Repeatability and Reliability of a New Phoria Test Using Flashed and Auto-flashed Presentation

Byeong-Yeon Moon[1], Sun-Young Park[2], and Hyun Gug Cho[1]*

[1]*Department of Optometry, Kangwon National University, Samcheok 245-907, Korea*
[2]*Johnson & Johnson Vision Care, Seoul 140-702, Korea*

The purpose of this study was to assess repeatability and reliability of a new phoria test as compared to established phoria tests and to assess the possibility of mobile or online testing using the auto-flashed presentation. Near dissociated phoria was measured using the von Graefe method, the Howell card test, the modified Thorington test, and a new dichromatic card test (Red-Blue Phoria card test; RBP card test) on 109 subjects. The inter-examiner difference and test-retest difference were calculated. With the auto-flashed presentation of the RBP card test, near dissociated phoria was measured and the intraclass correlation coefficient and test-retest repeatability was assessed on 26 subjects. The variation in inter-examiner repeatability was the smallest for the modified Thorington test (+3.1/-2.6). The RBP card test was +3.0/-3.1, the Howell card test was +3.5/-2.7, and the von Graefe test was +6.2/-6.0. The variation in test-retest repeatability was the smallest for the RBP card test (+0.4/-1.4). The modified Thorington test was +1.3/-1.3, the Howell card test was +1.27/-1.45, and the von Graefe test was +1.59/-2.20. The ICC was 0.919 for the auto-flashed presentation of the RBP card test. The RBP card test is considered a highly repeatable method clinically and can be made a useful application for testing in mobile or online.

Keywords : New phoria test, RBP card, Auto-flashed presentation, Mobile, Online
OCIS codes : (170.4460) Ophthalmic optics and devices; (330.1400) Vision-binocular and stereopsis; (170.1610) Clinical applications

I. INTRODUCTION

Measurement of phoria is a routine test in optometric examination and plays an important part in the diagnosis of patient symptoms such as blur, headache, eyestrain, or diplopia and in the management of accommodation vergence disorders [1, 2]. Clinically accurate and repeatable measurement of phoria is critical when diagnosing and managing such conditions [2]. There are numerous methods used clinically to measure a phoria. For measurement of phoria, binocular fusion has to be eliminated. The basis of these methods was as follows: by occlusion of one eye, by displacement of a single image using a prism, by distortion of a single image, or by the use of dissimilar targets [3]. The tests on the basis of these methods are called the cover test, the von Graefe method, the Maddox rod test, the Thorington method, a modified Thorington test, stereoscopes, and the Howell phoria card method. Most methods of phoria measurement are subjective. The variability, reliability and repeatability of phoria tests have been published in the magnitude of prism diopters by each method [4-9]. Variability in results of phoria tests was attributed to possible examiner bias, test technique variations, inadequate accommodative control, and subject response skill [10]. The ideal test should have good reliability and validity and should be simple to administer [11].

We tried to ascertain another test which is fast, accurate and can be applicable to online or mobile testing. We have found a new test that can quantify the phoria without a prism or a special lens. A dichromatic color filter and a new phoria card which can be seen as dissimilar targets based on color composite image were used. The key point of this new test is to be able to self-test with an auto-flashed presentation via a programmed display. The purpose of this study was to assess repeatability and reliability of a

---

*Corresponding author: hyung@kangwon.ac.kr

Color versions of one or more of the figures in this paper are available online.

new phoria test as compared to established phoria tests and to assess the possibility of mobile or online application using the auto-flashed version of a new phoria test.

II. METHODS

One hundred nine undergraduate optometry students (Korean, average age 23.8 ± 1.8 years; 73 males and 36 females) volunteered to participate in this study. Each subject provided informed consent to participate in this study based on oral explanation. The study was conducted in adherence to the tenets of the Declaration of Helsinki. Subjects had monocular visual acuity of at least 0.9 with best correction, normal stereopsis, and no physical problems. Preliminary tests were individually completed in one week.

Two faculty members from the Dept. of Optometry at Kangwon National University performed the phoria measurement procedure. Each examiner evaluated the phoria at near distance (40 cm or 33 cm) using four different methods: von Graefe method, Howell card test, modified Thorington test, and RBP card test. For the analysis of inter-examiner difference, the first examiner performed the four methods in order then the second examiner immediately repeated the procedure to evaluate phoria. After 24 hours from the first test, the two examiners measured the same subjects again by the same procedure for an analysis of test-retest difference.

