Comparison of Four Dissociated Phoria Tests: Reliability & Correlation with Symptom Survey Scores

- **David A. Goss, O.D., Ph.D.**
- **Jennifer L. Reynolds, O.D.**
- **Rebekah E. Todd, O.D.**

School of Optometry
Indiana University
Bloomington, IN

Abstract

Previous studies have found the modified Thorington test and the Howell phoria card test to be more repeatable than the von Graefe test. Another potential way to assess the value of a test could be to correlate test results with the level of symptoms. Near dissociated phorias were measured with the modified Thorington, Howell phoria card, and von Graefe procedures with 50 subjects between the ages of 18 and 35. The von Graefe test was performed with two different targets, thus making a total of four different phoria tests. All testing was performed using a phoropter. Test findings were correlated with symptom scores from a questionnaire. The means on the four tests ranged from 2.7Δexo to 3.7Δ exo. Correlations of phoria findings with symptom scores were not statistically significant for any of the four tests in subjects with exophoria. In a small number of subjects with esophoria, significant correlations were found between phoria and symptom scores for the modified Thorington test and the Howell phoria card test but not with the von Graefe test. The modified Thorington test and the Howell phoria card test were more repeatable than the von Graefe test.

Key Words

binocular vision, clinical tests, heterophoria, phoria, repeatability, validity, vergence

INTRODUCTION

Dissociated phoria testing is an important part of the clinical evaluation of binocular vision. Probably the most commonly used subjective dissociated phoria test is the von Graefe test.[1,2] Another common test using a tangent scale and a Maddox rod over one eye is referred to as the modified Thorington test.[3,4] Studies have consistently found the modified Thorington test to be more repeatable than the von Graefe test.[5-11] An additional test, attributed to Prentice,[9] uses prism dissociation of a tangent scale. One study reported the repeatability of that test using a card marketed by Howell to be comparable to that for the modified Thorington.[9]

To be of most value to the practitioner, clinical tests should be both reliable (repeatable) and valid. Comparisons of the repeatability of von Graefe and modified Thorington dissociated phoria tests have been examined in previous publications, but to our knowledge, comparisons of the validity of these tests have not been done. One method whereby the validity of a test might be assessed could be by determining whether the results of the test in question correlate with symptoms. This paper reports a study using a symptom questionnaire to examine the correlation of symptoms with findings on von Graefe, modified Thorington, and Howell card dissociated phoria tests. Data for repeatability of each the tests and for comparisons among the tests are also provided.

METHODS

Fifty people volunteered to participate in this study. Subjects were between the ages of 18 and 35, had a best corrected visual acuity of 20/25 or better in each eye, and did not have self-reported strabismus. The subjects were predominantly optometry students.

Subjects were seated comfortably behind a phoropter with foreheads against the phoropter forehead rest. The subjects' habitual spectacle corrections determined by lensometer reading were placed in the phoropter lens banks. Contact lens wearers wore their contact lenses during testing, and for them the phoropter lens power was set at plano.

Odd numbered subjects were tested in the following order: von Graefe with Look Here card, von Graefe with Borish card, modified Thorington, Howell card. Even numbered subjects were tested as follows: von Graefe with Borish card, von Graefe with Look Here card, Howell card, modified Thorington test. The von Graefe tests were performed before the Howell and modified Thorington tests because it was thought that the results of the latter tests might bias the examiner during the performance of the von Graefe tests. The phoropter Risley rotary prisms were used for testing. Measurements were recorded in prism diopters. The test sequence was repeated a second time with 4 prism diopters base in over the subject's right eye. One examiner performed the test blind to the amount of prism placed in front of the subject and unaware that the amount of prism used was invariant from subject to subject. The other examiner held the undisclosed prism in front of the subject (behind the phoropter) using a lens holder with a handle. The examiner who held the

prism also recorded the measurements. After the tests were completed the subjects completed a survey regarding reading eyestrain symptoms.

