A COMPARISON OF TWO AUTOMATED DEVICES THAT MEASURE REFRACTIVE ERROR

by

Ashley Jacobs

A Thesis submitted to the

School of Graduate Studies

in partial fulfillment of the requirements for the degree of

Master of Science

Psychology Department

Memorial University of Newfoundland

May 2017

St. John's, Newfoundland and Labrador

Abstract

The present study compared the reliability, validity, and screening effectiveness of the Welch Allyn SureSight Autorefractor (WASS) and the PlusoptiX S09 Vision Screener (PS09). Eighty-nine children attending appointments at a pediatric ophthalmology clinic were tested twice each with the PS09 and the WASS. Each child then completed the gold standard examination of refractive error, cycloplegic retinoscopy, with one of two pediatric ophthalmologists. Refractive error scores from the two devices were compared to cycloplegic retinoscopy. Results indicated that the PS09 yielded better reliability than the WASS on both spherical refractive error (Coefficients of Repeatability [CORs] = 1.21D v. 1.63D) and cylindrical refractive error (CORs = 0.50D v. 0.58D). Although the PS09 yielded better agreement with cycloplegic retinoscopy than did the WASS on spherical refractive error (CORs = 3.53D v. 4.19D), the validity of both devices was quite poor. Furthermore, both devices significantly underestimated hyperopia. Compared to the PS09, the WASS yielded slightly better agreement with cycloplegic retinoscopy on cylindrical refractive error (CORs = 0.87D v. 1.06D). In terms of screening effectiveness, the WASS yielded superior sensitivity (WASS = 69%; PS09 = 46%), but the PS09 yielded superior specificity (PS09 = 90% v. 54%). These results demonstrate the impact of the selection of pass/fail criteria, and therefore, the screening accuracy of each device was also calculated. The PS09 was the more accurate device (PS09 = 66%; WASS = 62%). In all, the analyses suggest that the PS09 is the superior device, but only by a very small margin.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my Supervisor Dr. James Drover for the continuous support of my Masters study and research. His patience, and immense knowledge made the research, testing and writing of this thesis possible.

Additionally, I would like to thank the rest of my thesis committee: Dr. Brent Snook and Dr. Russ Adams for their insightful comments and suggestions.

Finally, my sincere thanks go to Dr Bautistia and Dr Bramwell for allowing me to test in their clinic for eight months. Also, a special thanks to Alick Tsui and Dana French for providing me a space to test and updates on the clinics schedule.

Table of Contents

List of Tables

List of Figures

A Comparison of Two Automated Devices that Measure Refractive Error

The newborn visual system is extremely underdeveloped and consequently, their visual abilities are far from mature (Pan et al., 2009). Whereas normal adult visual acuity is approximately 20/20 in Snellen notation, a newborn is legally blind with visual acuity of approximately 20/800 (Cavallini, et al., 2002; Maurer & Lewis, 2001). The typical newborn is also hyperopic (i.e., farsighted) as the eye is too short to match the focal power of the lens and cornea and therefore, images are focused behind the retina (Friedburg & Klöppel, 1996). The first decade of life constitutes a sensitive period of development marked by dramatic improvement in functional vision. Visual acuity improves to adult levels at roughly 6 years of age (Drover et al., 2008). By approximately 6 to 8 years of age, the eye has elongated such that its length matches its focal power, producing perfect optics or emmetropia (Adams, Dalton, Murphy, Hall, & Courage, 2002; Pan et al., 2009). Underlying this functional improvement is an increase in the number of synaptic connections from the eye to the visual cortex. The formation of these connections can be disrupted by visual deficits (Ciuffreda, Levi & Selenow, 1991). If a visual deficit is present during this sensitive period, new connections may not form and existing connections may regress (Maurer, Lewis, & Brent, 1989; Odom, Hoyt, & Marg, 1981). If the deficit is not treated during the sensitive period, it may lead to amblyopia, a permanent loss of visual acuity (Ciuffreda, Levi & Selenow, 1991).

Amblyopia is a cortical disorder characterized by a reduction in vision in the absence of any detectable optical or retinal abnormalities (Ciuffreda, Levi & Selenow 1991). Although it can exist bilaterally (i.e., in both eyes), it is usually unilateral (Ciuffreda, Levi, & Selenow, 1991). Amblyopia results from a disruption in normal visual

experience due to an early visual insult, i.e., amblyogenic factor. There are three types of amblyopia categorized by the type of amblyogenic factor. Strabismic amblyopia is the most common form of amblyopia, and is caused by strabismus (i.e., a misalignment of the two eyes; Cuiffreda, Levi, & Selenow, 1991). One or both eyes may be misaligned in either an inward or outward direction. Thus, the eyes receive a different visual image, which can lead to double vision (Economides, Adams, & Horton, 2012). Image degradation amblyopia results from an optical obstruction, such as a cataract (a cloudy opacity in the lens) or ptosis (the drooping of the upper eyelid; Smith et al., 2007). In either case, the obstruction prevents the formation of a sharp image on the retina. Further, anisometropic amblyopia is caused by anisometropia (i.e., a large difference in refractive error between the two eyes). There are two types of refractive error. Spherical refractive error refers to a difference in the focusing power of the eye and its focal length (i.e., the length of the eye). If the focal length is too short for the focusing power of the eye, one is hyperopic. If the focal length is too long for the length of the eye, one is myopic. Cylindrical refractive error refers to astigmatism in which the cornea is misshapen causing different degrees of focusing power along one of the angles or meridians of the cornea (Eva, Pascoe & Vaughan, 1982). Anisometropic amblyopia can be caused by differences in either spherical or cylindrical refractive error between the eyes (Joly & Frankó, 2014).