To assess the possibility of mobile or online application for the auto-flashed presentation of a new phoria test, 26 new subjects (Korean, average age 24.3 ± 3.0 years; 19 males and 7 females) were tested. The subjects measured near phoria by themselves using the auto-flashed presentation on the monitor. All conditions for subjects and measurement were the same as the prior testing. A description of the procedure was presented well before the tests; each test was performed two times at 24 hour intervals. Test-retest repeatability and intraclass correlation coefficient were evaluated.

2.1. Von Graefe Method − Continuous Presentation [12]

A 6 Δ base-up was placed over the left eye and a 12 Δ base-in was placed over the right eye in the subject seated behind the phoropter. A vertical arrow target was held at 40 cm from the subjects by the examiner. Subjects were instructed to keep the target clear at all times. If subjects didn't see target dissociation vertically, the examiner adjusted the Risley prism over the left eye until dissociation occurred. The examiner then checked the relative direction of the upper target to the lower one, right or left. The horizontal prism over the right eye was moved slowly until the subject said that two targets were aligned vertically. At this point, prism power and base direction were recorded.

### 2.2. Howell Card Test − Flashed Presentation [1]

The Howell near phoria card was held at 33 cm from the subjects by the examiner. An 8 Δ base-up was placed over the subject's left eye using a trial frame. The examiner placed the occluder over the left eye and subjects were instructed to keep the downward-pointing arrow clear at all times. When the occluder was removed, subjects were asked to instantaneously report what number the top arrow pointed to on the bottom line. The arrow pointing to an odd number on a yellow background indicated an esophoria while an even number on the blue background indicated exophoria.

### 2.3. Modified Thorington Test − Flashed Presentation [10]

The MIM card (Bernell Muscle Imbalance Measure card) was used for this method. The MIM card was held at 40 cm from the subjects by the examiner. The card consists of horizontal and vertical scales of numbers graduated by 1 Δ at 40 cm. There is a hole at the intersection of vertical and horizontal scales through which a penlight is shown to the subject. A red Maddox rod with striations horizontally was placed over the subject's left eye using a trial frame. The penlight through the hole shows a red vertical line by the Maddox. The examiner placed the occluder over the left eye and subjects were instructed to keep the numbers clear at all times. When the occluder was removed, subjects were asked to instantaneously report the number on the red line passed. A red line passing on the left of the penlight indicated an esophoria and on the right of the penlight indicated exophoria.