von Graefe Test with Look Here Card

Figure 1. Look Here target used for the von Graefe Test

Twelve prism diopters base in was placed in front of the right eye and six prism diopters base up in front of the left eye using the phoropter rotary prisms. The subjects viewed the nearpoint card target shown in Figure 1 at 40 cm and were instructed to keep the stationary target clear. Base in prism was decreased, and the subject reported when one target appeared to be directly above the other. The examiner decreased the base in prism further (or increased base out) and then instructed the subject to indicate alignment again as the prism was moved back in the base-in direction. The average of the two readings was recorded.

von Graefe Test with Borish Card

Figure 2. Target on the Borish Nearpoint Card for the von Graefe Test.

The Borish Nearpoint Card includes a diamond shaped target with letters inside (Figure 2) for von Graefe dissociated phoria test. This test was started with the same rotary prism settings as the von Graefe test with the Look Here card. The subjects viewed a Borish card at 40 cm and were instructed to keep the stationary image clear and indicate when the two diamonds were aligned vertically. The
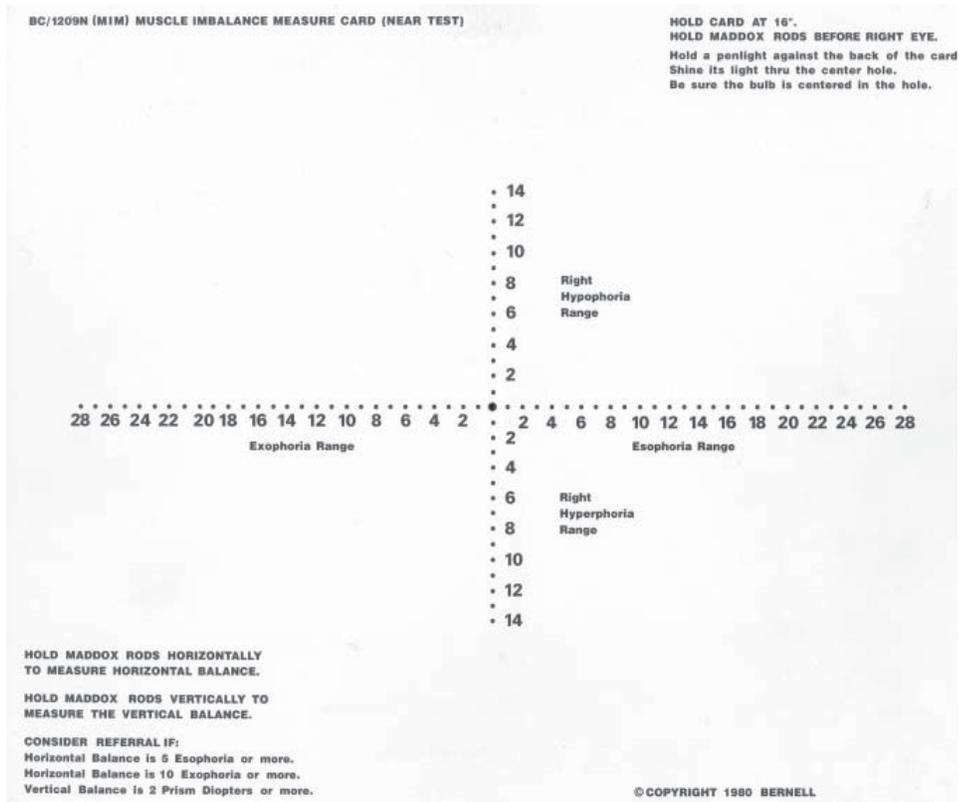

Figure 3. The Bernell Muscle Imbalance Measure card, with which the modified Thorington test was performed.
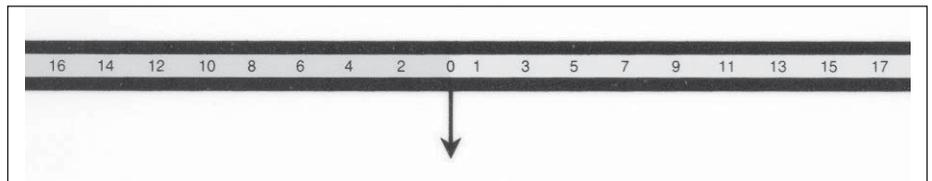

Figure 4. The scale on the Howell card used for the phoria test.

procedure was the same as that for the von Graefe test with the Look Here card.