In each type of amblyopia, the image formed on one or both eyes is/are suppressed by the central nervous system, and cortical connections are lost or not formed at all. The longer the period of deprivation, the greater the deficit and the more likely that the deficit will be permanent (Joly, & Frankó, 2014). This highlights the importance of

early vision screening to allow prompt detection, and subsequently, treatment of these disorders before the deficits are permanent. Indeed, early detection and treatment of amblyopia tends to yield better prognoses (Friedburg & Klöppel, 1996; Kaur et al., 2016; Pediatric Eye Disease Investigator Group, 2005; Sanchez, Ortiz-Toquero, Martin, Juan, 2016; Tailor, Bossi, Greenwood, Dahlmann-Noor, 2016).

The measurement of refractive error is arguably the most effective way to detect amblyopia (Rotsos, Grigoriou, Kokkolaki, & Manios 2009). The gold standard test for refractive error measurement is cycloplegic retinoscopy (Rajavi et al., 2015). Following this procedure, cycloplegic eye drops are administered to temporarily paralyze the muscles controlling the lens to prevent accommodation. A retinoscope is then used to shine light into the eye, and the movement of the reflected light and its shadow on the retina are observed to measure refractive error. Importantly, cycloplegic retinoscopy requires the expertise of an optometrist or ophthalmologist and therefore, is not feasible in the context of vision screening, which is typically done by lay people with no specific vision training (e.g., a public health nurse). Another disadvantage of cycloplegic retinoscopy is that children dislike the cycloplegic drops and find the experience unpleasant (Freedmen & Preston, 1992).

In light of these limitations, researchers have developed two technologically advanced devices, namely, photoscreeners and autorefractors that provide objective measures of refractive error. Autorefractors direct a low intensity beam of infrared light into the participant's optical system as he/she fixates on a target. The light reflects to the device to determine the extent to which it is out of focus, thereby providing an automatic measure of refractive error (Huang et al., 2013). Early versions of autorefractors were

large tabletop devices that were suitable for hospitals and clinics to provide initial estimates of refractive error, but immobile and impractical for vision screening. Therefore, smaller and more portable handheld devices such as the Nikon Retinomax were developed (Cordonnier & Kallay, 2001). Alternatively, photoscreeners consists of a camera and a flash source. The participant must fixate on the flash while an image is captured of the reflected light returning from his/her optical system. The size and position of the reflected light on the participant's pupil is analyzed to determine the type and degree of refractive error (Lowry, Wang & Nyong'o, 2014). Early versions of photoscreeners were somewhat crude as pictures were taken on Polaroid film, and the area of the reflected light was measured using a ruler (Watts, Walker, Beck, 1999). Recent versions of photoscreeners are more advanced as they obtain digital images of reflected infrared light that are then analyzed automatically by specialized computer software to provide estimates of refractive error (Arthur, Riyaz, Rodriguez, & Wong, 2009).

Autorefractors and photoscreeners possess three important advantages over the traditional method of assessing functional vision (i.e., visual acuity testing). First, the child must simply stare at a visual display during testing and thus, there are no cognitive requirements. This makes the procedure suitable for children of all ages, including infants and toddlers. Second, whereas the visual acuity testing typically requires three minutes or more to complete, testing with autorefractors and photoscreeners requires less than one minute. Third, the tests are completely automated and therefore, are objective. As such, little tester expertise is required and test scores are not affected by shyness or other personality aspects of the child.

Early versions of both photoscreeners (e.g., MTI photoscreener, Otago photoscreener, iScreen Photoscreener, Eyecor Photoscreener, Kodak DC photoscreener, Visiscreen 100 photoscreener) and autorefractors (e.g. the Nikon Retinomax) have been evaluated in several vision screening studies. As indices of effectiveness, these studies typically report two dependent measures, sensitivity and specificity. Sensitivity refers to the percentage of children with a disorder who are identified accurately by the device. High sensitivity is essential to ensure that those who have a disorder receive prompt treatment. Specificity refers to the percentage of those with normal vision who are identified correctly by the device as having normal vision. High specificity is important to prevent over referrals, that is, healthy children being referred unnecessarily for eye exams. A summary of results from these studies are provided in Table 1 (Arnold Armitage, 2014; Barry & König, 2001; Cooper et al., 1996; Cordonnier & Kallay, 2001; Cordonnier & Dramaix 1998; 1999; Enzenauer, Freeman, Larson & Williams, 2000; Freedman & Preston, 1992; Granet, Hoover, Smith, Brown, Batsch, & Brody, 1999; Guo, Jia & Guo, 2000; Kennedy & Sheps, 1989; Kennedy, Sheps & Bagaric, 1995; Kennedy & Thomas, 2000; Morgan & Johnson 1987; Ottar, Scott & Holgado, 1995; Silbert, Noelle, & Matta, 2013; The VIP Study Group, 2011; Watts, Walker & Beck, 1999; Weinand, Graf & Demming, 1998). Although there is a broad range in measures, the devices yield mean sensitivity from 70% to 81 %, and mean specificity from 85% to 88%.

Table 1.

Mean sensitivity and specificity for early photoscreeners and an early autorefractor (Nikon Retinomax).

