MRL Report # 97

COMPARATIVE STUDY OF MEASURES OF HETEROPHORIA

by

Lt. Comdr. John H. Sulzman, (MC), USNR
Lt. Comdr. Ellsworth B. Cook, H(S), USNR
Lt. (jg) Neil R. Bartlett, H(S), USNR

Progress Report No. 3
on
Bureau of Medicine and Surgery
Research Project No. X-493 (Av-263-p)

22 February 1946

APPROVED: Captain C. W. Shilling, (MC), U. S. NAVY, MC-in-C.

48

COMPARATIVE STUDY OF MEASURES OF HETEROPHORIA

by

Lt. Comdr. John H. Sulzman, (MC), USNR
Lt. Comdr. Ellsworth B. Cook, H(S), USNR
Lt. (jg) Neil R. Bartlett, H(S), USNR

Progress Report No. 3.

Bureau of Medicine and Surgery

Research Project No. X-493 (AV-263-p)

"Comparison of Various Screening Devices
With Standard Medical Procedures."

Medical Research Department
U. S. Submarine Base
New London, Connecticut.

22 February 1946

Acknowledgments

The authors wish to emphasize that the findings in this report were possible through the wholehearted cooperation of many observers and test operators. They also desire to express their thanks to Chief Pharmacist's Mate Robert E. L'Etoile for attending to the endless details of conducting the experiment, and to Marian B. Elliott for supervising the statistical computations. The authors note that it was possible to submit this report on schedule because of the untiring competence of Yeoman David Fruits in typing the many drafts of this manuscript.

GENERAL SUMMARY

1. Certain criteria for a valid measure of heterophoria are presented.

2. Analysis of the results obtained by measuring 100 subjects by three clinical tests indicates that the Maddox rod method for lateral heterophoria meets these criteria adequately.

3. Three visual screening devices were used to measure heterophoria on 121 subjects who were also tested by the Maddox rod method.

4. A comparison of the data from the two experiments suggests that for the measurement at the near point, the angular deviation of the eyes from the horizontal position is apparently critical. The results obtained when the eyes are in the position usually employed in reading are more consistent than when the near measurement is made with the gaze directed horizontally forward.

5. The data from the two experiments suggests that clinical measurements of heterophoria are more consistent when the examiner adjusts the phorometer prisms than when the subject makes the adjustment.

6. There does not seem to be any justification for the use of arbitrary designations for heterophoria scores in non-clinical terms. Such use of coded digits complicates any statistical analysis of data obtained from visual screening devices.

7. A comparison of the measurements obtained by the Maddox rod method with those obtained by the use of visual screening devices indicates:

a. That visual screening devices are apparently as reliable as the clinical method for the measurement of heterophoria.

b. That the instrument tests do not correlate well with the clinical test; it is possible that this is due to some defect in construction of the devices or in the method of measurement.

c. That further research is indicated to establish why there is no more correspondence between the clinical and instrument measures.

8. The procedures described in the "Manual of Instructions Heterophoria Testing" of the Army-Navy-NRC Vision Committee are satisfactory for service use; a change in the procedure is proposed in paragraph five above.

COMPARATIVE STUDY OF MEASURES OF HETEROPHORIA

General Introduction

The present report summarizes the experimental studies of muscle balance which were indicated from earlier experiments to which reference will be made.

In addition to pursuing certain lines of investigation, it was desired to evaluate the clinical method of phoria measurement described in the "Heterophoria Testing Manual of Instructions" (1). This Manual was developed by the Subcommittee on Procedures and Standards for Visual Examinations of the Army-Navy-OSRD Vision Committee.

Some of the references published between 1939 and 1944 will be found in the "Bibliography of Visual Literature" (2) compiled with the assistance of members of this activity. Additional information regarding the terms used herein will be found in Progress Report No. 1 to which reference will be made, and in the Minutes of the Twelfth Meeting of the Vision Committee (3). Another discussion of the subject will be found in the Minutes of the Fourteenth Meeting of the Vision Committee (4). The latter paper contains much of the preliminary experimental work upon which the present study is based.

- - - - - - - -

There has been a great deal written on the function of the extra-ocular muscles. Anyone who makes an effort to review the literature is overwhelmed by the number of published articles and by the great variety of tests advocated. Indeed, there are so many methods for measuring heterophoria that the impartial investigator might conclude that no one test has outstanding merit. To put it another way, one can find published authority for almost any viewpoint on this important subject. As Cridland (5) has expressed the matter:

"The diversity of the available methods is only equalled by the diversion which some of them engender. The fantastic elaborations of the instrument maker have introduced further complexity, so that the patient, when faced by an imposing mass of machinery, may be easily forgiven for his failure to relax, and the element of voluntary effort becomes a thunder-cloud obscuring the entire heterophoric horizon. The wonder is that the results of measurement by different tests are not even more contradictory than they are."

The present paper is justified only if some necessary clarification of the current confusion is achieved.

In the interests of brevity, reference to specific authorities will be held to a minimum. The investigator who desires to pursue any particular phase will find an abbreviated bibliography appended.

As has been indicated, the number of available tests is extensive. A paper by Wheeler (6) reviews the measures which were first developed. Some of the early investigators devised methods which are worth study, and the principles laid down at the beginning have changed very little. It may be stated that, at the present time, a majority of ophthalmologists employ either the cover test and its modifications, the Maddox rod test in some form, or a combination of these.

The problem under consideration differs from that of the practicing civilian eye specialist. These clinicians have the advantage of much experience with all degrees of ocual motility, including the obviously abnormal. However, the problem which confronts the military services is rarely one of differential diagnosis of ocular muscle conditions. The main objective is the screening of large numbers of men in order to reject those who are unfit. Additional requirements for a measure which is desirable from the military standpoint will be discussed later.

In the present experiment no attempt has been made to evaluate all of the available measures of heterophoria. Such an encyclopedic study would present a monumental task and certainly would not be worth the effort. The limitation can be indicated in titles for the two parts into which this paper is divided. These are as follows:

Part I, Comparison of three clinical tests for lateral heterophoria: Maddox rod, screen-Maddox rod, and screen and parallax.

Part II, Comparison of the Maddox rod test for lateral and vertical phoria, with phoria tests incorporated in three visual screening devices: Keystone "Telebinocular", Bausch & Lomb "Ortho-Rater" and the American Optical "Sight Screener".

Part I.

COMPARISON OF THREE CLINICAL TESTS FOR LATERAL HETEROPHORIA

Introduction

Mention has already been made of the general categories into which the most accepted measures of heterophoria can be grouped. Two of the techniques evaluated in this report are in common use by ophthalmologists generally (5)(7)(8)(9) but with some qualifications by individual writers (10)(11)(12)(13)(14). These are: (1) the Maddox rod and (2) the screen and parallax (cover test). A third method, (3) the screen-Maddox rod, is prescribed for the aviation medical examination for the Army and Navy (15)(16).

It might be stated parenthetically that tests which employ any type of stereoscopic instrument have been criticised by various authorities (13). One of the most common objections has been stated by Verhoeff (14) as follows:

> "Tests for presumptive heterophoria made by use of a stereoscope of any kind are especially unreliable because the observer has no accurate idea of the distances concerned."