Modified Thorington Test

The target was the Bernell Muscle Imbalance Measure (MIM) card (Figure 3) consisting of vertical and horizontal axes marked with numbers calibrated to read in prism diopters when the card is at 40 cm. A Maddox rod was placed with horizontal bars in front of the right eye. The examiner placed a penlight in the central hole. With the card at 40 cm, the subject indicated where the red vertical line appeared to be on the lateral scale. Numbers to the right of the light indicated eso, and numbers to the left indicated exo.

Howell Card

Dissociated phoria testing with the Howell card (Figure 4) uses prism-induced diplopia of a numbered scale. With the Howell card in place at 33 cm, six prism diopters base up was placed over the left eye to induce dissociation. To confirm dissocia-

tion, the subjects were asked if they saw two rows of numbers and two downward pointing arrows. Subjects were told to keep the numbers clear at all times. The subjects looked at the top arrow and reported where that arrow pointed on the bottom row of numbers, as well as the corresponding color. Yellow numbers to the right of zero indicated esophoria, and blue numbers to the left of zero indicated exophoria. Testing with the Howell card was performed at 33 cm since its scale is calibrated for that distance, and the numbers indicate the phoria in prism diopters.

Eyestrain Symptom Questionnaire

A questionnaire similar to that used in studies on convergence insufficiency and validated for the assessment of the level of symptoms[12] was used in this study. Each of the following questions was rated on a scale (never, infrequently, sometimes, fairly often, always) by the subject:
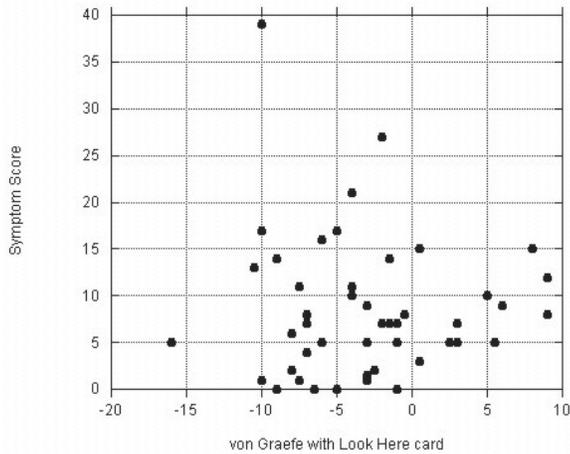
Figure 5. Scatterplot of symptom score with dissociated phoria from the von Graefe test using the Look Here card.
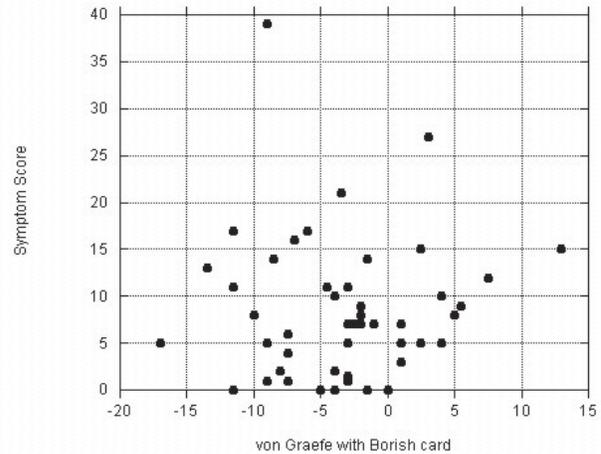

Figure 6. Scatterplot of symptom score with dissociated phoria from the von Graefe test using the Borish card.

Table 1. Pearson product-moment correlation coefficients between phoria and symptom score. Asterisks indicate statistical significance of the correlation coefficients (*p<0.02; **p<0.01).