2.4. RBP Card Test − Flashed Presentation

The new dichromatic phoria card for this test was named the red and blue phoria card (RBP card). The RBP card consists of a red horizontal line and a blue vertical line with scales of numbers graduated by 1 Δ at 40 cm on a black background (Fig. 1). The intersection of blue and red lines was zero. On the right of the red line and top of the blue line positive numbers were printed, and on the left of the red line and the bottom of the blue line negative numbers were printed. Wearing a blue filter over the eyes, the red horizontal line on the card was hardly distinguishable from the black background but the blue vertical line was clearly visible. Conversely, wearing a red filter over the eyes, the blue vertical line on the card was hardly distinguishable from the black background but the red horizontal line was clearly visible. The RBP card was held at 40 cm from the subjects by the examiner. Subjects wore red (right) and blue (left) filter glasses, a red line was shown in the subject's right eye and a blue line was shown in the left eye. The examiner placed the occluder over the right eye and subjects were instructed to keep the numbers of the blue line clear at all times. While the examiner performed an alternating cover technique continuously (1 cycle per second) subjects were asked to instantaneously report the number on the red line at the point of intersection with the blue line. The blue line intersected on the positive
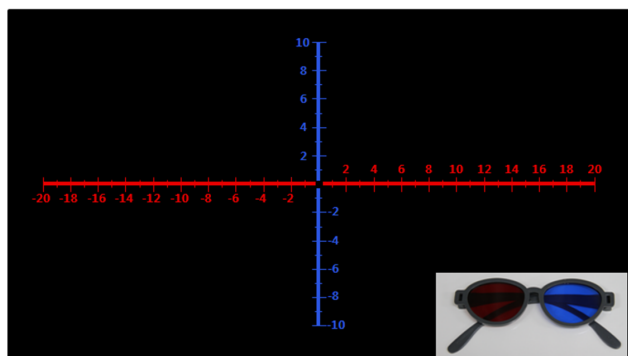
(a)
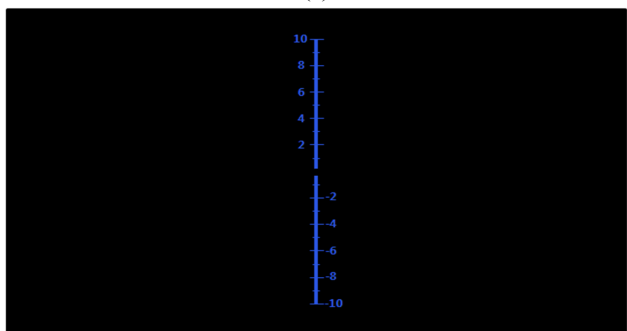
(b)
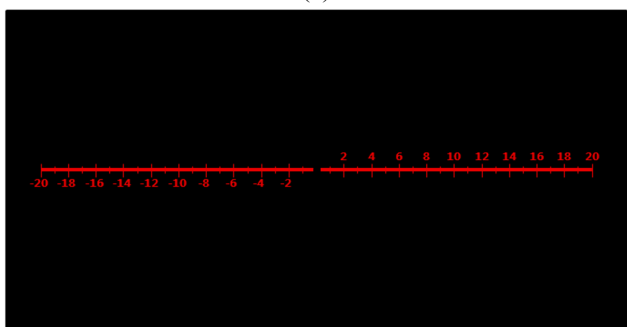
(c)

FIG. 1. (a) Red-Blue Phoria card (RBP card) consists of red horizontal line and blue vertical line with prism scales at 40 cm on black background. On red filter (right eye) and blue filter (left eye), (b) blue line is shown to left eye only and (c) red line is shown to right line only.

number of the red line indicated an exophoria while intersection on the negative number of the red line indicated esophoria.

2.5. RBP Card Test − Auto-flashed Presentation

The auto-flashed version of a RBP card was made by PowerPoint 2007 (Microsoft Co., Redmond, USA), which displayed a blue vertical line and a red horizontal line alternately (1 cycle per second) on a computer monitor. The procedure of the test was equal to RBP card-flashed presentation with the exception of the examiner's alternating cover technique.

2.6. Statistical Analysis

Data analysis was performed with MedCalcTM (Med-Calc Software, Maria Kerke, Belgium) using a Bland and Altman plot [13] to assess the level of agreement between examiners and between test-retest. The mean difference, the standard deviation (S.D.), and the 95% limits of agreement (the lower and upper limits of agreement = $\pm 1.96 \times$ S.D. of mean difference) were determined. The reliability of averages of k ratings was estimated by intraclass correlation coefficient (ICC). Descriptive statistic, paired t-test, and one-way ANOVA test were performed in a 95% confidence level with SPSS for windows (SPSS Inc., Chicago, USA).

III. RESULTS

For each examiner, mean and standard deviation for each phoria test are given in Table 1. Positive and negative values represent esophoria and exophoria respectively. The mean by the von Graefe method was significantly the highest exophoria ($p < 0.05$), then followed by the modified Thorington test, Howell card test, and RBP card test. The range of mean exophoria was between 6.87 Δ / 7.03 Δ (test / retest) by the von Graefe method and 3.42 Δ / 3.67 Δ by RBP card test with all tests. Inter-examiner difference was not significant statistically in each method.