Test	No. of Studies	Mean Sensitivity	Mean Specificity
Photoscreener	23	81 (54-100)	85 (52-99)
Retinomax Autorefractor	12	70 (52-80)	88 (58-98)

Note: Numbers in parentheses represent ranges

Recently, two modern devices have been developed to provide more objective and rapid automatic measures of refractive error, the Welch-Allyn SureSight Autorefrator (WASS; Skaneateles, N.Y., U.S.A.) and the PlusoptiX S09 Vision Screener (PS09; PlusoptiX, Nuremberg Germany). The WASS and the PS09 are shown in Figures 1 and 2, respectively. The WASS is a handheld device that emits continuous beams of infrared light into the participant's optical system as he/she fixates a circular pattern of lights. The reflected light is analyzed to determine the extent to which it is out of focus, thereby providing monocular measures of spherical and cylindrical refractive error. The PS09 is a digital photoscreener that consists of a computer connected to a video recorder. A low intensity infrared light is shone into the participant's eyes and a digital image is captured of the reflected light from both pupils. A linked computer contains software that analyzes the reflected light, providing automatic measures of refractive error in both eyes simultaneously. Note that because the device provides binocular measures of refraction, the eyes must be aligned, and therefore, estimates cannot be obtained in the case of a

misalignment of the eyes, i.e., strabismus.

Figure 1. *The Welch-Allyn SureSight Autorefractor*.

Figure 2**.** *The PlusoptiX S09 Vision Screener.*

Numerous studies have evaluated the WASS by comparing scores obtained with this device to those obtained with gold standard measures of refractive error including cycloplegic retinoscopy and cycloplegic refraction. Note that in cycloplegic refraction, cycloplegic drops are administered and refractive error is measured with an autorefractor. This technique also provides accurate estimates of refractive error (Kulp et al., 2011; Kulp et al., 2014). Collectively, these studies indicate that compared to cycloplegic retinoscopy and cycloplegic refraction, the WASS tends to overestimate the degree of myopia and underestimate the degree of hyperopia, an outcome that is referred to as the "myopic shift" (Buchner, Schnorbus, Grenzebach, Busse, 2004; Iuorno, Grant, Noel, 2004; Jost et al., 2014; Kemper, Keating, Jackson, Levin, 2005; Kulp et al., 2011; Kulp et al., 2014). In addition, although estimates of cylindrical refractive error are more accurate, it does tend to overestimate the degree of astigmatism (Kulp et al., 2011; Rowatt, Donahue, Crosby, Hudson, Simon, 2007). Finally, cases of anisometropia have

the potential of going undetected, but the WASS will identify most cases of visual impairment (Harvey, Dobson, Miller, Clifford, Donaldson, 2009; Rowatt et al., 2007).

A number of studies have also evaluated the effectiveness of the WASS for screening for amblyopia and/or amblyogenic factors (Buchner et al., 2004; Iurno Grant, Noel, 2004; Rogers, et al., 2008; Schmidt et al., 2004, The VIP Study Group, 2005a; 2005b**)**. These studies are summarized in Table 2 below, as well, data are presented for the PS09, other versions of the PlusoptiX Photoscreener (the PlusoptiX S04 [PS04], the PlusoptiX S08 [PS08], the PlusoptiX S12 [PS12]), and PlusoptiX autorefractors (PA09, PA12; the PlusoptiX data are discussed below). In all, the WASS studies yielded a mean sensitivity of 70% and a mean specificity of 75%. There is however, a great deal of variability. Specifically, the sensitivity of the WASS ranges from 35% to 97%, whereas specificity ranges from 5% to 94%. It is important to note that the broad ranges on these measures are due, in large part, to the different pass/fail criteria chosen to classify children as positive or negative for amblyogenic factors (Rowatt et al.2007; Silverstein, Lorenz, Emmons, Donahue, 2009; Ying et al., 2005). Specifically, if the criteria are lenient, the test will be easy to pass, leading to low sensitivity and high specificity. If the criteria are more conservative, the test is difficult to pass, leading to high sensitivity and low specificity.

Table 2.

Mean sensitivity and specificity scores for the WASS and PlusoptiX devices. Note that numbers in parentheses represent the range. Ranges are not provided for the PS09 as only one study has been conducted with this device.

Test	No. of Studies	Mean Sensitivity	Mean Specificity
WASS	9	70	75
		(35-97)	(5-94)
PS04/PS08/PS12	10	81	83
		(45-100)	(39-100)
PS09	1	88	96
PlusoptiX Autorefractors	6	88	81
		(75-98)	(68-97)

Given that the PS09 is a relatively new device, few studies have evaluated its validity in measuring refractive error. Lim, Bae and Shin (2014) compared refractive measurements using the PS09 to scores obtained using cycloplegic refraction in 134 children. There was a significant difference between the PS09 and cycloplegic refraction on spherical refractive error, but no difference on cylindrical refractive error. The PS09 overestimated the degree of myopia and underestimated the degree of hyperopia (i.e., myopic shift). Lim et al. (2014) also evaluated the screening effectiveness of the PS09 to amblyopia and amblyogenic factors. The PS09 had a sensitivity of 88% and a specificity of 96%. Thus, the device appears to be an effective method of detecting amblyogenic risk factors.

While there is limited research on the PS09, much more research has been devoted

to comparing refractive error estimates from previous generations of PlusoptiX photoscreeners (i.e., PlusoptiX S04 [PS04] and PlusoptiX S08 [PS08]) to estimates obtained from cycloplegic refraction. Collectively, this research indicates that the previous photoscreeners tend to underestimate the degree of hyperopia and/or overestimate the degree of myopia in children (Arnold & Armitage, 2014; Bloomberg & Suh, 2013; Demirci, Arsian, Ozsutcu, Eliacik, & Gulkilk, 2014; McCurry, Lawrence, Wilson & Mayo, 2013; Singman, Matta, Tian, & Silbert, 2013; Yilmaz Ozkaya, Alkin, Ozbengi & Yazici, 2015). That is, these devices are susceptible to the myopic shift. At the same time, estimates of cylindrical refractive error obtained using the previous generations of the PlusoptiX photoscreener are not significantly different from those obtained with cycloplegic refraction (Rajavi, Parsafar, Ramezani & Yaseri, 2012).