Phoria Method	Exophoria	Esophoria
von Graefe, Look Here card	-0.12 (n=39)	0.35 (n=11)
von Graefe, Borish card	-0.15 (n=38)	0.31 (n=12)
modified Thorington	-0.07 (n=38)	0.71 (n=12)**
Howell card	-0.15 (n=41)	0.76 (n=9)*

Table 2. Means and standard deviations of each phoria test in the no added prism condition. Units are prism diopters (negative=exo).

Phoria Method	Mean ± Standard Deviation
von Graefe, Look Here card	-3.0 ± 5.4
von Graefe, Borish card	-3.4 ± 5.8
Modified Thorington	-2.7 ± 4.8
Howell card	-3.7 ± 5.2

1. Do your eyes feel tired when reading or doing close work?
2. Do your eyes feel uncomfortable when reading or doing close work?
3. Do you have headaches when reading or doing close work?
4. Do you feel sleepy when reading or doing close work?
5. Do you have double vision when reading or doing close work?
6. Do the words seem to move or jump when reading or doing close work?
7. Do your eyes ever hurt when reading or doing close work?
8. Do your eyes ever feel sore when reading or doing close work?
9. Do you feel a "pulling" around your eyes when reading or doing close work?
10. Do you notice the words blurring or going in and out of focus when reading or doing close work?

The responses were scored according to the following scale: never=0, infrequently=1, sometimes=2, fairly often=3, always=4. The responses were summed to give an overall score potentially ranging from 0 to 40.

Statistical Analysis

The primary analysis was a comparison of phoria measurements to the symptom score. For this analysis, the initial phoria measurement (without the addition of the 4Δ BI prism) was used. The relationship of symptom score to phoria was examined using Pearson correlation coefficients. Because it would be expected that symptoms would increase with an increase in esophoria and with an increase in exophoria beyond normal levels; correlations were calculated separately for esophoria and exophoria. Ortho findings were included with the exo grouping. In calculations, exophoria was negative and esophoria was positive.

The second analysis was repeatability of each of the tests. The first and second measurements with each test were compared after the 4Δ BI had been subtracted out of the result from the second measurement. A mean difference and standard deviation of the difference was determined for each test, a lower standard deviation of the differences indicating a better repeatability.

The third analysis was comparison between tests. In this analysis, the initial phoria measurement (without the addition of the 4Δ BI prism) was used. Mean differences between tests and standard deviation of differences were calculated. A coefficient of agreement (COA) between tests was calculated by multiplying the standard deviation of the differences by 1.96. The COA represents the range of differences that could be expected between the two tests 95% of the time. A low COA would thus indicate better agreement.

Correlation coefficients of the individual differences between two tests with the mean of the two tests for each individual were calculated. This method was recommended by Altman and Bland,[13] to assess whether differences between tests changed as the magnitude of the measurements varied. A low correlation of differences with means would thus be more desirable because it would indicate that the difference between tests was fairly constant over the full range of test findings.