Bland and Altman recommended calculating the mean difference and 95% limits of agreement between examiners as the most appropriate indication of inter-examiner test repeatability. Mean difference comparisons are represented

TABLE 1. Mean ± standard deviation for the phoria tests for each examiner and the mean of the two examiners (n=109)

Methods	Test			Retest		
	Examiner 1	Examiner 2	Mean	Examiner 1	Examiner 2	Mean
von Graefe	-6.86 ± 4.62	-6.89 ± 3.80	-6.87 ± 4.22[a]	-6.99 ± 4.93	-7.06 ± 4.03	-7.03 ± 4.49[a]
Howell card	-3.55 ± 2.67	-3.85 ± 2.85	-3.70 ± 2.77[b]	-3.53 ± 2.80	-3.95 ± 3.06	-3.74 ± 2.93[b]
Modified Thorington	-4.41 ± 3.23	-4.57 ± 3.43	-4.50 ± 3.34[b]	-4.38 ± 3.22	-4.62 ± 3.67	-4.50 ± 3.45[b]
RBP card	-3.48 ± 3.26	-3.36 ± 3.17	-3.42 ± 3.22[b]	-3.69 ± 3.37	-3.65 ± 3.31	-3.67 ± 3.33[b]

All units are in prism diopters.
Plus sign denotes esophoria, and minus sign denotes exophoria.
[a,b] : a subgroup by Duncan test of one-way ANOVA ($p < 0.05$)

in the scatterplots of Fig. 2 through 5. The figures were constructed by plotting the mean phoria of two examiners for each subject against the difference between the examiner results. Horizontal lines are drawn at the mean difference, and the dashed lines show 95% limits of agreement. The agreement limits were narrowest for the modified Thorington test calculated as ±2.9 Δ while the von Graefe method calculated as ±6.1 Δ had the widest agreement limits. The agreement limits of the RBP card test and the Howell card test were ±3.1 Δ.

Figures 6 through 9 show the mean difference comparisons of test-retest repeatability in phoria. The test and retest were performed by the same examiners and by the same procedure to the same subjects. Horizontal lines are drawn at the mean difference, and the dashed lines show 95% limits of agreement. The agreement limits were narrowest for the RBP card test calculated as ±0.91 Δ while the von Graefe method calculated as ±1.90 Δ had the widest agreement limits. The agreement limits of the modified Thorington test was ±1.27 Δ and the Howell card test was ±1.36 Δ.

