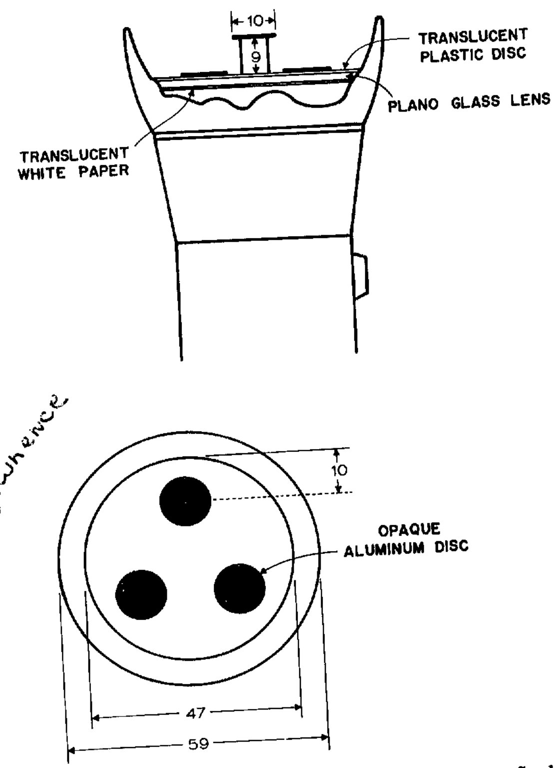

Fig. 1. Cutaway illustrating the diastereo test flashlight.

In the test procedure the examiner aimed the flashlight with one hand towards the subject's eyes and exposed the face of the flashlight for a period of one to two seconds by temporarily removing a large card held in front of the flashlight with the other hand. After each exposure the subject was asked to report which of the three discs, or spots, was nearest to him. Prior to each exposure the examiner rotated the flashlight randomly so that the protruding disc would be in one of eight positions, up, down, left, right, up and left, up and right, down and left, or down and right. Markers on the outside of the shield indicated these positions exclusively to the examiner. Ordinary but

Approved For Release 2003/05/14 : CIA-RDP78B05171A000600070046-7

consistent care in aiming the flashlight toward the subject's eyes during exposure seemed to be adequate to prevent the subject's use of any nonstereopsis clues. At any time that the examiner suspected the influence of any nonstereopsis clues he would check by having the subject cover one eye, a technique which invariably resulted in complete loss of stereo judgment, and thus assured the examiner that the correct binocular responses were in fact attributable to binocular clues.

The data for this report were collected by two high school seniors* on a group (I) of other high school seniors, and by a high school freshman† on a group (II) which included mostly children between ages 6 and 15 and several teachers.

In group I, 31 subjects were actually run through the test, but for this report only the data of 19 were used, those who showed acuity of 20/20 or better in each eye and binocular stereopsis. Of these 19, 13 wore glasses. Their ages ranged from 15 to 17, inclusive. Only three of the 19 were females.

Each subject in group I was given six exposures or trials at the test distance of 5 ft., six trials at 7 ft., and six each at 9, 11, 13, 15, 17, 19 and 21 ft., respectively. Then he was given six trials at 22 ft., six at 20 ft., and six each at 18, 16, 14, 12, 10, 8, and 6 ft., respectively. At each test distance the number of correct responses out of six trials was recorded without informing the subject as to the correctness of his answers.

The interpupillary distance of each subject was also measured. This ranged from 57 mm. to 66 mm., with a mean of 62 mm.

In group II, 45 subjects with binocular stereopsis were tested, but the record sheets for 24 of the subjects were inadvertently destroyed before all of the tallies and computations were completed, so that a part of the analysis of this group is based on all 45 subjects and a part on only 21 subjects. Only six of the 45, and two of the 21, wore glasses. The acuity was not measured, but the relatively high socio-economic level of the population for the school at which these tests were made and the high attention given to proper vision care in the school district strongly indicate that virtually all of the subjects in group II had 20/20 vision. Approximately half were males and half females.

Each subject in group II was given five exposures or trials successively at each of the test distances 6, 8, 10, 12, and 14 ft. The six adults in group II were tested also at 16 and 18 ft. At each test distance the number of correct responses out of five trials was recorded without informing the subject of the correctness of his answers.