In order to appreciate fully the extremes of viewpoint, one writer (17) assumed that occlusion of one eye for a week is necessary to elicit all the presumptive heterophoria: Fortunately, this opinion has not been entertained seriously by military surgeons.

Out of such a variety of usage and opinion, the experimental work of Cridland (5) and of Scobee (4) are outstanding as probably the first attempts to evaluate several methods of measuring heterophoria by assessing them on the basis of statistical evidence and analysis.

From the evidence of experimental studies, Captain Scobee recommended:

1. That since the correlation of the Maddox rod test with the screen and parallax test is as high as the correlation of the screen and parallax test with the screen-Maddox rod test, the Maddox rod test may be used for measuring heterophoria.

2. That the Maddox rod rest is easier to standardize and should be substituted for the more complicated screen-Maddox rod method.

3. That the factors of placing the Maddox rod before the dominant eye and the brightness of the room do nct affect the measures elicited.

4. That the Maddox rod should be white instead of red.

5. That phorias should be measured at distances of 20 feet and 13 inches.

In the light of Scobee's studies, it was decided that before visual screening devices could be evaluated for their competence to measure heterophoria, three of the most acceptable clinical methods for this purpose should be compared. The most promising clinical test then might be applied as a validation criterion for the comparison of instruments.

Procedure

One hundred subjects were chosen at random from among the candidates for the Submarine School. Groups of ten men were tested twice during the same half-day, each time at distances of 20 feet and 13 inches, using the following order in each case:

1. Maddox rod (white)
2. Screen-Maddox rod (white)
3. Screen and parallax.

All measurements were made by one examiner in a lighted room. In each case, a one c.m. light was viewed at a distance of twenty feet without glasses and the examiner adjusted the Risley prisms of the phorometer. Since the vertical deviations were not being measured, the Stevens phorometer attachment was not used. A procedure wherein the examiner adjusted the prisms was selected in order to compare such measures with those to be performed later wherein the subject adjusted the prisms. Prior to the test the phorometer frames were accurately centered and levelled and the white Maddox rod was rotated into position before the right eye.

The first measure was made with the Risley prism in place behind the Maddox rod, the prism being set so that the subject perceived a vertical white streak of light to the left of a white light. It was demonstrated to the subject that the streak could be moved laterally, and he was directed to inform the examiner when the streak lay directly through the light. Next, the test was repeated while the right eye was covered intermittantly, thus adding the screening factor to the Maddox rod test. The reading for the second test was likewise the prism position at which the subject reported that the line covered the light. The third test was performed after the Maddox rod had been rotated out of the line of view so that the subject viewed the light with each eye; a cover was placed before each eye alternately while the prism was adjusted until the subject

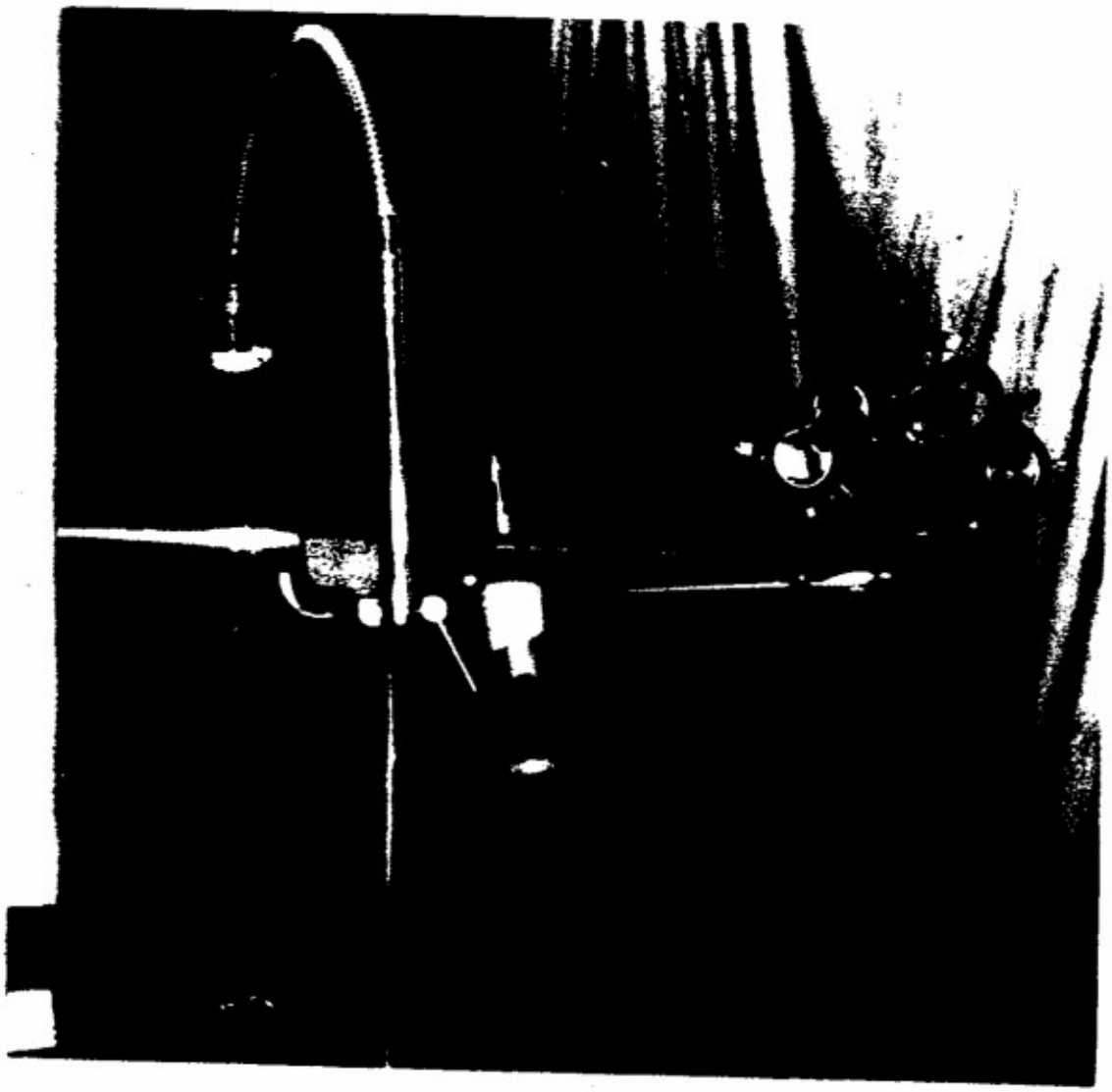


Fig. 1.

Ophthalmoscope light attached to phorometer for Maddox rod measure at 13 inches.

reported no apparent lateral movement of the light. The prism position was then taken as the third reading.

At the completion of the above tests at the twenty foot distance, the light was occluded and the subject's gaze was directed at the bulb of a lighted ophthalmoscope at a distance of thirteen inches from his eyes, exactly in the mid-line and at eye level. The ophthalmoscope handle was secured by means of tape to the bar attached to the phorometer (Figure 1.) and swung into position after the distance testing had been completed. The same three tests were then performed at the near point, in the corresponding order and in exactly the same manner.