Several studies have also investigated the screening effectiveness of the PS04, the PS08, and the PS12, a new version of the PlusoptiX vision screeners (see Table 2; Arthur et al., 2009; Matta, Arnold, Singman, Silbert, 2011; Matta, Singman, & Silbert, 2010; Moghaddam et al., 2012; et al., 2012; Singman et al., 2013; Ugurbas, Alpay, Tutar, Sagdik, Ugurbas, 2001). Collectively, these studies yielded a mean sensitivity of 81% and a mean specificity of 83%. As with the WASS, these studies tend to report broad ranges of sensitivity (45% to 100%) and specificity (39% to 100%; Arthur et al., 2009; Matta, et al., 2011; Matta et al., 2010; Moghaddam et al., 2012; Singman et al., 2013; Ugurbas et al., 2001). However, once again it is likely that the broad ranges can be attributed to the different pass/fail criteria across studies.

Finally, Silbert, Matta, and Ely (2014) compared the WASS to the PlusoptiX A09 (PA09). The PA09 is an autorefractor that is very similar to the PS09. The medical

records of 216 children who had been tested with the WASS, the PA09, and cycloplegic retinoscopy were examined. The PA09 yielded a sensitivity of 89% and a specificity of 80%, while the WASS had a sensitivity of 89% and a specificity of 71%. The authors concluded that both devices are effective screening devices.

Despite the broad ranges of scores reported, the PlusoptiX devices yielded superior screening effectiveness compared to the WASS (see Table 2). Importantly however, no single study has compared the PS09 and WASS directly. The present study will directly compare the PS09 and the WASS on reliability, validity, and screening effectiveness to determine the superior device. These two devices have been chosen for two reasons. First, although the WASS is no longer manufactured, it is perhaps the most widely used hand-held autorefractor to date (see Table 2 for a list of research studies that have evaluated this device). Second, at the time of testing in this thesis, the PS09 represented the most recent in a line of PlusoptiX photoscreeners which have provided very promising screening results (Arthur et al, 2009; Matta et al., 2010; Moghaddam et al., 2012; Rajavi et al., 2012; Singman et al., 2013; Ugurbas et al., 2001). This is an important endeavor because as noted above, the measurement of refractive error is arguably the best way to detect amblyopia. Furthermore, both devices are expensive (WASS ~ $5000 Cdn; PS09 ~ $8000 Cdn) and thus, it is likely that only one would be purchased for a vision screening program. To assess the reliability of each device, children who are attending an exam with one of two pediatric ophthalmologists will be tested twice with both the WASS and the PS09. To determine validity, scores from each device will be compared to scores obtained by pediatric ophthalmologist using cycloplegic retinoscopy. Finally, measures of sensitivity and specificity for each device

will be determined using established pass/fail criteria, and comparing pass/fail classifications to formal diagnoses.

Methods

Participants

Participants included 89 children who were attending eye exams with one of two pediatric ophthalmologists at the Janeway Hospital in St. John's, NL. Ethical approval was obtained from The Interdisciplinary Committee on Ethics in Human Research before testing. An additional 39 participants were tested but excluded from the final analyses because they could not complete testing with the PS09 (n = 28), the WASS (n = 7), or both (n = 4). Children ranged from 0.7 to 12.9 years of age (M = 7.5 years, SD = 3.2 years).

Design

Each participant was tested twice each with the WASS and the PS09 Vision Screener. The order of the testing with the WASS and the PS09 was counterbalanced. Each child then underwent cycloplegic retinoscopy with one of the pediatric ophthalmologists. Note that testing with the two devices was typically brief, taking three to five minutes.

Materials and Procedure

Welch-Allyn SureSight Autorefractor. Each participant was tested with the Welch-Allyn SureSight (WASS; see Figure 1), a hand-held autorefractor that provides rapid estimates of spherical and cylindrical refractive error along with the axis of astigmatism. To use the WASS, the tester looks through an aperture and moves the device towards the front of the participant's face. The tester is guided to the 35 cm test distance

by the device's audible feedback system. A low powered infrared light is directed through the pupil into the child's eye. This light then reflects back to the device, which determines the extent to which it is out of focus, thereby providing estimates of refractive error along with the axis of astigmatism. In all, the device takes 5 to 7 rapid measurements of refractive error. The device also provides a measure of reliability for the set of measurements.

PlusoptiX S09 Vision Screener. Each participant was tested using the PlusoptiX S09 Vision Screener (PS09; see Figure 2), one of several automated photoscreening devices (i.e., PS04; PS08, PS12) manufactured by PlusoptiX. The PS09 consists of an infrared video recorder linked to a portable lightweight computer. At a test distance of 1.5m, the video recorder projects an infra-red light through the participant's pupils onto the retina. An image is then taken of both of the participant's eyes. If the participant has a refractive error, the reflected light forms a specific brightness pattern within the pupil. This image is then relayed to a computer, which contains software that estimates the type and degree of refractive error.