Three subjects, aged 4, 8, and 9, who failed at six feet also failed at four and three feet and were not included among the 45 in group II. The reasons for their failure were not definitely ascertained, but there were indications that the 8 and 9 year-olds were squinters and that the 4 year-old did not understand the instructions.

For the purposes of this report all test distances were computed in seconds of stereopsis angle according to the following formula, in which the interpupillary distance is assumed to be 64 mm.:

$$\text{Stereopsis angle in seconds} = 1280/(\text{test distance in ft.})^2$$

Results

The results for group I are shown in the 19 individual graphs in Fig. 2. The ordinate values represent the number of correct responses out of the total of twelve trials at 5 and 6 feet averaged as 5.5 feet, at 7 and 8 feet averaged as 7.5 feet, etc. The abscissa is the log value of the seconds of stereopsis angle, whereby 5.5 feet becomes 1.63 log seconds, 7.5 feet becomes 1.36 log seconds, etc. The abscissa value in seconds is shown in the scale at the top of the figure.

The combined per cent of correct responses for the whole group at each test distance is shown in the curve in the upper right corner of Fig. 2. The dots represent the series at 5, 7, 8, ... 21 feet in that (receding) order of testing, while the circles represent the subsequent series at 22, 20, 18, ... 6 feet in that (aprpoaching) order of testing. The differences appear negligible and opposite to what might have been expected as a learning effect.

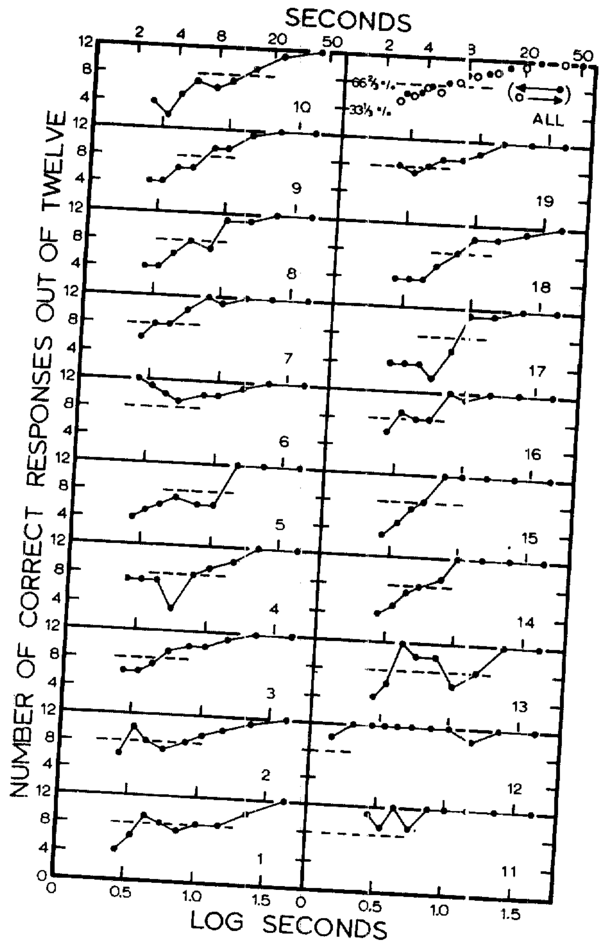

Fig. 2. Stereopsis test response curves for 19 subjects in group I, high school seniors. The horizontal dashed line represents the 66 2/3% correct response level, which corresponds to the 50% accuracy threshold level. For three subjects, Nos. 6, 11 and 12, the test was not carried out far enough to determine the 50% accuracy threshold level. The test distances are represented on the abscissa in log seconds of equivalent parrallactic angle.

In the upper right curve, the per cent correct responses are for all subjects at each test distance. The circles represent serially approaching (decreasing difficulty) tests; the dots represent serially receding (increasing difficulty) tests.