RESULTS

Correlation of Test Results with Level of Symptoms

Correlation coefficients are given in Table 1. Figures 5 to 8 are scatterplots of symptom scores with phoria. Symptom scores did not show statistically significant cor-
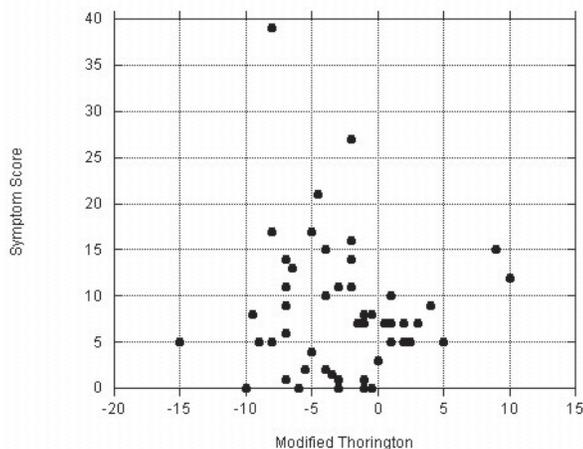
Figure 7. Scatterplot of symptom score with dissociated phoria from the modified Thorington test.
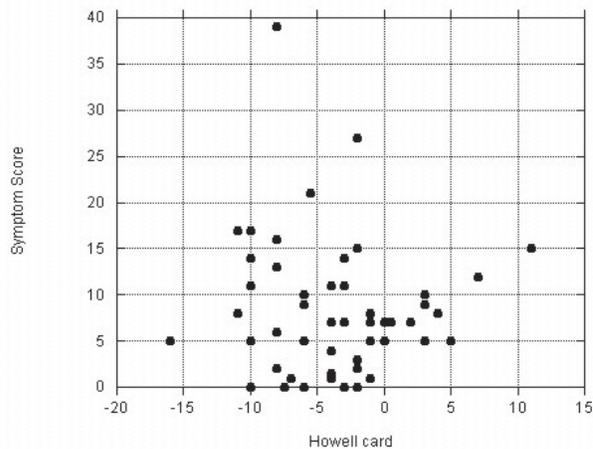

Figure 8. Scatterplot of symptom score with dissociated phoria from the Howell card.

Table 3. Comparison of the first (no added prism) and second (with addition of 4Δ BI) performance of each test: Mean differences and standard deviation of the differences.

Phoria Method	mean difference	Standard deviation of the differences
von Graefe, Look Here card	+1.6	±4.0
von Graefe, Borish card	+1.0	±3.6
Modified Thorington	+1.2	±2.7
Howell card	+1.4	±3.0

The 4Δ BI was subtracted out for the results of the prism altered condition.
The differences are the first test minus the second test. The units are prism diopters.
(minus = exo)

relations with exophoria on any of the phoria tests, with the r values ranging from -0.07 to -0.15. In subjects with esophoria, the von Graefe methods produced correlation coefficients of 0.31 and 0.35. The modified Thorington test produced a correlation coefficient of 0.71 with symptom score and is statistically significant (n=12; p<0.01). The Howell card's coefficient of correlation with symptom score was 0.76, which is also statistically significant (n=9; p<0.02).

Repeatability of Tests

Means and standard deviations for the initial performance (with no added prism) of each of the tests are listed in Table 2. Phoria tests unaltered by prism had means between -2.7 and -3.7Δ. The means and the standard deviations of the differences between the first and second performance of each test are given in Table 3. The lowest standard deviations of the differences, indicating better repeatability, were found with the modified Thorington test (2.7Δ) and the Howell card (3.0Δ). The highest standard deviations of the differences were found on the von Graefe test with the

Borish card (3.6Δ) and the von Graefe test with the Look Here card (4.0Δ).

Comparison of Tests

Results of the analyses of the comparison of tests can be found in Table 4. The differences between means were all one prism diopter or less. The coefficients of correlation of the various tests with each other were all fairly high, with a range of 0.84 to 0.93 (Table 4). The coefficients of correlation of the differences of two tests with the means of those same two tests were all fairly low, ranging in magnitude from 0.07 to 0.37. The comparison showing the greatest correlation of mean with difference (r=-0.37) was the comparison of modified Thorington with the von Graefe using the Borish card. That correlation would indicate a trend toward the modified Thorington showing less convergent phorias in esophoria cases and less divergent phorias in exophoria cases compared to the von Graefe test done with the Borish card. That trend toward less variability on the modified Thorington than on the von Graefe can be seen in the standard deviation for the modified Thorington (4.8Δ) being less than that

for the von Graefe with the Borish card (5.8Δ).

The lowest coefficient of agreement (COA) between tests was found for the comparison of modified Thorington with Howell card (Table 4). The highest COA was found for comparison of modified Thorington with von Graefe test using the Look Here card.