Table 2 shows mean and standard deviation of phoria measurements, the reliability by ICC, and test-retest repeatability for the flashed presentation and the auto
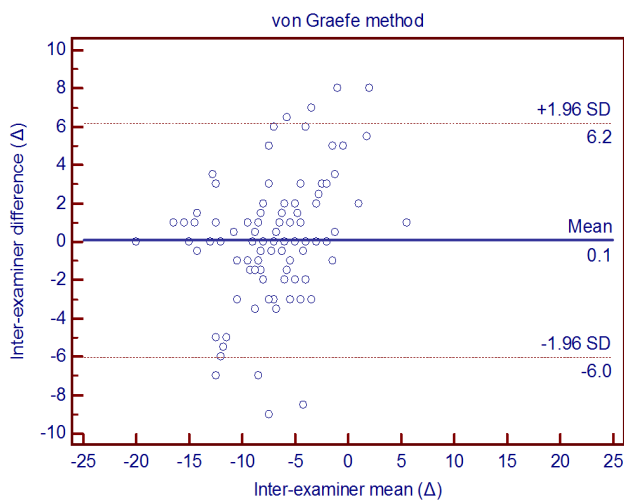
FIG. 2. Inter-examiner difference plotted against the inter-examiner mean for the continuous presentation of the von Graefe method. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.
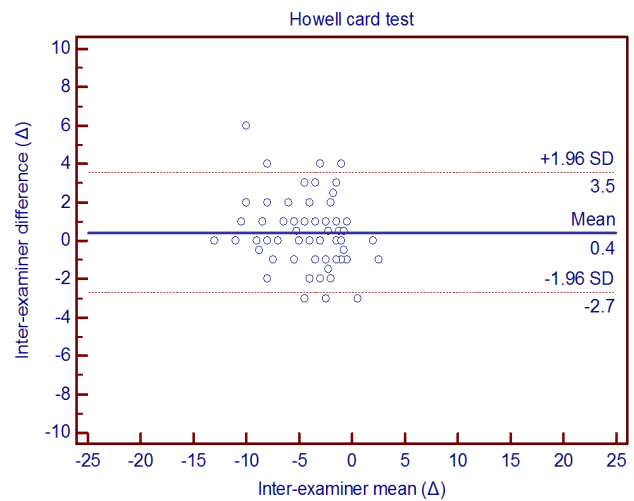
FIG. 3. Inter-examiner difference plotted against the inter-examiner mean for the flashed presentation of the Howell card. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.
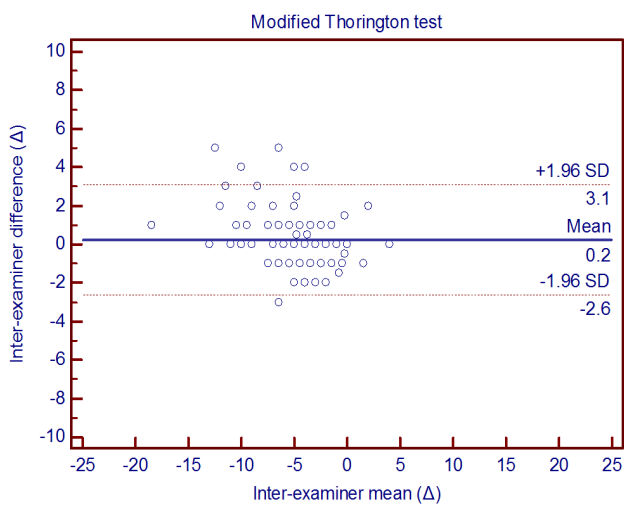
FIG. 4. Inter-examiner difference plotted against the inter-examiner mean for the flashed presentation of the modified Thorington's MIM (Bernell Muscle Imbalance Measure) card. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.
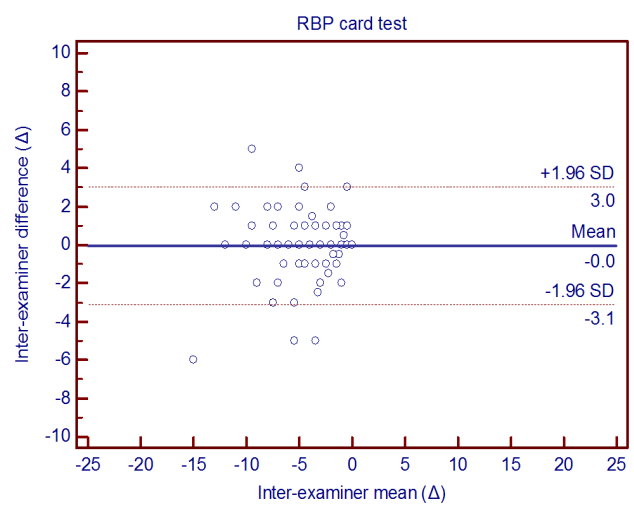
FIG. 5. Inter-examiner difference plotted against the inter-examiner mean for the flashed presentation of the RBP card. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.
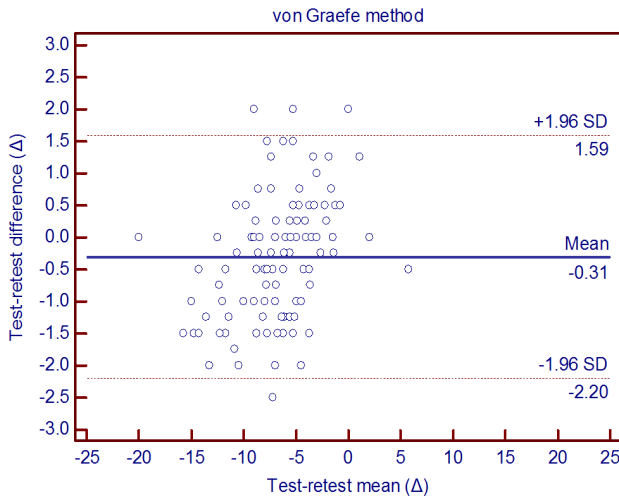
FIG. 6. Test-retest difference plotted against the test-retest mean for the continuous presentation of the von Graefe method. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.
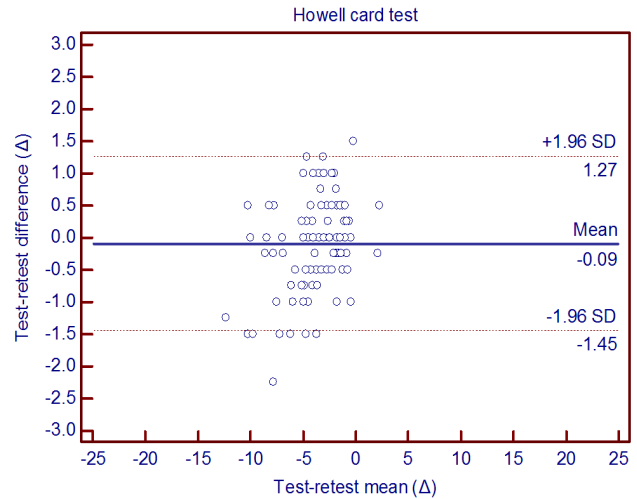