At the completion of the test-series, the data was plotted on scattergrams and subjected to statistical analysis.

Results

The test-retest consistency data are summarized in Table I and include product-moment reliability coefficients*, and test and retest means with standard deviations. Except for the first column of figures, the values given are prism-diopters of lateral heterophoria. The letter "E" indicates esophoria, "X" indicates exophoria.

- -

*The coefficient of correlation is a value, which in the mathematical theory of probability, serves to measure the degree of correlation between two attributes or measures of a group of individuals. When the value is zero, there is no correlation, and when the value is one, the correlation is perfect.

TABLE I.

Test - Retest Consistency

A. Lateral heterophoria at twenty feet.

Method	Correlation Coefficient	Test Mean	Test S.D.	Retest Mean	Retest S.D.	Standard error of retest estimate
Maddox rod	0.87	E .88	2.295	E .88	2.368	1.18
Screen-Maddox rod	0.89	E 1.37	2.784	E 1.30	3.106	1.43
Screen and parallax	0.72	E 1.42	2.741	E 0.47	2.579	1.79

B. Lateral heterophoria at thirteen inches.

Maddox rod	0.74	X 3.87	4.50	X 4.99	4.896	3.33
Screen-Maddox rod	0.82	X 4.24	4.984	X 5.40	5.654	2.25
Screen and parallax	0.67	X 6.06	6.782	X 7.71	5.638	4.14

The data for test intercorrelations are presented in Table II. The values are the product-moment reliability coefficients for each pair of tests.

TABLE II.

Test Intercorrelations

A. Lateral heterophoria at twenty feet.

	Maddox rod	Screen-Maddox rod	Screen and parallax
Maddox rod	0.87	0.92	0.74
Screen-Maddox rod		0.89	0.85
Screen and parallax			0.72

B. Lateral heterophoria at thirteen inches

Maddox rod	0.74	0.89	0.70
Screen-Maddox rod		0.82	0.76
Screen and parallax			0.67

TABLE III

Table III presents the distributions of scores for test and retest by each of the three methods. Each vertical column signifies one prism-diopter of lateral heterophoria, with separate scales used for the two distances. The values in each column indicate the number of individuals found to be in each step in the scale.

Distributions of Scores for Lateral Heterophoria

A. Twenty feet Distance

Prism diopters	Esophoria																					Exophoria							
	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7
Maddox rod																														
Test									1	0	0	0	0	0	0	1	0	2	4	8	16	16	34	12	4	0	1	0	0	1
Retest									1	●	0	0	0	0	1	0	1	2	2	5	20	20	31	12	1	1	1	1	0	1
Screen-Maddox rod																														
Test					1	0	0	0	0	0	0	0	0	0	0	1	0	2	4	6	8	17	23	21	11	4	0	1	1	
Retest	1	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	1	2	5	9	15	21	23	15	3	0	1	0	0	1
Screen and parallax																														
Test					1	0	0	0	0	0	0	0	0	0	0	0	2	5	14	4	21	15	15	10	11	2				
Retest							1	0	0	0	0	0	0	0	1	0	0	3	5	10	15	14	11	15	14	3	6	2		

B. Thirteen inches Distance

Prism Diopters	Esophoria										Exophoria																			
	10	9	8	7	6	5	4	3	2	1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Maddox rod																														
Test					1	2	4	2	8	2	18	6	8	10	6	7	7	2	2	6	3	3	0	1	1	0	1			
Retest					1	2	2	2	6	7	3	8	12	8	6	8	6	4	9	3	2	4	2	1	0	1	2	1		
Screen-Maddox rod																														
Test				2	2	0	1	5	6	6	13	10	4	8	10	1	2	6	7	4	4	0	5	3	0	1				
Retest	1	0	0	0	0	2	4	3	3	6	8	3	11	9	8	3	7	5	3	3	5	4	2	2	2	2	1	2	1	
Screen and parallax																														
Test		1	1	0	1	1	1	2	5	6	6	5	4	7	12	4	9	0	7	4	6	3	3	2	4	0	3	1	2	
Retest					1	0	3	2	3	3	3	3	4	3	12	4	7	8	10	3	9	1	2	6	5	2	0	1	4	1

Discussion

No discussion of measures for heterophoria is complete unless the inherent variability of this function is emphasized. One authority (18) states the situation as follows:

> "If heterophoria tests are repeated several times under apparently equal conditions, the results obtained in those tests will very frequently differ, and this not only with respect to the amount but occasionally also with respect to the kind of existing heterophoria."

Since retesting heterophoria yields results of uncertain variation, a coefficient of test-retest correlation is not the only criterion of consistency which should be applied. The differences between test and retest means must be evaluated, and the reliability coefficient must be considered in the light of relative variabilities.

From the data presented in Table I above, the Maddox rod test compares favorably with the other two methods at both distances. One sees from the relative standard errors of estimate that the results of a second distance examination, with the Maddox rod can be predicted after a first examination, with more precision than a second screen-Maddox rod from a first screen-Maddox rod test. Note that it is assumed that one test of muscle balance does not affect the score for a later test; the order of testing was not altered from one individual to the next. There is some indication from the means in Table I that this assumption is not valid for the near tests.

In Table II, one notes that the intercorrelations are as high as the inherent reliability in each measure allows; in other words, insofar as one can judge from these data, all three tests may be measuring identicl functions.

Another aspect of fundamental importance must be stressed; that is, the question of what a method measures. There is a distinction between methods which employ simultaneous binocular fixation, such as the Maddox rod, and those in which the amount of ocular deviation is measured when binocular fixation is interrupted, as in covering each eye alternately. The screen-Maddox rod and the screen and parallax tests are examples of the latter type. In other words, the validity of measurement is critical. Unfortunately, there is as yet no standard which can be recognized as generally valid.

The selection of one test over others as the best measure of heterophoria is possible then, only when the method selected demonstrates clear-cut superiority in respects agreed upon previously as crucial. While the desirable qualities for a valid measure may not have been agreed upon universally, or even achieved, this question must be addressed by those in ophthalmological research. However, for the present needs of the military services, if not for clinicians in civilian practice, it may be assumed that the best test is one in which experimental studies indicate satisfactory sensitivity, a mean and range which are in agreement with ophthalmological experience, and an adequate reliability or consistency of measurement. In addition, the method of choice should employ the type of ocular innervation which differs least both in degree and in kind from that normally employed in binocular fixation. Finally, the preferred method of measuring heterophoria should be at once as simple as possible, and at the same time capable of standard administration.

Summary

1. The findings of Seabee, and the results of the experiment which are presented above, recommend the Maddox rod test as a military measure of lateral heterophoria which is as satisfactory as the other two methods. Its advantages are:

A. The consistency of measurement is at least as great as that with the other tests studied.

B. The distribution of scores indicates that small deviations can be elicited over a wide range of measurements.

C. The correspondence between means and variability measures for first and second examinations compares favorably with the correspondence found with two other clinical measures.

D. Standardization of technique can be achieved since the method of administration is simple and there is a minimum of interference by the examiner.