DISCUSSION

Correlation of Test Results with Level of Symptoms

The correlation of test results with symptom scores was examined as a possible means of evaluating the validity of the tests. The von Graefe test has been used as a standard subjective dissociated phoria test for many years. However, the better repeatability with other tests, such as the modified Thorington and the Howell card has led to the question of whether the modified Thorington or Howell phoria card should be preferred to the von Graefe test. If a given test is more repeatable but less valid than another test, it may not represent a better test. The modified Thorington and Howell card showed statistically significant correlations with symptom scores in subjects with esophoria, but the von Graefe test did not. The magnitudes of the correlation coefficients of phoria with symptom score were low for all tests in subjects with exophoria.

The fact that correlations with symptom scores for the modified Thorington test and the Howell phoria card test were similar to (exophoria) or better than (esophoria) those for the von Graefe test suggests that they might be preferred over the von Graefe. That recommendation should be tempered, of course, by the low number of

Table 4. Comparison of tests.
For each comparison, the mean difference is for the first test listed minus the second test listed.

	Mean difference (minus = exo).	Standard deviation of difference	Coefficient of agreement	Correlation of test results (r)	Correlation of difference with mean (r)
von Graefe, Borish card; von Graefe Look Here	-0.4	±2.1	4.1	0.93	0.20
Modified Thorington; von Graefe, Look Here card	0.3	±3.0	5.8	0.84	0.22
Howell; von Graefe, Look Here card	-0.7	±2.4	4.8	0.89	0.07
Modified Thorington; von Graefe, Borish card	0.7	±2.8	5.5	0.87	0.37
Howell; von Graefe, Borish card	-0.3	±2.4	4.7	0.91	0.24
Howell; modified Thorington	-1.0	±2.0	3.9	0.92	0.22

Units are prism diopters.

Table 5. Means (with standard deviations in parentheses) from studies comparing von Graefe near phorias with modified Thorington near phorias and/or Howell card near phorias.

Study	Prism dissociation/ measurement with prism (von Graefe)	Maddox rod/ tangent scale (modified Thorington)	Prism dissociation/tangent scale with Howell card
Scobee and Green[21]	3.0 exo (5.2) (13 inches)	2.8 exo (4.3) (13 inches)	--
Hirsch and Bing[5]	4.6 exo (5.4)	3.4 exo (4.2)	--
Hirsch,[22] students	4.0 exo (6.4)	3.2 exo (6.2)	--
Hirsch,[22] instructors	4.3 exo (6.1)	3.1 exo (5.9)	--
Rainey et al.,[8] examiner 1	4.3 exo (4.7)	2.2 exo (3.0)	--
Rainey et al.,[8] examiner 2	5.0 exo (6.3)	2.1 exo (3.2)	--
Wong et al.[9]	2.2 exo (4.2)	2.8 exo (3.7)	2.3 exo (4.2) (33 cm)
Goss et al.,[20] examiner 1	1.9 exo (4.4)	2.2 exo (3.6)	--
Goss et al.,[20] examiner 2	2.4 exo (5.1)	2.3 exo (4.1)	--
Maples et al.[23]	6.3 exo (5.7) (33 cm)	--	2.1 exo (3.1) (33 cm)
Present study	3.0 exo (5.4) 3.4 exo (5.8)	2.7 exo (4.8)	3.7 exo (5.2) (33 cm)

(Some studies examined effects of variations in testing procedures, in which case data in this table are those with small test target letters and with continuous viewing of the test target. Test distances are 40 cm unless noted otherwise.) Units are prism diopters.

subjects in this study. Replication of such findings with additional studies would be helpful.

The low correlation between symptom scores and phorias for subjects with exophoria is consistent with studies by Sheedy and Saladin.[14,15] They found symptoms in exophoria to be better correlated with a combination of phoria and base-out vergence range (Sheard's criterion) than phoria alone.