Statistical Analyses

Only right eye measures were used to assess reliability and validity. To determine the reliability of each device, coefficients of repeatability (COR) were calculated. The COR is $\pm$ 1.96 times the standard deviation of the differences scores calculated for test 1 - test 2 scores obtained with each device, and therefore provides the 95% limits of agreement for these differences. CORs were calculated instead of correlation coefficients because whereas the latter provide an index of association, they often do not provide an accurate measure of agreement, which is an essential requirement for a clinical tool. In

addition, CORs are reported in the same units as those obtained by the test (i.e., diopters), allowing them to be interpreted much more easily than correlation coefficients (Bland & Altman, 1986; Reeves, Wood, & Hill, 1991). To determine whether test 1 and test 2 scores differed with each device, test 1 and test 2 scores were compared directly. Because these data were not distributed normally (Kolmogorov-Smirnov, $p < 0.05$), these analyses were conducted using a Wilcoxon Signed-Rank Test.

To assess the validity of each device, refractive error measurements obtained on test 1 were compared to gold standard cycloplegic retinoscopy measurements obtained by one of the two participating ophthalmologists. Specifically, CORs were also calculated. In this instance, the COR is ± 1.96 times the standard deviation of the difference between participant's scores obtained with the device and the gold standard scores. As noted above, it provides the 95% limits of agreement. In addition, the refractive error measurements obtained with each device were compared directly to the gold standard measurement. Once again, as these data were not distributed normally (Kolmogorov-Smirnov, $p < 0.05$), Wilcoxon Signed-Rank Tests were conducted.

Finally, to assess the screening effectiveness of each device, each participant was graded as pass/fail on test 1 based on standard criteria (see Table 3, Nathan & Donahue, 2010; Schmidt et al., 2004), and these results were compared to the final refractive error diagnosis based on cycloplegic retinoscopy following the standard criteria of Donahue, Arnold, Ruben, and the AAPOS Vision Screening Committee (2003). The sensitivity, specificity, and accuracy of each device were then calculated. As noted above, sensitivity refers to the proportion of those with disorders who are correctly identified by the device. Specificity refers to the proportion of those with healthy vision who are correctly

identified by the device. Lastly, accuracy refers to the proportion of classifications (pass/fail) based on the device, that agree with the ophthalmologists' diagnosis.

Table 3.

Pass/fail criteria in diopters (D) for the WASS and the PS09. These criteria are based on Nathan and Donahue (2010) and Schmidt et al. (2004).

Device	Hyperopia	Myopia	Astigmatism	Anisometropia
WASS	≥ 4.00	≤ -1.00	≥ 1.5	≥ 3.00
PS09	≥ 3.50	≤ -3.00	≥ 2.00	≥ 1.50

Results

Reliability

The means and standard deviations for each device's test 1 and test 2 scores are provided in Table 4 below. For the WASS, spherical refractive error measures were significantly more hyperopic on test 2 compared to test 1 (0.47D v 0.31D, $p = 0.017$). For cylindrical measures obtained with the WASS, test 1 scores were significantly higher than for test 2 (0.97D v. 0.87D, $p = .0033$). However, for the PS09, spherical measures on test 1 and test 2 were not significantly different (0.46D v. 0.43D respectively, $p = .61$). Similarly, cylindrical measures on test 1 and test 2 did not differ (0.79 v. 0.83, $p = 0.28$).

Table 4.

Mean scores for test 1 and test 2 in dioptres (D). Numbers in parentheses represent standard deviations.

PS09				WASS			
Sphere		Cylinder		Sphere		Cylinder	
Test 1	Test 2	Test 1	Test 2	Test 1	Test 2	Test 1	Test 2
0.46	0.43	0.79	0.83	0.31	0.47*	0.97	0.87*
(1.36)	(1.33)	(0.66)	(0.68)	(1.81)	(1.78)	(0.72)	(0.70)

*Test 1 and test 2 were significantly different ($p < 0.05$)

The CORs, which represent the level of agreement between test 1 and test 2 scores, are provided for each device in Table 5. The agreement between test 1 and test 2 scores for each device are illustrated in Bland-Altman plots in Figures 3 and 4. Bland-Altman plots are the standard method of representing the agreement between two clinical measures visually (Hanneman, 2008). Specifically, Bland-Altman plots illustrate the mean difference between the two measures, along with the 95% limits of agreement. The plots in Figures 3 and 4 show that the limits of agreement for spherical refractive error were 0.03 ± 1.21D for the PS09 and 0.16 ± 1.63D for the WASS (see Figure A). The 95% limits of agreement for cylindrical refractive error were 0.04 ± 0.50D for the PS09 and 0.10 ± 0.58D for the WASS. Thus, for both spherical and cylindrical refractive error, the PS09 yielded better reliability and agreement.

Table 5.

Coefficients of repeatability (COR) for test 1 and test 2 scores in dioptres (D).

PS09		WASS	
Sphere	Cylinder	Sphere	Cylinder
Test 1 v. Test 2	Test 1 v. Test 2	Test 1 v. Test 2	Test 1 v. Test 2
1.21	0.50	1.63	0.58