E. The amount of time required for explanation, demonstration and testing is minimal.

2. The reliability of each of the lateral heterophoria measures at twenty feet exceeds its reliability at thirteen inches.

Conclusion

The Maddox rod test is as satisfactory a measure for lateral heterophoria as either the screen-Maddox rod or the screen and parallax tests. Accordingly, the Maddox rod test is accepted as a standard clinical measure for evaluating the visual screening devices studied in Part II below.

PART II.

COMPARISON OF THE MADDOX ROD TEST FOR LATERAL AND VERTICAL PHORIA WITH PHORIA TESTS INCORPORATED IN THREE VISUAL SCREENING DEVICES (Keystone "Telebinocular", American Optical "Sight Screener" and Bausch & Lomb "Ortho-Rater")

Introduction

Data from a previous study (19) which was also reported to the Army-Navy-CSRD Vision Committee (20) indicated that visual screening devices might be acceptable for ocular tests in routine physical examinations of military personnel.

The results reported in the first portion of the present experiment are in agreement with those reported by Scobee - that the Maddox rod test for lateral heterophoria is an adequate measure of horizontal ocular muscle imbalance. That test takes advantage of the fact that the eyes will superimpose two relatively different uniocular impressions into a single binocular perception.

Duane (7) makes a distinction between tests which employ alternating "Binocular fixation", such as the screen-Maddox rod and the screen and parallax (cover) tests, and those which involve "binocular vision". The Maddox rod test requires simultaneous macular perception of relatively dissimilar objects (a light streak and a dot of light) and while the screening devices involve heterophoria targets and a method of presentation that are different from those in the Maddox rod method, the basic principle is essentially the same. The similarity is stressed by redefining the concept "heterophoria"; from an operational standpoint it is the deviation of the eyes which is overcome and rendered latent when fusion is permitted (21).

Description of Tests

The Maddox rod test displays a light to the left eye and a streak of light to the right eye. On the other hand, the phoria targets in the instrument tests display arrows or lines for one eye and dots or steps for the other eye (Figs. 2, 3, 4, 5, 6, 7).

The Ortho-Rater incorporates an additional feature, in that while the right eye views a row of dots on the targets for lateral phoria, the left eye is presented with an arrow pointing to the middle one of three horizontal dots. The dots seen by the left eye fuse with those seen by the right eye and thus tend to minimize shifting of the arrow laterally. Minimizing the lateral swinging in this fashion is presumed to make the measure more stable.

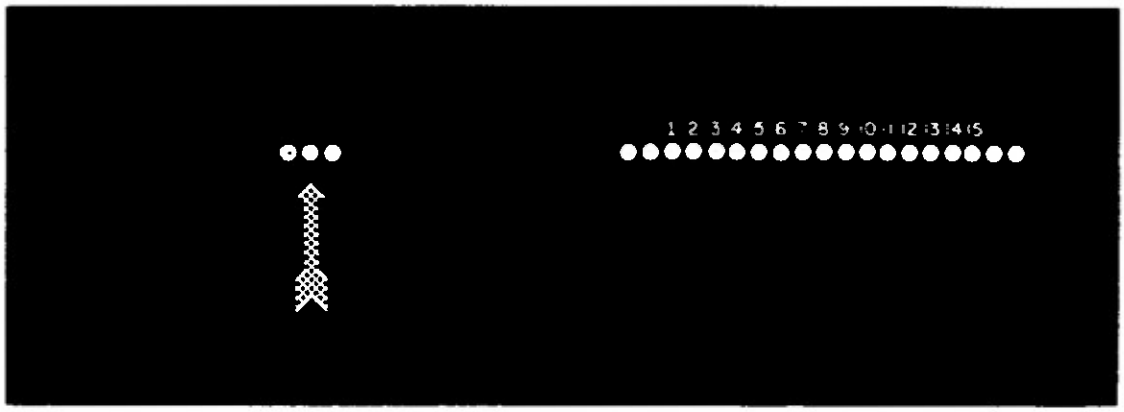
Another difference in the instruments lies in the illumination. The Telebinocular targets are illuminated from the front, while the Sight Screener and Ortho-Rater targets are transilluminated from behind. In addition, the Sight Screener is unique in that the Polaroid vectograph principle is employed for presenting individual targets to each eye.

Decentered convex lenses are a part of the optical system for the equivalent of the 20 foot distance in each of the instruments. In the Sight Screener, however, the targets for near are viewed at an actual 14 inch distance without the interposition of prismatic lenses.

Procedure

The present experiment was an integral part of the procedures described in Progress Report No. 2 (22) which has been submitted for approval to the Bureau of Medicine and Surgery. For details of the order of testing, the conditions which obtained, and the general conduct of the test, reference should be made to that report.

Test F-1 (Phoria) Vertical

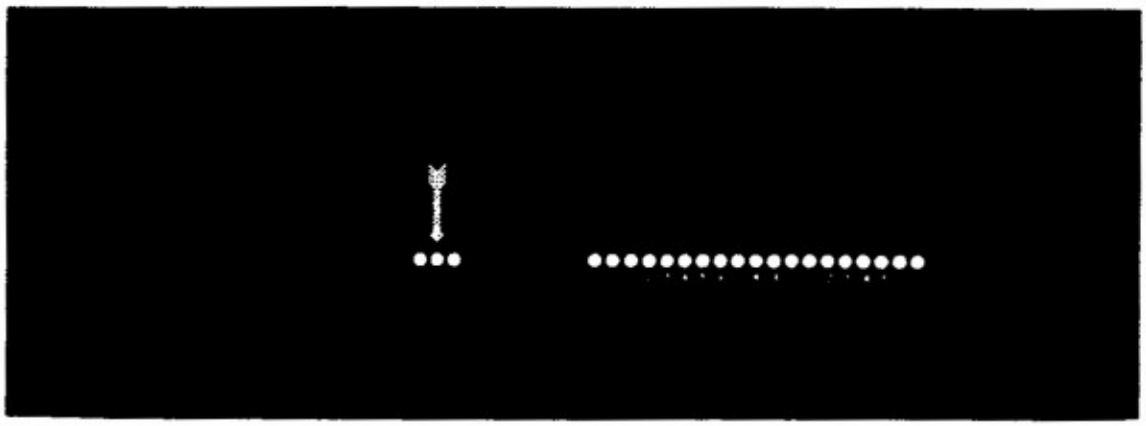


Test F-2 (Phoria) Lateral

Fig. 2.

Ortho-Rater targets for vertical and lateral phoria at the optical equivalent of 20 feet.

Test N-4 (Phoria) Vertical



Test N-5 (Phoria) Lateral

Fig. 3.

Ortho-Rater targets for vertical and lateral phoria at the optical equivalent of 13 inches.

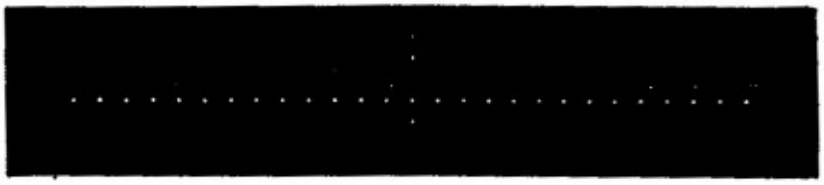


Fig. 4.