In all of the curves in Figs. 2 and 3, 33⅓%, or 4 correct responses out of 12, represent the frequency of correct responses when the binocular clues are totally inadequate; 100% would represent the frequency when the binocular clues are more than adequate; 66⅔%, or 8 out of 12, would indicate the absolute threshold of stereopsis at the 50% accuracy level, as computed from formula (1). In Figs. 2 and 3 the absolute threshold response level, 66⅔% (50% accuracy), is shown by a horizontal series of dashes in each graph; the intersection of this with the trend curve indicates the log second value of the absolute threshold on the abscissa scale. It is readily seen that the absolute threshold values for subjects number 6, 11, and 12 were smaller than that included within the maximum test distance of 22 feet. For number 12 a few trials were made at 25 and 30 feet without attaining the threshold.

The "log second" abscissa scale was adopted after considerable experimental plotting to find a scale which would give a normal increasing frequency of correct answers as represented in the theoretical curve in Fig. 3. Neither a "test distance" scale, a "stereopsis angle" scale, nor a "stereopsis angle reciprocal" scale gave the symmetry of Fig. 3 as faithfully as did the "log second" scale.

The distribution of threshold values for the 19 subjects in group I is shown in Fig. 4 on a rank scale. The lowest curve in Fig. 4 represents the log second equivalents of the greatest mean distance at which each subject gave eight correct responses in 12 successive trials. Since the tests were not carried out to this level of performance for three subjects, the curve starts with rank "4" and continues to rank "19". This representation of the stereopsis values on the ordinate in log seconds produced a curve which, though incomplete, closely resembles the theoretical curve of equal cumulative area intervals of a normal curve as shown in Fig. 5. From this it may be inferred that the designation of the stereopsis threshold in log seconds produces a normal distribution. The plotting of these thresholds on a "test distance", "stereopsis angle," or "stereopsis angle reciprocal" scale did not produce curves so nearly like the corresponding theoretical curve in Fig. 5.

The middle and upper curves in Fig. 4 are derived in the same way as the bottom curve except for the adoption of a higher criterion of passing. This permitted a ranking of all 19 subjects for the 100%, 12 correct responses out of 12, criterion, and all but the two best performers for the 83⅓% (75% accuracy threshold), 10 correct responses out of 12 criterion.

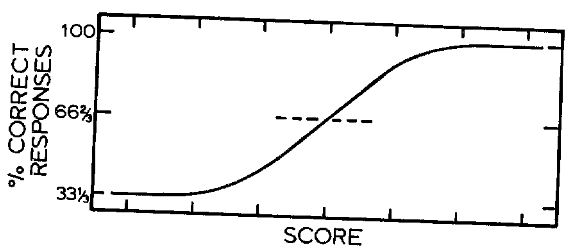

Fig. 3. Curve showing per cent correct responses for normally distributed hypothetical diastereo test score values of decreasing difficulty.

These two additional curves for the same group indicate that the lower ends of the curves have downward tails like that in Fig. 5. The upper end of the absolute threshold (66⅔% correct responses, or 50% accuracy) curve has an upswing like that in Fig. 5, as does also the 83⅓% correct responses, or 75% accuracy, curve, but this feature is not apparent in the 100% curve. This lack of upswing in the 100% curve could be a statistical artifact of the 100% criterion; it could be a clustering of the several poorest performers at a single level by reason of the large step to the next response level; or it could represent the invasion of a secondary clue at these poorer response levels. Whatever the explanation or significance of this feature, it is not eliminated by the choice of ordinate scale.

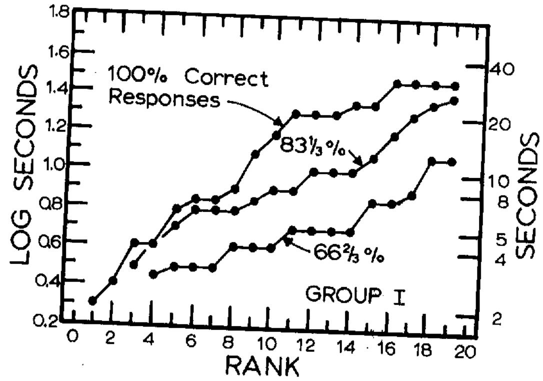

Fig. 4. Ranked stereopsis threshold values for 19 subjects in group 1.