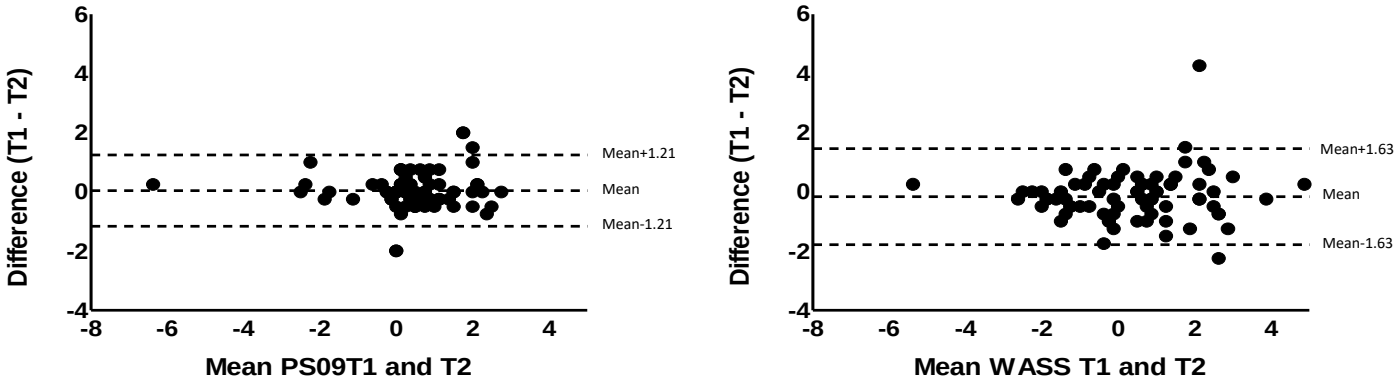

Figure 3. *The 95% limits of agreement between test 1 and test 2 scores on spherical refractive error for the PS09 and the WASS. All scores are in diopters.*

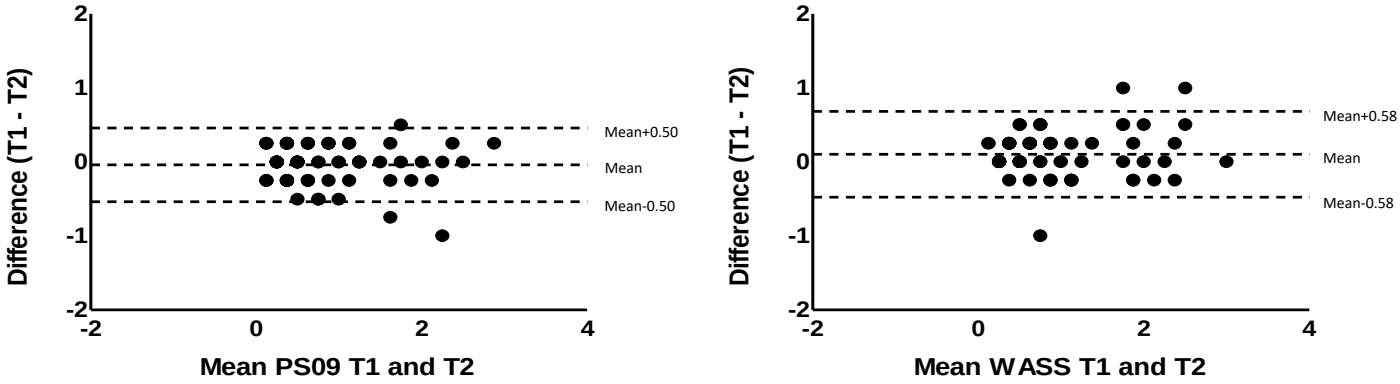

Figure 4. *The 95% limits of agreement between test 1 and test 2 scores on cylindrical refractive error for the PS09 and the WASS. All scores are in diopters.*

Validity

Means and standard deviations for refractive error measures on test 1 for each device and for the gold standard exam are provided in Table 6. Both the PS09 (0.53D, $Z = -6.38$, $p < .001$) and WASS (0.36D $Z = -5.93$, $p < .001$) provided spherical refractive error measures that were significantly different from the gold standard measure (2.04D). Specifically, each device underestimated the degree of hyperopia. The cylindrical refractive error scores for both the PS09 (0.78D; $Z = -2.70$, $p < .01$) and the WASS (0.95D, $Z = -5.71$, $p < .001$) were significantly higher than the gold standard (0.60D).

Table 6.

Mean Gold Standard and test 1 scores in dioptres (D). Standard deviations are in parentheses.

Gold Standard v. PS09				Gold Standard v. WASS			
Sphere		Cylinder		Sphere		Cylinder	
GS	PS09	GS	PS09	GS	WASS	GS	WASS
2.04	0.53*	0.6	0.78*	2.04	0.36*	0.6	0.95*
(2.25)	(1.37)	(0.66)	(0.65)	(2.25)	(1.79)	(0.66)	(0.73)

*Gold standard scores and scores from the device were significantly different ($p < 0.05$)

The CORs showing the agreement between each device and the gold standard are presented in Table 7. The Bland-Altman plots illustrating the 95% limits of agreement between each device and the gold standard on spherical and cylindrical refractive error are shown in Figures 5 and 6. The results show that compared to the WASS, the PS09 demonstrated better agreement with the gold standard exam on spherical refractive error (95% limits of agreement for spherical refractive error: 1.51±3.53D for the PS09; 1.68±4.19 for the WASS). Conversely, the WASS showed better agreement with the gold standard exam on cylindrical refractive error. The 95% limits of agreement on cylindrical refractive error was -0.18 ±1.06D for the PS09 and -0.25 ±0.87D for the WASS.

Table 7.

Coefficients of repeatability (COR) for Gold Standard and test 1 scores in dioptres (D).