FIG. 7. Test-retest difference plotted against the test-retest mean for the flashed presentation of the Howell card. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.
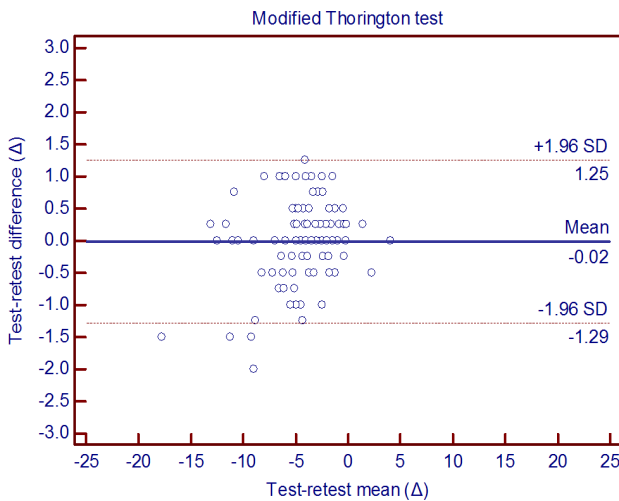

FIG. 8. Test-retest difference plotted against the test-retest mean for the flashed presentation of the modified Thorington's MIM (Bernell Muscle Imbalance Measure) card. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.
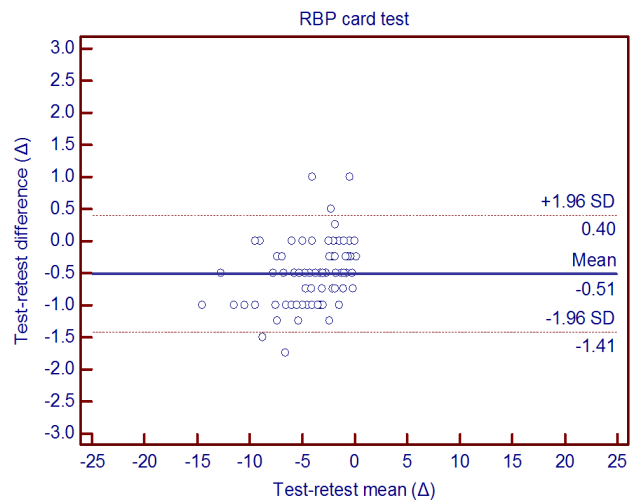

FIG. 9. Test-retest difference plotted against the test-retest mean for the flashed presentation of the RBP card. The dashed lines show 95% limits of agreement as ±1.96 · S.D. of mean difference.

TABLE 2. Mean ± standard deviation, ICC, and test-retest repeatability for flashed and auto-flashed presentation of RBP card test (n=26)

Presentation	Phoria measurement			ICC	Test-retest repeatability	
	1st	2nd	Mean		Mean difference	±1.96 · SD
Flashed	-5.36 ± 3.29	-4.95 ± 3.16	-5.15 ± 3.20	0.932	-0.4	±3.20
Auto-flashed	-5.12 ± 2.88	-4.39 ± 2.46	-4.75 ± 2.68	0.919	-0.7	±2.85

All units of mean ± standard deviation for phoria measurement are in prism diopters.
Plus sign denotes esophoria, and minus sign denotes exophoria in phoria measurement.

-flashed presentation of the RBP card test. The mean of the auto-flashed presentation was 4.75 Δ and the flashed presentation was 5.15 Δ. The ICC of the auto-flashed

presentation (0.919) was smaller than those of the flashed presentation (0.932), but the agreement limits of the auto-flashed presentation (±2.85 Δ) was narrower than those of

the flashed presentation (±3.20 Δ).