Repeatability of Tests

For the study of repeatability the second measurement with each test was done with the placement of a 4 Δ BI prism in front of the subject. This was done in order to reduce the possibility of the subjects remembering the numbers they reported the first time on the modified Thorington and Howell tests and having a bias toward reporting the same number the second time. The test measurements obtained with the added base-in prism would be expected to be more eso or less exo than without the prism. The average esophoric shifts were 2.4Δ for von Graefe test with Look Here card, 3.0Δ for von Graefe test with Borish card, 2.8Δ for modified Thorington, and 2.6Δ for Howell. Using a similar technique, Morris[6] found an average difference of 1.2Δ more exo when he used 2Δ BO with the modified Thorington test. The fact that the phoria did not change the same amount as the power of the prism placed in front of the eye might be explained in part on the basis of prism effectivity.[16-18] Due to prism effectivity, the effect of a prism on eye position for a near point object is less than the labeled power of the prism.[19]

Repeatability in the present study was assessed using the standard deviation of the differences between the first and second times each test was performed. The present study was consistent with the findings of several previous studies,[5-10] in finding the modified Thorington test to be more repeatable than the von Graefe method. In the current study, as in the Wong et al.[9] study, repeatability was better with the Howell card than with the von Graefe test. The two von Graefe methods in this study showed similar repeatability results.

Casillas Casillas and Rosenfield[10] found dissociated phoria tests to generally be more repeatable with trial frame testing than in a phoropter. The Howell phoria

card and modified Thorington cards are usually used out of the phoropter, but in this study all tests were done in the phoropter to eliminate the presence or absence of the phoropter as an additional variable. Even in the phoropter, the modified Thorington was more repeatable than the von Graefe, consistent with the findings of Casillas Casillas and Rosenfield.[10]

Comparison of Tests

The differences between the means for the four tests were all 1Δ or less. The correlation coefficients between tests were all high, ranging from 0.84 to 0.93. The Pearson correlation coefficients between the sum of two tests and the difference of the same tests were generally low, suggesting that the differences in the phoria test measurements did not vary significantly with the magnitude of the phoria. Only one of these correlations was statistically significant, although low at r=-0.37, that for the comparison of modified Thorington with von Graefe using the Borish card. The direction of the correlation showed a trend toward less convergent measurements in higher eso cases and less divergent measurements in higher exo cases using the modified Thorington than with the von Graefe using the Borish card. A previous study comparing von Graefe and modified Thorington tests included an Altman and Bland analysis and found significant correlations of the test differences and their means for both distance and near phorias.[20] A conclusion of that study was that: "For higher exo or higher eso, the von Graefe method tends to yield higher values than the modified Thorington test."[20]

Several studies have compared results on modified Thorington and von Graefe tests. Results from those studies are summarized in Table 5. Standard deviations were consistently higher on the von Graefe than on the modified Thorington, again showing more variability on the von Graefe.

The present study found differences of less than 1Δ between Howell card test mean results and von Graefe mean findings. That is consistent with the findings of Wong et al.[9] A larger difference between the two was found by Maples et al.[23] (see Table 5). In the present study and in the Wong et al. study, the Howell card was used at 33 cm and the other tests were performed at 40 cm. In the Maples et al. study, both von Graefe and Howell card testing were done at 33 cm.

CONCLUSIONS

Consistent with the findings of previous studies, the modified Thorington test and the Howell phoria card test were more repeatable than the von Graefe test. Correlations of questionnaire symptom scores with test findings on the modified Thorington test and the Howell phoria card test was similar to (exophoria) or better than (esophoria) those with the von Graefe test. The study findings appear to support the value of the modified Thorington and Howell card phoria tests, although additional study with larger numbers of subjects with esophoria is needed to confirm the correlation with symptom score findings.

Note

The authors have no proprietary interest in any of the equipment used in this study.