Sight Screener target for Lateral Phoria at both the optical equivalent of 20 feet and at an actual 14 inches.

Fig. 5.

Sight Screener target for Vertical Phoria at both the optical equivalent of 20 feet and at an actual 14 inches.

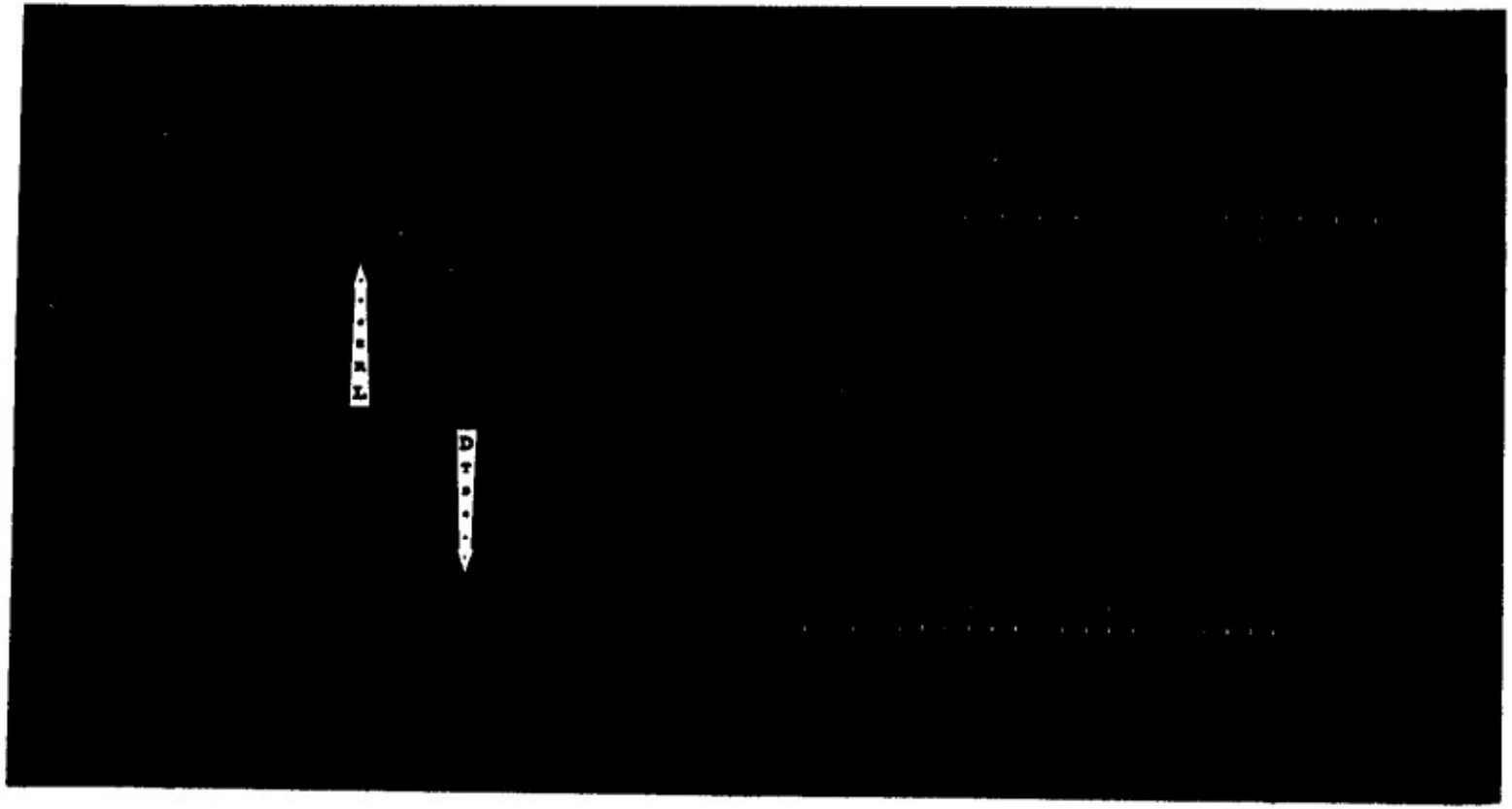


Fig. 6.

Telebinocular target for lateral phoria
for the optical equivalent of 20 feet
(upper scale) and 13 inches (lower scale).