The fact that the use of a log second scale results not only in a normal distribution of the responses for individual subjects as shown in Fig. 2, but also in a normal distribution of the threshold values for the group, as shown in Fig. 4, permits an evaluation of test reliability by conventional statistical methods. In order to incorporate the test results of all 19 subjects at the 66⅔% response level (50% accuracy level) in the computation of a reliability coefficient two such thresholds were derived for each subject, one from the series of receding test trials (5, 7, 9, ... 21 feet), and one from the series of approaching test trials (22, 20, 18, ... 6 feet). The threshold in each series for each subject was the greatest mean distance at which 12 correct responses were obtained in 18 trials. Thus, a subject who gave six correct responses out of six at both 20 and 22 feet could be considered as having given at least 12 correct responses out of 18 even if he gave all wrong responses in six trials at 24 feet (at which he was not tested), hence his threshold would be identified as 22 feet or 0.41 log seconds. Notwithstanding the imposition of such limitations for deriving threshold values, the product moment coefficient of correlation for reliability was 0.49±0.17 s.d. The scatterplot of these values is shown in Fig. 6.

A scatter plot of the threshold values for the subjects in group I against the interpupillary distances showed no apparent relationship, but the limited number of subjects does not exclude the possibility of such a correlation in a larger sample.

The distribution of threshold values for 21 subjects in group II is shown in Fig. 7 on a rank scale. The lowest curve (70% correct responses, or 55% accuracy threshold) represents the log second equivalents of the greatest mean distance at which each subject gave seven correct responses in 10 successive trials. Since the tests were not carried out to this level of performance for 10 of the 21 subjects, the curve starts with rank "11" and continues to rank "21". The representation of the

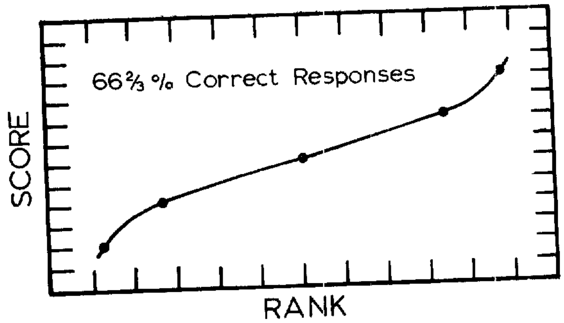

Fig. 5. Theoretical curve of normally distributed scores of a random sample of subjects plotted according to rank. The five dots on the curve represent the mean and the first and second standard deviations on either side of the mean.

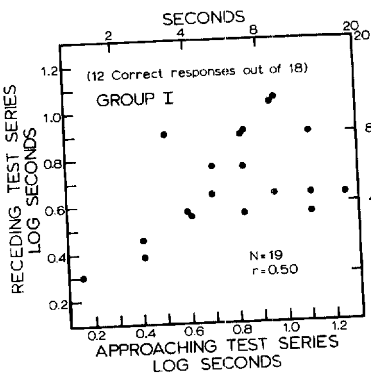

Fig. 6. Scatterplot of 50% accuracy threshold stereopsis values of 19 subjects in group 1.

stereopsis values on the ordinate in log seconds produced a curve which, though only half complete, closely resembles the theoretical curve of equal cumulative area intervals of a normal curve, as shown in Fig. 5. From this it may again be inferred, as for Fig. 4, that the designation of the stereopsis threshold in log seconds produces a normal distribution. Similarly, the plotting of these thresholds on a "test distance", "stereopsis angle", or "reciprocal of stereopsis angle" scale did not produce curves so nearly like the theoretical curve in Fig. 5.

The middle (80% correct answers, or 70% accuracy, threshold) and upper (100% correct answers, or 100% accuracy, threshold) curves in Fig. 7 are derived in the same way as the bottom curve except for the adoption of higher criteria of passing. The lower ends of these two curves clearly resemble the lower end of the theoretical curve in Fig. 5. The lack of an upswing at the upper end of the 100% curve corresponds to the same characteristic in Fig. 4.