Gold Standard v. PS09		Gold Standard v. WASS	
Sphere	Cylinder	Sphere	Cylinder
3.53	1.06	4.19	0.87

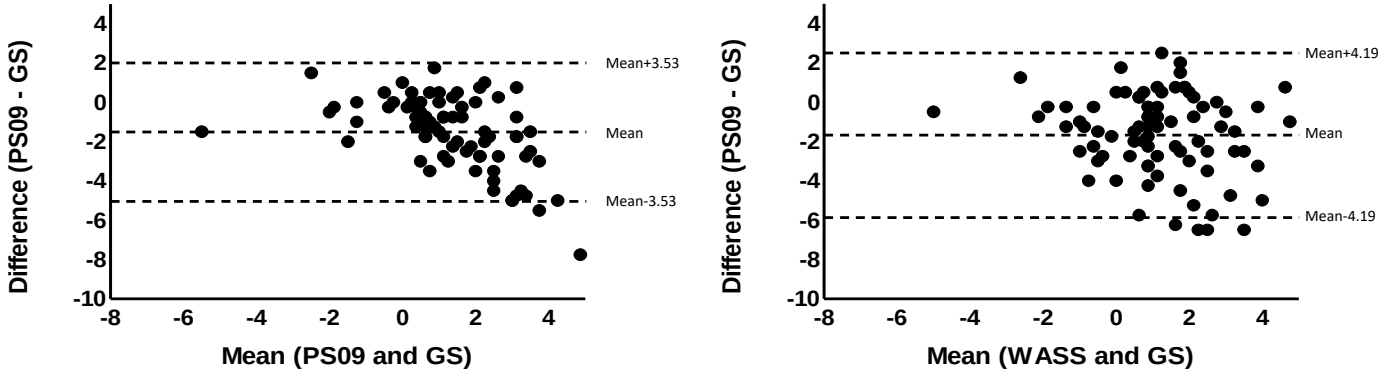

Figure 5. *The 95% limits of agreement between the PS09 and the gold standard exam (GS), and between the WASS and the gold standard exam on spherical refractive error. All scores are in diopters.*

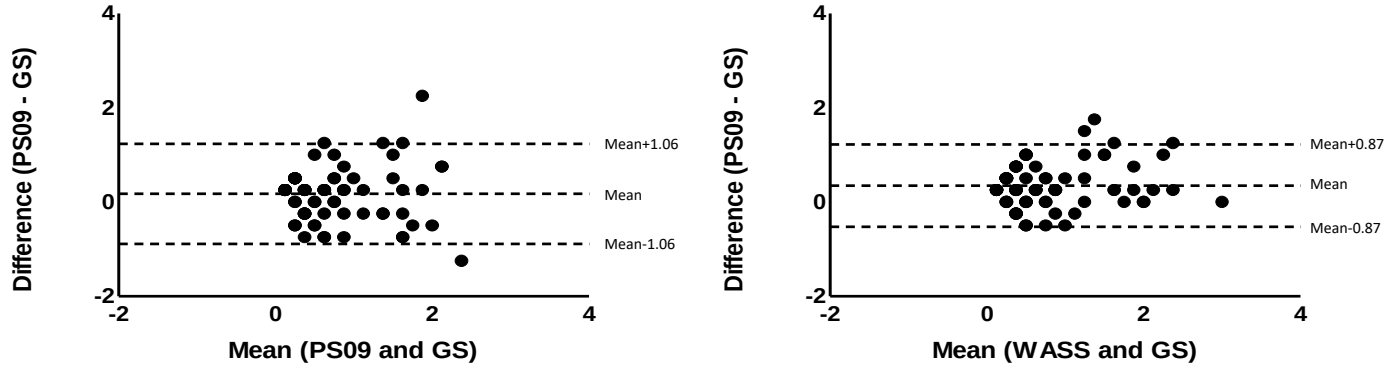

Figure 6. *The 95% limits of agreement between the PS09 and the gold standard exam (GS), and between the WASS and the gold standard exam on cylindrical refractive error. All scores are in diopters.*

Screening Effectiveness

The sensitivity, specificity, and accuracy of each device are provided in Table 8. Screening effectiveness analyses demonstrated that the sensitivity of the WASS was higher than that of the PS09 (69% v. 46%) indicating that the WASS detected more

participants with ametropia (i.e., significant refractive errors). Conversely, the specificity of the PS09 was higher than that of the WASS (PS09 = 90%, WASS = 54%). The specificity results suggest that the PS09 correctly identified a higher percentage of children with normal, healthy vision than did the WASS. Taken together, these results indicate that the differences in sensitivity and specificity might be related to the pass/fail criteria utilized for each device. That is, the criteria for the PS09 were lenient making the test easy to pass, and thus, sensitivity was poor and specificity was high. On the other hand, the pass/fail criteria for the WASS were strict making the test difficult to pass, allowing for better sensitivity, but poorer specificity. Thus, to limit the effect of pass/fail criteria, the accuracy yielded by each device was also calculated. The accuracy of PS09 was slightly higher than that of the WASS (66% vs 62%). The higher accuracy score suggests that pass/fail outcomes of a PS09 showed better agreement with the gold standard diagnosis than did WASS outcomes.

Several additional screening effectiveness analyses were also conducted. First, because test 2 refractive error measures obtained using the WASS were significantly different from test 1 refractive error measures, screening effectiveness scores were also calculated based on test 2 refractive error scores (see WASS Test 2 in Table 8). This analysis indicates that in comparison to test 1, screening effectiveness scores based on test 2 scores were superior on sensitivity (69% v. 70%, respectively), specificity (54% v. 63%, respectively), and accuracy (62% v. 67%, respectively).

Table 8.
Screening effectiveness of the WASS and the PS09. WASS test 1 denotes effectiveness measures based on refractive error scores from test 1. WASS test 2 denotes effectiveness measures based on refractive error scores from test 2.