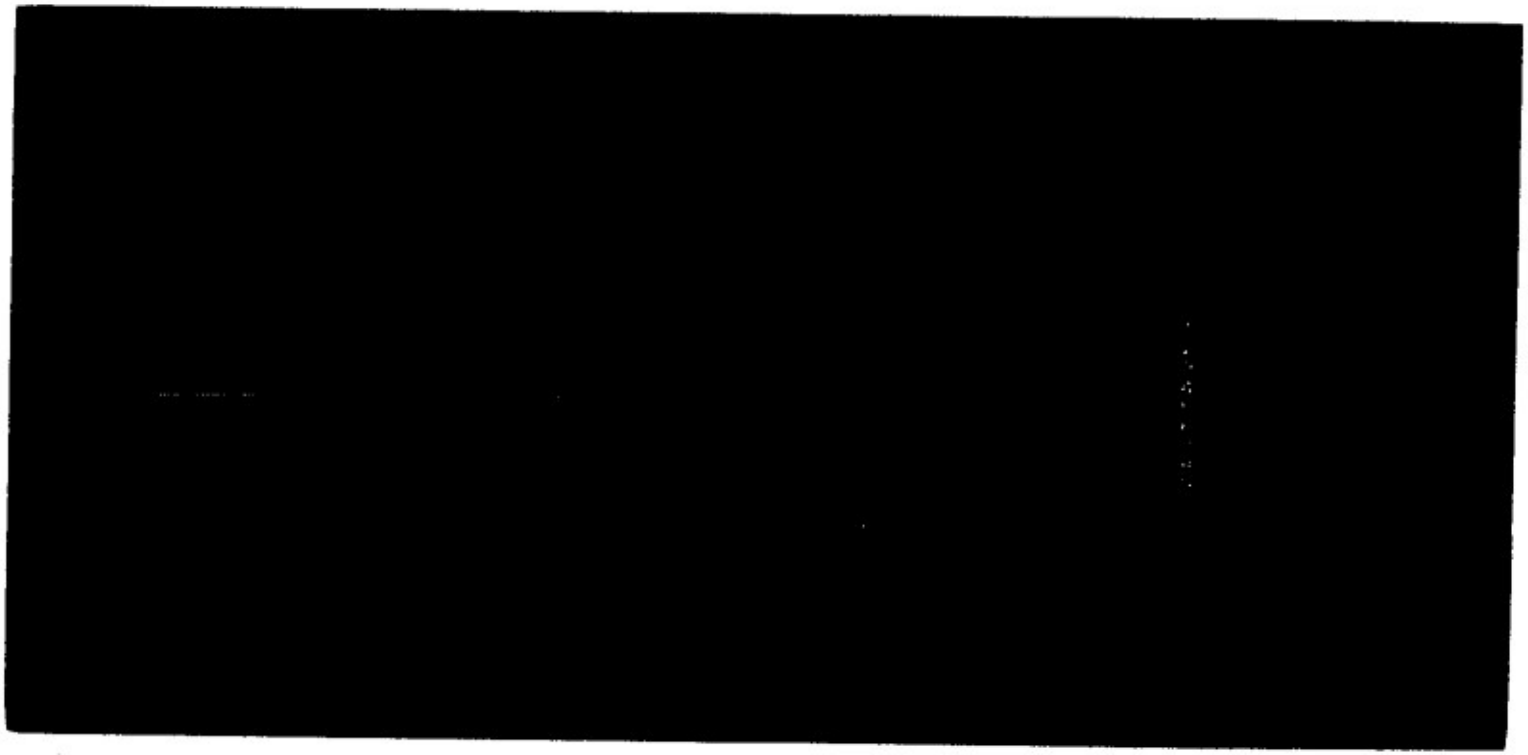


Fig. 7.

Telebinocular target for vertical phoria

at the optical equivalent of 20 feet.

In recapitulation, it may be stated that the present experiment involved the measuring for ocular muscle imbalance of 121 individuals without eye glasses. For the Maddox rod portion of the test, the procedure described in the "Manual of Instructions for Testing Heterophoria" (1) were followed exactly. This method employs the Maddox rod, and differs from that reported in the first part of this paper in that the subject adjusts the Risley prisms of the phorometer himself, and the near target is depressed to a level of 6 inches below the level of the subject's eyes at the 13 inch distance. This latter change has been recommended in order to approximate the visual axes of the eyes more closely to the reading position, instead of looking directly ahead for the near point measurement.

For the Maddox rod test, the wall brightness was measured by a Macbeth illuminometer at eye level as 0.5 apparent foot candles. Wall brightness around the 20 foot light from the observer's position was 0.2 apparent foot candles.

In measuring lateral heterophoria with the Maddox rod, the streak of light is placed on the exophoria side of the zero mark on the Risley prism scale before the measurement is begun. The instructions to the subject as to how he would adjust the prisms were exactly as prescribed in the Manual. One objective of the experiment was, after all, an assessment of a simple method requiring a minimum of instruction and supervision. In most cases the subject rotated the prisms during the course of the measurement without pausing or reversing the direction of movement until he indicated that the end-point had been reached. In some cases, however, a subject would reverse the direction of prism movement spontaneously one or more times before the measurement was completed.

The procedure described in the Bausch and Lomb Ortho-Rater "Standard Practice" manual (23) was followed in using the Ortho-Rater. However, since the Sight

Screener procedure measures the extreme of excursion as a measure of lateral heterophoria, a similar procedure was adopted arbitrarily with one Ortho-Rater for purposes of comparison. A second Ortho-Rater was studied, and in using this instrument only the standard procedure was followed. The data from these two methods and from the standard administration of the two instruments are compared. In the Telebinocular tests, the subject read the values at which the indicator came to rest.

It was assumed that any error of phoria measurement inherent in the optical system of the instruments would be maximal for subjects with extremes of interpupillary distance, especially for measurements at the near point, (24). Accordingly, the data for near lateral phoria were divided according to the subjects' interpupillary distance as determined by five measures using an N.D.R.C. interpupillometer (19) (Figure 8).

Finally, an effort was made to determine correlation between heterophoria and age.

Since the unit graduations of heterophoria measurement varied for lateral and vertical deviations, it was necessary to encode the results of the clinical test so that the data could be handled conveniently. However, statistical results from the Maddox rod have been decoded and are presented in clinical ophthalmological terms.

Interpretation of Scores

The Bausch & Lomb "Ortho-Rater" results are expressed in terms of an arbitrary score as read from the instrument and must be converted into clinical units in accordance with the values stated in Table I (25).

Fig. 8.

N.D.R.C. Interpupillometer

TABLE I.

Standard Equivalents for Ortho-Rater Scores

A. Vertical Phoria Far

	Left Hyperphoria									Right Hyperphoria
*O-R Score	1	2	3	4	5	6	7	8	9	
**PDE	2	1½	1	½	0	0	½	1	1½	

B. Lateral Phoria Far

	Esophoria														Exophoria
O-R Score	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
PDE	6½	5½	4½	3½	2½	1½	½	½	1½	2½	3½	4½	5½	6½	7½

C. Vertical Phoria Near

	Left Hyperphoria									Right Hyperphoria
O-R Score	1	2	3	4	5	6	7	8	9	
PDE	1½	1	½	0	0	½	1	1½	2	

D. Lateral Phoria Near

	Esophoria														Exophoria
O-R Score	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
PDE	6½	5	3½	2	½	1	2½	4	5½	7	8½	10	11½	13	14½

*O-R - Ortho-Rater.
**PDE - Prism Diopter Equivalent.

The A. O. "Sight Screener" values are quoted from a key supplied with the instrument as indicated in Table II.

TABLE II.

Standard Equivalents for Sight Screener Scores

A. Vertical Phoria Near and Far	"Muscle Balance (Vertical) Space between two Adjacent Dots Equals ½ Prism Diopter Line for Orthophoria Crosses Fourth Dot From Top"	* Left * Hyperphoria * * — Orthophoria * * Right * Hyperphoria *

B. Lateral Phoria, Far and Near

"Muscle Balance (Lateral) Space Between Two Adjacent Dots Equals 1 Prism Diopter. Line for Orthophoria crosses at the numbered 15".

2 4 6 8 10 12 14 16 18 20 22 24 26 28
* *

Esophoria ! Exophoria

!

!

Orthophoria

The Keystone "Telebinocular" values for lateral heterophoria for far and near are read directly from the target in terms of prism diopters. The values for far vertical phoria are as given in Table III.

TABLE III

STANDARD EQUIVALENTS FOR TELEBINOCULAR VERTICAL PHORIA SCORES

	Right Hyperphoria					Left Hyperphoria			
Telebinocular	1	2	3	4	5	6	7	8	9
P.D. Equiv.	2	1½	1	½	0	½	1	1½	2

Results

The data from the present experiment are extensive. In order to delineate the necessary comparisons, data are presented in many individual tables.

Table IV summarizes the consistency of the measurements for lateral and vertical heterophoria at 20 feet distance (or its optical equivalent). The following values are listed: (1) product moment reliability coefficient (r), (2) test mean in terms of the values described above (T.M.), (3) test standard deviation (T.S.D.), (4) retest mean (R-T.M.), (5) retest standard deviation (R-T.S.D.), (6) standard error of retest estimate (E.E.). The letter "E" indicates esophoria, "X" indicates exophoria, "LH" and "RH" indicates left hyperphoria and right hyperphoria respectively.

TABLE IV

Test - Retest Consistency at 20 Feet.