References

1. Daum KM. Heterophoria and heterotropia. In: Eskridge JB, Amos JF, Bartlett JD, eds. Clinical Procedures in Optometry. Philadelphia: Lippincott, 1991:72-90.
2. Grosvenor T. Primary Care Optometry, 5th ed. St. Louis: Butterworth Heinemann Elsevier, 2007:224-26.
3. Scheiman M, Wick B. Clinical Management of Binocular Vision: Heterophoric, Accommodative, and Eye Movement Disorders, 2nd ed. Philadelphia: Lippincott Williams & Wilkins, 2002:8,37-38.
4. Carlson NB, Kurtz D. Clinical Procedures for Ocular Examination, 3rd ed. New York: McGraw-Hill, 2004:211-13.
5. Hirsch MJ, Bing LB. The effect of testing method on values obtained for phoria at forty centimeters. Am J Optom Arch Am Acad Optom 1948;25:407-16.
6. Morris FM. The influence of kinesthesis upon near heterophoria measurements. Am J Optom Arch Am Acad Optom 1960; 37: 327-51.
7. Schroeder TL, Rainey BB, Goss DA, Grosvenor TP. Reliability of and comparisons among methods of measuring dissociated phoria. Optom Vis Sci 1996; 73: 389-97.
8. Rainey BB, Schroeder TL, Goss DA, Grosvenor TP. Inter-examiner repeatability of heterophoria tests. Optom Vis Sci 1998;75:719-26.
9. Wong EPF, Fricke TR, Dinardo C. Interexaminer repeatability of a new, modified Prentice card compared with established phoria tests. Optom Vis Sci 2002;79:370-75.
10. Casillas Casillas E, Rosenfield M. Comparison of subjective heterophoria testing with a phoropter and trial frame. Optom Vis Sci 2006;83:237-41.
11. Goss DA, Penisten DK, Pitts KK, Burns DA. Repeatability of prism dissociation and tangent scale heterophoria measurements in straightforward gaze and in downgaze. In: McCoun J, Reeves L, eds. Binocular Vision: Development, Depth Perception, and Disorders. Hauppauge, NY: Nova Science, 2010.
12. Borsting EJ, Rouse MW, Mitchell GL, Scheiman M, et al. Validity and reliability of the revised Convergence Insufficiency Symptom Survey in children aged 9 to 18 years. Optom Vis Sci 2003;80:832-38.
13. Altman DG, Bland JM. Measurement in medicine: The analysis of method comparison studies. Statistician 1983;32:307-17.
14. Sheedy JE, Saladin JJ. Phoria, vergence, and fixation disparity in oculomotor problems. Am J Optom Physiol Opt 1977;54:474-78.
15. Sheedy JE, Saladin JJ. Association of symptoms with measures of oculomotor deficiencies. Am J Optom Physiol Opt 1978;55:670-76.
16. Alpern M. The position of the eyes during prism vergence. Arch Ophthalmol 1957;57:345-53.
17. Ogle KN, Martens TG. On the accommodative and the proximal convergence. Arch Ophthalmol 1957;57:702-15.
18. Pratt CB. The variation in phorias with time after dissociation and magnitude of convergence. Am J Optom Arch Am Acad Optom 1962;39:257-63.
19. Brown WL. Optical principles of prism. In: Cotter SA, ed. Clinical Uses of Prism: A Spectrum of Applications. St. Louis: Mosby, 1995:1-40.
20. Goss DA, Moyer BJ, Teske MC. A comparison of dissociated phoria test findings with von Graefe phorometry and modified Thorington testing. J Behav Optom 2008;19:145-49.
21. Scobee RG, Green EL. Reliability of tests, comparison between tests, and effect of changing testing conditions. Trans Am Acad Ophthalmol Otolaryngol 1947; 51: 179-97.
22. Hirsch MJ. Clinical investigation of a method of testing phoria at forty centimeters. Am J Optom Arch Am Acad Optom 1948; 25: 492-95.
23. Maples WC, Savoy RS, Harville BJ, Golden LR, et al. Comparison of distance and near heterophoria by two clinical methods. Optom Vis Dev 2009;40:100-06.

Corresponding author:
David A. Goss, O.D., Ph.D
School of Optometry, Indiana University, Bloomington, IN 47405
dgoss@indiana.edu
Date accepted for publication:
April 29, 2010