The combined per cent of correct responses at each test distance is shown in Fig. 8 for each of three age subgroups of group II. The average of the six adults shows a 50% accuracy threshold of less than four seconds; the same threshold for the 20 teenagers is eight seconds, and the corresponding threshold for the 6 to 10 year olds is 11 seconds. It is noteworthy that the older teenagers in group I gave a corresponding

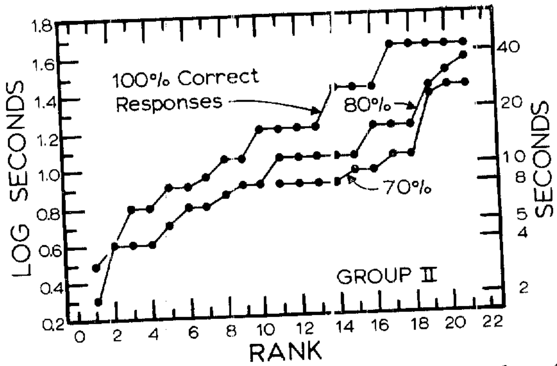

Fig. 7. Ranked stereopsis threshold values for 21 subjects in group II.

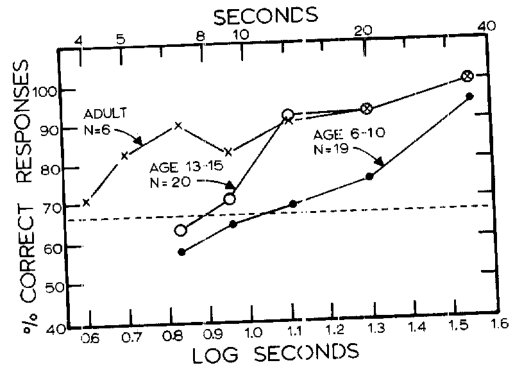

Fig. 8. Per cent correct responses at each test level for the group II subjects in each of three age subgroups. The horizontal dashed line represents the 50% accuracy threshold.

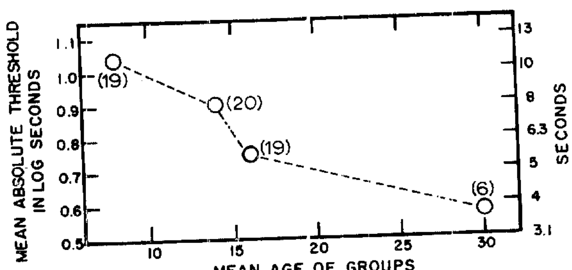

Fig. 9. Average 50% accuracy thresholds for each of four groups of subjects plotted against their average ages. The numbers in parentheses represent the number of subjects in each group.

mean threshold value of 5.5 seconds, as shown in the upper right curve of Fig. 2. This could have been interpolated quite accurately from the trends with age in group II. These average values for the four age groups are plotted in Fig. 9.

Discussion

To provide a basis for the comparison of these results with those of other investigators, it is possible to derive

Approved For Release 2003/05/14 : CIA-RDP78B05171A000600070046-7

a broad statement of the absolute threshold values for the whole group of subjects in this study by inspection of the bottom curves in Figs. 4 and 7. In Fig. 4 the bottom or 50% accuracy curve centers at about 0.6 log seconds and shows about two-thirds of the sample between 0.4 and 0.9 log seconds. In Fig. 7, the bottom curve, which can be presumed to be just slightly higher than a 50% accuracy curve, centers at about 0.8 log seconds and shows about two-thirds of the sample between about 0.4 and 1.1 log seconds. A combination of these two observations suggests a mean absolute threshold value of about 0.7 log seconds with a standard deviation of about 0.3 log seconds. This range, 0.7 ± 0.3 log seconds, would represent a mean of 5.0 seconds and a range from 2.5 to 10 seconds. The inclusion of two standard deviations from the mean would give a range in seconds from 1.3 to 20.

This range of results compares very favourably with the 2 to 4 seconds thresholds obtained by Berry[4] on three subjects. Howard[5] obtained a range of values between 1.8 and 7.3 seconds for 85 of his 106 subjects, while the other 21 showed a range from 10.6 to 136.2 seconds. Howard believed the latter poor scores to be attributable to physical factors interfering with the subject's vision, presumably inadequate visual acuity or absence of binocular vision.