A. Lateral Heterophoria.

Method	r	T.M.	T.S.D.	R-T.M.	R-T.S.D.	E.E.
Maddox Rod	.793	E 1.30	3.78	E 1.12	3.89	2.37
Ortho-Rater	.872	6.79	2.72	6.59	2.53	1.22
Sight Screener	.796	14.94	1.48	14.60	1.93	1.16
Telebinocular	.755	E 1.48	2.31	E 2.20	3.58	2.35

B. Vertical Heterophoria

Maddox rod	.623	LH 0.08	0.54	RH 0.28	.39	.30
Ortho-Rater	.630	5.17	1.07	5.11	1.01	.78
Sight Screener	.613	3.06	.37	3.02	.29	.23
Telebinocular	.602	4.71	.84	4.73	.76	.61

Table V summarizes the consistency of the measurements for lateral and vertical heterophoria at 13 inches, or its optical equivalent, by each method. The columns are the same as these in Table IV.

TABLE V

Test - Retest Consistency at 13 inches

A. Lateral Heterophoria

Method	r	T.M.	T.S.D.	R-T.M.	R-T.S.D.	E.E.
Maddox rod	.867	X 3.84	6.02	X 3.04	5.65	2.82
Ortho-Rater	.924	8.08	2.76	7.51	2.69	1.03
Sight Screener	.831	17.07	3.33	16.30	2.97	1.65
Telebinocular	.880	E 0.48	5.69	E 1.16	5.46	2.59

B. Vertical Heterophoria

Maddox rod	.677	RH 0.50	.61	RH 0.50	.57	.42
Ortho-Rater	.624	4.50	.83	4.59	.83	.65
Sight Screener	.551	3.02	.34	3.06	.39	.33

(There was no Telebinocular near vertical phoria measure.

Table VI presents the reliability coefficients for the four Maddox rod tests. Shown also are reliability coefficients for each of the several tests under consideration, and coefficients between instrument tests and the corresponding clinical measure.

TABLE VI

Phoric Relisability, Validity and Inter-correlation Data

	Maddox rod				Ortho-Rater				Sight Screener				Telebinocular		
	Lateral		Vertical		Lateral		Vertical		Lateral		Vertical		Lateral		Vertical
	Far	Near	Far	Near	Far	Near	Far	Near	Far	Near	Far	Near	Far	Near	Far
Maddox rod															
Latteral															
Far	.793	.698	.151	.173	.564				.370				.371		
Near		.317	.155	.159		.674				.543				.683	
Vertical															
Far			.523	.446			.286				.279				.426
Near				.677				.343				.337			
Ortho-Rater															
Lateral															
Far					.872										
Near						.924									
Vertical															
Far							.630								
Near								.624							
Sight Screener															
Lateral															
Far									.796						
Near										.831					
Vertical															
Far											.613				
Near												.551			
Telebinocular															
Lateral															
Far													.755		
Near														.880	
Vertical															
Far															.602

In Table VII are listed the values for (1) Ortho-Rater No. 1 employing standard procedure for lateral and vertical heterophoria measurements, (2) the same instrument employing the "excursion" type of measurements for far and near lateral heterophoria, (3) values for lateral and vertical heterophoria by Ortho-Rater No. 2 employing standard procedure and operated by a second examiner. The values are in Ortho-Rater units.

TABLE VII

Test Reliabilities of Two Ortho-Raters
and
Two Methods of Procedure

A. Heterophoria at 20 feet.

Ortho-Rater Method	r	T.M.	T.S.D.	R-T.M.	R-T.S.D.
			Lateral		
Excursion #1	.630	7.37	3.14	7.03	2.24
Standard #1	.872	6.79	2.72	6.59	2.56
Standard #2	.844	6.81	2.95	6.46	2.69
			Vertical		
Standard #1	.630	5.17	1.07	5.11	1.01
Standard #2	.592	5.00	.86	4.88	.66

B. Heterophoria at 13 inches.

			Lateral		
Excursion #1	.759	8.19	2.91	7.33	2.72
Standard #1	.924	8.08	2.76	7.51	2.69
Standard #2	.858	6.46	2.80	5.82	2.68
			Vertical		
Standard #1	.623	4.50	.83	4.59	.83
Standard #2	.715	4.70	.87	4.76	.84

Table VIII presents the mean near lateral reading for each of the three instruments, tabulated for each of three groups: (1) 23 subjects whose IPD was under 62 mm., (2) 63 subjects whose IPD was between 62 and 65.9 mm., and (3) 35 subjects whose IPD was 66 mm. or above. The values given are the means for each of the measures for lateral heterophoria at 13 inches.

It is to be noted that the Sight Screener is the only one of the instruments studied in which the near targets are viewed without the interposition of prismatic lenses.

TABLE VIII

Lateral Heterophoria at 13 inches

Mean Scores for Each IPD Group

	Under 62 mm.	62 - 65.9 mm.	66 mm. or over
A. Measurement with Prismatic Lenses			
Ortho-Rater	8.26	7.63	8.46
Telebinocular	8.96	8.51	9.03
B. Measurement without Prismatic Lenses			
Sight Screener	16.96	17.49	17.66
Maddox rod	X 4.20	X 3.12	X 4.38

An effort to find significant age correlates in this study was essentially fruitless.

Discussion

A preceeding section on Interpretation of Scores was included in this paper for the reason that its need is at once apparent when one attempts to evaluate the data. The difficulty is that with the exception of the Telebinocular slide for lateral imbalances, the scale units on the phoria targets supplied with each instrument are designated by arbitrary code digits. For each code the manufacturer of the device supplies a conversion key which is adequate for clinical purposes.

However, there is no certainty as to whether the conversion values for phoria score interval limits should be assigned to the mid-points or to the limits of the scale units. Until details of the scale conversions are available, one is not justified in converting arbitrary instrument score values into prism diopters. In this paper no decoded instrument values are presented.

The data presented in Tables IV and V suggest that the stabilizing feature incorporated in the Ortho-Rater lateral phoria slides might account for the enhanced reliability for this method in comparison with other methods of measuring lateral heterophoria. It is perhaps obvious that a stabilizing feature in measures of lateral phoria improves consistency. The question of the validity of such a feature will be considered later.

The reliability of the Maddox rod test for distance is not as high when the subject adjusts the prisms, as that found in the first part of this paper when the examiner makes the adjustment. For the near Maddox rod test the reverse is found. It should be remembered, however, that there is also a change in the viewing angle for near in the present experiment.

Tables IV and V indicate that, with the exception noted above, each of the four methods correspond with the others in consistency and there is little choice among them. All of the lateral phoria measures are more consistent than the vertical, in terms of reliability coefficients, means, standard deviations and errors of retest estimate.

The reliability and validity data for phorias in Table VI merit extended discussion.

(a) Each of the near lateral reliability coefficients is higher than the corresponding value for far. The opposite had been anticipated, on the basis of the results of the experiment described in the first part of this paper.

(b) The Maddox rod reliability coefficients for lateral phoria at 20 feet and at 13 inches are significantly larger than the inter-correlation coefficient between the near and far measures. A similar finding by use of the Ortho-Rater has been reported (19).