IV. DISCUSSION

Studies on the reliability and repeatability of measuring dissociated phoria have been conducted and methods of measuring dissociated phoria were varied and the results of reliability and repeatability were slightly different according to the researchers [1, 6-11, 14-19]. The reliability and repeatability of phoria measuring methods may be affected by several factors. Examiner and patient bias, instrument calibration, inconsistent testing technique or testing environment will result in variations among measurements [20, 21]. The reliability of the near modified Thorington test is greater than those of the near von Graefe method according to results from Hirsch and Bing [11] and from Morris [22]. Morris suggested that the reliability of the von Graefe method is greater than those of the Maddox rod test. The 95% limits of agreement were 2.0 Δ or less for the near modified Thorington test, 2.5 to 3.3 Δ for the near von Graefe test, and 1.8 Δ to 4.6 Δ for the near Maddox rod test. Wong et al. [1] have documented that the Howell phoria card continuous presentation and the MIM card (both continuous and flashed presentations) had significantly less validity in inter-examiner differences than the Howell phoria card flashed presentation and the von Graefe method. In results of inter-examiner repeatability of the heterophoria test from Rainey et al. [10], the modified Thorington test had the highest inter-examiner correlation, and the smallest 95% limits of agreement, suggesting it was the most repeatable of the methods evaluated. The subjective cover test also had a high correlation and a small 95% limits of agreement range [10]. The repeatability of the distance Howell card test is greater than that of the distance von Graefe method and distance modified Thorington test [11].

This experimental result was similar to that of a previously published study [17], the von Graefe method had the highest mean phoria and the RBP card test had the smallest mean phoria. The von Graefe method had the highest mean deviation and the Howell card test had the smallest mean deviation. The variation in inter-examiner repeatability, the 95% limits of agreement (upper limit/ lower limit), exhibited the widest variation for the von Graefe method (+6.2/-6.0) and showed the narrowest variation for the modified Thorington test (+3.1/-2.6). In the Howell card test (+3.5/-2.7) and the RBP card test (+3.0/-3.1), the variation of two tests was similar practically. The variation in test-retest repeatability, the 95% limits of agreement showed the widest variation for the von Graefe method (+1.6/-2.2) and the narrowest variation for the RBP card test (+0.4/-1.4). The modified Thorington test (+1.3/-1.3) was narrower than the Howell card test (+1.3/-1.5) in the variation. Results, in the 95% limits of agreement, show that the RBP card test is considered a repeatable method for phoria measurement and is a useful method clinically.

No current mobile or online applications available at this time provide any testing for phoria measurements. In the case of the RBP card test, if the alternating cover technique can be achieved without examiners, subjects can measure their phoria themselves by just wearing a red/blue filter, unlike the case of that the von Graefe method and the Howell card test require a prism lens and the modified Thorington test requires a Maddox rod for phoria measurement. The auto-flashed version of the RBP test showed flashed presentation automatically on a computer monitor not executed by an examiner but by the program. The mean of phoria measurement using the auto-flashed presentation was 0.4 Δ smaller than those of the flashed presentation test. With coefficients greater than 0.90, the reliability of phoria measurement is generally good [19]. Results that the ICC of average measures was 0.919 and the 95% limits of agreement was +2.1/-3.6 for the auto-flashed presentation of the test indicate the possibility of phoria testing in a mobile application or online using a RBP card.

V. CONCLUSION

A phoria test in optometric examination plays an important part in the diagnosis of patient symptoms such as blur, headache, eyestrain, or diplopia. In this study, a new phoria test using a flashed presentation with a RBP card was assessed as having a high repeatability and the auto-flashed version made by PowerPoint showed high reliability. Clinically, the new phoria test is a useful method and has potential application in mobile or online phoria measurement.