The presently reported results also compare favourably with those of Bourdon[6] (5″), Crawley[7] (2.3″ and up), Anderson and Weymouth[8] (1.64″ and up), Früböse and Jaensch[9] (3.2″ to 6.6″), Langlands[10] (1.8″ to 7.3″), and Münster[11] (5″), all of whom carried out their testing in well-controlled laboratory settings.

The results obtained by the more typical screening techniques are not so impressive, however. Probably the most inclusive collection of such data are those of Sloan and Altman.[12] On both the standard and a modified Stereopter they obtained a continuum of scores on 68 subjects ranging from 10 seconds to 132 seconds, with modes at about 25″. These were based on a 7 out of 8 correct responses or 81% accuracy instead of 50% accuracy. On the Armed Forces Vision Tester they obtained a mode value of 16 seconds for 42 subjects with 40% of the subjects failing the easiest test plate, which represented a parallactic angle value of 39-41 seconds. Weymouth and Hirsch[13] obtained similarly high thresholds for a large share of the 65 subjects on a telebinocular stereopsis test. Even the "100%" performance level on the scales devised by Shepard and Fry[14] for use with stereoscope test slides represents 16 seconds of parallactic angle.

It is apparent that the diastereo test, even when used as a rapid screening instrument, measures stereopsis at a much more critical threshold level.

The matter of scaling stereopsis scores does not seem to have been given very analytical treatment except that skewness of typical data has been pointed out by Weymouth and Hirsch[15] who represented their data in relation to separation and/or parallactic angle thresholds. Similar skewing can be observed in virtually all published data, whether they are the frequency of correct response data on a single subject, as in Figs. 2 and 8, or the rank distribution of threshold values in a group of subjects as in Figs. 4 and 7. The transformation of such data to log second scales show substantial if not virtually complete elimination of skewness in the data of Howard[5], Crawley[7], Anderson and Weymouth[8], Langlands[10], Sloan and Altman[12], and Hirsch and Weymouth[15].

Such skewness appears to have prevented meaningful statistical correlation computation, although Weymouth and Hirsch[15] did attempt to derive correlation coefficients for some of their samples by omitting extreme scores. By this technique they derived reliability coefficients from which they concluded that, ". . . . the less-time-consuming rod-test (Howard-Dolman) and the telebinocular test are unreliable and invalid, respectively . . . . ". In the same vein Sloan and Altman [12] reported for the Howard-Dolman and the Stereopter test that. "The data suggest, however, that within the group showing good depth perception there is no close agreement in relative ranking on the two tests". Unfortunately, the data from both reports are not presented in raw form and so do not lend themselves to re-evaluation on a transformed log scale as was done in the present study (Fig. 6) showing a test-retest reliability coefficient of 0.5 for a group of 19 subjects all of whom showed good scores.

The indication of improved stereopsis with age appears to be practically uninvestigated. Tiffen[16,17] showed an increasing percentage of passing of a stereopsis test among adults up to about the age of 40. Twenty subjects in Crawley's[7] report, ranging in age from 4 to 70, showed an average of about 10 seconds around age 8 and a decrease to about 4 seconds at age 35. It is quite possible that the apparent agreement of these two reports with the present data is purely fortuitous, but it certainly justifies further investigation.

No published data showing a statistical relationship between interpupillary distance and stereopsis have been found. The theory that larger interpupillary distances should give better stereopsis scores is not confirmed in the presently reported data. Neither is the large apparent increase of stereopsis with age quantitatively attributable to an increase in interpupillary distance with age. Rather, these results suggest that a continuous stereopsis learning process may be involved, right up to full adulthood.

Summary

Disastereo test thresholds were determined on two groups of subjects, one a group of 31 high school students and the other a group of 45 subjects of a substantial range of ages, mostly children, all of whom had binocular vision. The two groups gave mean threshold values of 0.7 log seconds (5 seconds) and a standard deviation of ± 0.3 log seconds, representing a standard deviation range from 2.5 to 10 seconds of parallactic angle. The test-retest coefficient of reliability for one group was 0.5. The stereopsis scores showed no apparent trend with the interpupillary distances, but they showed a marked improvement with the increase of age into adulthood. The sample was not large enough to establish the statistical significance of the latter relationship.