(c) The inter-correlation data for the instruments versus the criterion Maddox rod test are not as high as the reliability coefficients for either the instruments or the Maddox rod test. It is possible that the instruments are not measuring the identical functions determined by the Maddox rod method. Verhoeff (14) has made a statement quoted in the first part of this paper that for tests by use of stereoscopic devices "the observer has no accurate idea of the distances concerned". At least, the observer has a feeling of nearness of the targets when using screening devices, and this attitude may influence by association, the impulse to accommodate.

(d) In general, the lateral phoria measures elicited by the instruments correlate slightly better with the Maddox rod lateral phoria equivalents than do the corresponding values for the vertical measures.

(e) The comparison of two procedures in administering the lateral phoria test by the Ortho-Rater indicates that the standard procedure is more consistent than the so-called "excursion" procedure. Two different Ortho-Raters with standard procedure, elicit nearly identical data for means and standard deviations for distance and near.

(f) No correlation of phoria scores with interpupillary distance is noted. From the IPD analysis for this small population, it happened that the group with narrow IPD was slightly more exophoric (0.06 standard deviation units from the total group mean); the middle IPD group was slightly more esophoric (0.12 standard deviation units from the total group average); the wide IPD group was also more exophoric (0.09 standard deviation units from the total group mean). These differences in Maddox rod scores are probably insignificant. But it is interesting that they are reflected also in the mean instrument scores shown in Table VIII.

(g) It has been noted in the section on Results that the data show no meaningful correlation between age and phoria measures. The correlation coefficients tended to be magnified with the oldest group, but this can be attributed to the enhanced variability of the older group.

Summary

1. Three visual screening devices were used to measure heterophoria on 121 subjects who were also tested by the Maddox rod method.

2. A comparison of the data from the two experiments suggests that for the measurement at the near point, the angular deviation of the eyes from the horizontal position is apparently critical. The results obtained when the eyes are in the position usually employed in reading are more consistent than when the near measurement is made with the gaze directed horizontally forward.

3. The data from the two experiments suggest that clinical measurements of heterophoria are more consistent when the examiner adjusts the phorometer prisms than when the subject makes the adjustment.

4. There does not seem to be any justification for the use of arbitrary designations for heterophoria scores in non-clinical terms. Such use of coded digits complicates any statistical analysis of data obtained from visual screening devices.

5. A comparison of the measurements obtained by the Maddox rod method with those obtained by the use of visual screening devices indicates:

   a. That visual screening devices are apparently as reliable as the clinical method for the measurement of heterophoria.

   b. That the instrument tests do not correlate well with the clinical test; it is possible that this is due to some defect in construction of the devices or in the method of measurement.

c. That further research is indicated to establish why there is no more correspondence between the clinical and instrument measures.

6. No significant correlation between heterophoria and either interpupillary distance or age grouping was noted.

BIBLIOGRAPHY

(1) "Heterophoria Testing Manual of Instructions" (Restricted) developed by the Subcommittee on Procedures and Standards for Visual Examinations, of the Army-Navy-OSRD Vision Committee, 1945.

(2) "A Bibliography of Visual Literature, 1939-1944, and Supplement" (Secret) Publication No. 11, Historical Library, Yale Medical Library, 1945.

(3) "Relationship of Interpupillary Distance to Performance on the Ortho-Rater" Sulzman, J.H., Minutes and Proceedings of the Army-Navy-OSRD Vision Committee (Restricted) Twelfth Meething - 12 June 1945, Washington, D. C.

(4) "Testing Vergence and Phoria" Scobee, R.G., Minutes and Proceedings of the Army-Navy-OSRD Vision Committee (Restricted) Fourteenth Meeting - 11-12 September 1945, New London, Conn.

(5) "The Measurement of Heterophoria" Cridland, N., British Jour. Ophth., Vol. 25, No. 5 (May) 1941.

(6) "Objective Strabismometry in Young Children" Wheeler, M.C., Arch. Ophth., Vol. 29, No. 5 (May) 1943.

(7) "Motor Anomalies of the Eye. The Relative Value of the Tests Used in Diagnosing Them" Duane, A., N. Y. Med. Jour. Vol. 99, No. 9 (February) 1914.

(8) "The Extra-Ocular Muscles" Peter, L.C., Third Ed., pub. Lea and Febiger, Philadelphia, 1941.

(9) "Detecting, Measuring, Plotting and Interpreting Ocular Deviations" Lancaster, W.B., Arch. Ophth., Vol. 22, No. 5 (November) 1939.

(10) "Routine Muscle Examination in its Practical Significance" White, J.W., Trans. Sect. Ophth., A.M.A., 1934.

(11) "The Screen Test and its Modifications" White, J.W., Penn. Med. Jour., (September) 1936.

(12) "Routine Examinations of the Ocular Muscles and Nonoperative Treatment" Berens, C., Losey, R.L., and Hardy, L.H., Am. Jour. Ophth., Vol. 10, No. 12 (December) 1927.

(13) "An Approach to the Phorias" Livingston, P.C., British Orthop. Jour., 1:71-104, 1939.

(14) "Heterophoria Tests Based on a New Principle" Verhoeff, F.H., Arch. Ophth., Vol. 22, No. 5 (November) 1939.

(15) "Manual of Ocular Tests" pub. Council on Education and Professional Guidance, Am. Optom. Assoc., 1944.

(16) "The Maddox Rod and Screen Test Combined" Dolman, F., Air Service Inf. Circ., 1:42 (August 15) 1920.

(17) "A Tentative Interpretation of the Findings of the Prolonged Occlusion Test on an Evolutionary Basis" Harlow, F.W., Arch. Ophth., 19:194 (February) 1938.

(18) "Lectures on Motor Anomalies" Bielchowsky, A., pub. Dartmouth College Publications, Hanover, N. H., 1940.

(19) "Comparison of Various Screening Devices with Standard Medical Visual Procedure" Progress Report No. 1, on Bu&Med Research Div. Project X-493, Medical Research Dept., U. S. Submarine Base, New London, Conn., 22 November 1945.

(20) "Comparison of Five Methods of Screening for Acuity and Muscle Balance" Sulzman, J.M., Farnsworth, D., Bartlett, N., and Kindred, M.I., Minutes and Proceedings of the Army-Navy-OSRD Vision Committee, (Restricted) Eleventh Meeting - 19 April 1945, Washington, D.C.

(21) "Terminology in Ocular Motility" Lancaster, W.B., Am. Jour. Ophth., Vol. 26, No. 2 (February) 1943.

(22) "Visual Acuity Measurements with Three Commercial Screening Devices" Progress Report No. 2 in Bulled Research Div. Project X-493, Medical Research Department, U. S. Submarine Base, New London, Conn., 7 February 1946.

(23) "Standard Practice in the Administration of the Bausch & Lomb Occupational Vision Tests with the Ortho-Rater" pub. Bausch & Lomb Optical Co., Rochester, N.Y., 1944.

(24) "Near Horizontal Phoria and Duction Tests" Abraham, S.V., Am. Jour. Ophth., Vol. 26, No. 3 (March) 1943.

(25) "The Influence of the Visual Tasks Required of Personnel in the 16 Weeks Fire Controlmen (O) Training Course upon Their Visual Efficiency" (Restricted) OSRD Report No. 3970, 1 August 1944, NDRC Applied Psychology Panel.