REFERENCES

1. E. P. F. Wong, T. R. Fricke, and C. Dinardo, "Interexaminer repeatability of a new, modified prentice card compared with established phoria tests," Optom. Vis. Sci. **79**, 370-375 (2002).
2. J. R. Griffin, *Binocular Anomalies: Procedures for Vision Therapy*, 2nd ed. (Chicago, Professional Press, USA, 1982), pp. 140-147.
3. D. A. Goss, *Ocular Accommodation, Convergence and Fixation Disparity: A manual of Clinical Analysis,* 2nd ed. (Butterworth-Heinemann, Boston, USA, 1995), pp. 11-12, p. 70.
4. P. A. Howarth and G. Heron, "Repeated measures of horizontal heterophoria," Optom. Vis. Sci. **77**, 616-619 (2000).
5. E. C. Casillas and M. Rosenfield, "Comparison of subjective heterophoria testing with a phoropter and trial frame," Optom. Vis. Sci. **82**, 237-241 (2006).
6. D. A. Goss, B. J. Moyer, and M. C. Teske, "A comparison of dissociated phoria test findings with von Graefe phorometry & modified Thorington testing," J. Behavioral Optom. **19**, 145-149 (2008).
7. M. J. Hirsch, "Clinical investigation of a method of testing phoria at forty centimeters," Am. J. Optom. Arch. Am. Acad.

Optom. **25**, 492-495 (1948).

8. B. B. Rainey, T. L. Schroeder, D. A. Goss, and T. P. Grosvenor, "Reliability of and comparisons among three variations of the alternating cover test," Ophthalmic Physiol. Opt. **18**, 430-437 (1998).

9. K. M. Daum, "Analysis of seven methods of measuring the angle of deviation," Am. J. Optom. Physiol. Opt. **60**, 46-51 (1983).

10. B. B. Rainey, T. L. Schroeder, D. A. Goss, and T. P. Grosvenor, "Inter-examiner repeatability of heterophoria tests," Optom. Vis. Sci. **75**, 719-726 (1998).

11. M. J. Hirsch and L. B. Bing, "The effect of testing method on values obtained for phoria at forty centimeters," Am. J. Optom. Arch. Am. Acad. Optom. **25**, 407-416 (1948).

12. H. Calvin, P. Rupnow, and T. Grosvenor, "How good is the estimated cover test at predicting the von Graefe phoria measurement?," Optom. Vis. Sci. **73**, 701-706 (1996).

13. J. M. Bland and D. G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," Lancet **397**, 307-310 (1986).

14. R. G. Scobee and E. L. Green, "Tests for heterophoria-reliability of tests, comparisons between tests, and effect of changing testing conditions," Trans. A. Acad. Ophthalmol. Otolarlyngol. **51**, 179-197 (1947).

15. F. W. Weymouth, "An experimental comparison of three common methods of measuring heterophoria," Am. J. Optom. Arch. Am. Acad. Optom. **40**, 497-503 (1963).

16. K. B. Lee, S. W. Jeon, H. J. Lee, S. H. Lee, W. B. Park, K. J. Mar, and S. B. Park, "The comparative analysis of various distance phoria tests," Korean J. Vis. Sci. **9**, 115-125 (2007).

17. S. H. Lee, S. Y. Park, S. Y. Kim, D. Y. Lee, K. K. Kim, J. S. Son, J. Y. Lee, and H. G. Cho, "Repeatability of new phoria test using color chart and color filter," J. Korean Oph. Opt. Soc. **15**, 269-274 (2010).

18. D. A. Goss, J. L. Reynolds, and R. E. Todd, "Comparison of four dissociated phoria tests: reliability & correlation with symptom survey scores," J. Behavioral Optom. **21**, 99-104 (2010).

19. T. L. Schroeder, B. B. Rainey, D. A. Goss, and T. P. Grosvenor, "Reliability of and comparisons among methods of measuring dissociated phoria," Optom. Vis. Sci. **73**, 389-397 (1996).

20. M. W. Morgan, "The reliability of clinical measurements with special reference to distance heterophoria," Am. J. Optom. Arch. Am. Acad. Optom. **32**, 167-179 (1955).

21. J. T. Thomson and D. L. Guyton, "Ophthalmic prisms. Measurement errors and how to minimize them," Ophthalmology **90**, 204-210 (1983).

22. F. M. Morris, "The influence of kinesthesis upon near heterophoria measurements," Am. J. Optom. Arch. Am. Acad. Optom. **37**, 327-351 (1960).