Analysis of the data in terms of the relative frequency of correct responses about the absolute threshold and in terms of the distribution of individual subjects threshold values clearly indicate the justification of a log second scale to represent stereopsis data. In other words, the log second scale produces the distribution characteristics of normal data and so permits the application of conventional statistical correlation formula. A review of previously reported stereopsis data supports the log second technique.

The diastereo test, though simple and quick in application, gives results fully comparable with the best stereopsis data previously reported for rigorous and time consuming laboratory techniques. The diastereo test results reported here appear substantially more valid and more reliable than those reported for other popular stereopsis screening instruments.

BIBLIOGRAPHY

[1] Pardon, H. R., *A New Testing Device for Stereopsis*, Journal of the American Optometric Association, Vol. 33, No. 7, February, 1962, pp. 510-512.

[2] Koetting, R. A., and R. C. Mueller, *Evaluation of a Rapid Stereopsis Test*, American Journal of Optometry and Archives of American Academy of Optometry, Vol. 39, No. 6, June, 1962, pp. 229-303.

[3] Reisman, M., *Evaluation of a Modified Stereopsis Test*, submitted to the Indiana Journal of Optometry for publication in 1965.

[4] Berry , R. N., *Quantitative Relations Among Vernier, Real Depth, and Stereoscopic Depth Acuities*, Journal of Experimental Psychology, Vol. 38, No. 6, Dec., 1948, pp. 708-721.

[5] Howard, H. J., *A Test for the Judgment of Distance*, The Optician and Scientific Instrument Maker, Vol. 63, No. 1505, Jan. 30, 1920, pp. 285-289, and No. 15/6, Feb. 6, 1920, pp. 297-299.

[6] Bourdon, B., *La Perception Visuell de L'espace*, Schleicher Freres, Paris, 1902.

[7] Crawley, C. W. S., *Stereoscopic Vision*, British Journal of Photography, Vol. 52, No. 2353, pp. 446-447.

[8] Anderson, E. E., and F. W. Weymouth, *Visual Perception and the Retinal Mosaic*, American Journal of Physiology, Vol. 64, No. 3, May, 1923, pp. 561-594.

[9] Frübose, A., and P. A. Jaersch, *Die Einfluss Verschiedener Faktoren auf die Tiefensehscharfe*, *Zeitschrift für Biologie*, Vol. 78, Ncs. 3-4, April, 1923, pp. 119-132.

[10] Langlands, N. M. S., *Experiments on Binocular Vision*, Transactions of the Optical Society (London), Vol. 28, No. 2, 1926-27, pp. 45-82.

[11] Münster, C., *Uber einige Probleme der Stereoskopischen Messung*, Zeitschrift für Instrumentenkunde, Vol. 62, November, 1942, pp. 346-357.

[12] Slcan, L. L., and A. Altman, *Factors Involved in Several Tests of Binocular Depth Perception*, American Medical Association Archives of Ophthalmology, Vol. 52, October, 1954, pp. 524-544.

[13] Weymouth, F. W., and M. J. Hirsch, *The Reliability of Certain Tests for Determining Distance Discrimination*, American Journal of Psychology, Vol. 58, No. 3, July, 1945, pp. 379-390.

[14] Fry, G. A., *Measurement of the Threshold of Stereopsis*, Optometric Weekly, Vol. 33, No. 37, October 22, 1942, pp. 1029-1032.

[15] Hirsch, M. J., and F W. Weymouth, *Distance Discrimination*, Archives of Ophthalmology, Vol. 39, No. 2, Feb., 1948, pp. 210-231.

[16] Hofstetter, H. W., *Industrial Vision*, Chilton Co., Philadelphia, 1956, pp. 77-79.

[17] Tiffin, J., *Industrial Psychology*, Prentice-Hall, Mc., New York, 1942.

Transactions, Dublin International Optical Congress 1965