Device	Sensitivity	Specificity	Accuracy
WASS Test 1	69	54	62
WASS Test 2	70	63	67
PS09	46	90	66

Next, given the broad age range tested in the present study and the fact the screening has been discussed in this theses in the context of early vision screening, a further analysis was conducted. Specifically, the sample was divided into two age groups; young children (5 years of age and younger; N = 32) and school-age children (i.e., 6 years of age and older; N = 57), and the screening effectiveness of the two devices was compared. In all, this analysis yielded the same general results as the overall analysis. Specifically, the WASS had better sensitivity than the PS09 (young children: WASS = 55%, PS09 = 36%; school-age children: WASS = 73%, PS09 = 49%), whereas the PS09 had better specificity than the WASS (young children: PS09 = 95%, WASS = 76%; school-age children: PS09 = 85%, WASS = 30%). Of the two devices, the PS09 yielded slightly higher accuracy (young children: PS09 = 75%, WASS = 69%; school-age children: PS09 = 61%, WASS = 58%).

Finally, the age-based screening effectiveness (i.e., young children and school-age children) of the WASS was determined based on test 2 refractive error scores. Compared to test 1 scores, test 2 scores yielded lower sensitivity (55% vs. 45%, respectively), higher specificity (76% vs. 81%, respectively), and equivalent accuracy (69%) for younger

children. Compared to test 1 scores, test 2 scores yielded higher sensitivity (73% vs. 78%, respectively), higher specificity (30% vs. 45%, respectively), and higher accuracy (58% vs. 66%, respectively) for school-age children.

Importantly, the age-based analysis reveals the same comparative trends as the overall analysis. That is, the WASS yielded higher sensitivity, whereas the PS09 yielded higher specificity and accuracy. However, it is noteworthy that compared to the overall analyses, these age-based analyses revealed lower sensitivity and higher specificity for both devices in younger children, and lower specificity for both devices in older children. Yet these results must be interpreted very cautiously as each group has very few participants for a screening study, and only 11 children in the young age group had ametropia, making it difficult to obtain a valid estimate of sensitivity. In light of this, these age-based results will not be discussed further.

Discussion

The present study is the first to compare the WASS and PS09 directly. Specifically, the devices were compared in terms of reliability, validity, and screening effectiveness. Both devices were designed to provide objective, noninvasive, noncycloplegic measures of refractive error. Screening for refractive error is important because it is arguably the most effective way to detect amblyopia (Rotsos et al., 2009). Given that both the WASS and PS09 are automated, they do not require the expertise of an eyecare expert, but can be used by testers with no formal training in vision testing. These latter two points suggest that the devices can potentially provide effective vision screening.

Reliability

Participants were tested twice with each device to determine reliability. The PS09 yielded better reliability than the WASS for spherical refractive error (CORs: 1.21D v. 1.63D, respectively). Furthermore, although test 1 and test 2 scores obtained using the WASS were significantly different, there was no significant difference between test 1 and test 2 scores obtained with the PS09. The reason for the superior reliability of PS09 is unclear. It is not likely due to testing conditions as for all participants, test 1 and test 2 were conducted in the same room under the same lighting conditions. Furthermore, given that the devices are automated and designed for use by lay screeners, it is unlikely that the differences are related directly to the tester. Instead, the difference in reliability may be due to a combination of test distance and a learning effect. During the measurement of noncycloplegic refractive error, participants often attempt to accommodate or focus on the device, changing the refractive power of the eye. This leads to an inaccurate measure of refractive error. Specifically, it leads to "myopic shift" in which the act of accommodating with the lens causes the device to underestimate hyperopia and/or overestimate myopia (Buchner et al., 2004; Iuorno, Grant, Noel, 2004; Kemper et al., 2005; Kulp et al., 2011 & 2014; Jost et al., 2014; Moghaddam et al., 2012). The tendency to accommodate has a greater effect on refractive error measurement with the WASS as it implements a much shorter test distance than the PS09 (35cm v 1.5m, respectively), and therefore, the eye must accommodate to a greater extent. There may also be a learning effect as participants become more relaxed and familiar with the procedure and the tester, thereby relaxing his/her gaze during the second test. In other words, the participant does not accommodate as much leading to less myopic shift. This explanation is plausible as

the results indicate that the spherical refractive error measures on test 2 were more hyperopic than measures on test 1 for the WASS (0.47D vs 0.31D, respectively).

On cylindrical refractive error, the reliability of the PS09 and WASS were very similar (CORs = 0.50D, 0.58D, respectively). Note that the CORs are much lower than those reported for spherical refractive error. This is not surprising as measures of cylindrical refractive error are not affected by accommodation (Rowatt et al., 2007). Once again, scores on test 1 and 2 obtained using the WASS were significantly different. This was not the case with the PS09, which demonstrated better reliability when compared to the WASS.

Surprisingly, the reliability of handheld autorefractors and of photoscreeners has rarely been evaluated. As such, it is difficult to compare the results reported here to those of other studies. Nevertheless, the CORs of the present study are similar to those reported for spherical and cylindrical refractive errors using the WASS, the Nikon Retinomax autorefractor, and the 2 win videorefractor (CORs for spherical refractive error range: 1.18 to 1.59; CORs for cylindrical refractive error range: 0.49D to 0.59D; Huang et al., 2013; Ogbuehi, Almaliki, AlQarni & Osuagwu, 2015). Ogbuehi et al. (2015) reported far lower CORs for both spherical and cylindrical refractive error (spherical refractive error = 0.70D; cylindrical refractive error = 0.44D). However, they tested participants with the Topcon KR8800, a large tabletop autorefractor. These devices possess superior technology over handheld devices due to their large size. Yet because of their size, they are not portable, and therefore, inappropriate for onsite vision screening.

Validity

To determine validity, test 1 scores obtained with each device were compared to estimates obtained from cycloplegic retinoscopy. On measures of spherical refractive error, the PS09 showed better agreement with cycloplegic retinoscopy than did the WASS (CORs = 3.53D v. 4.19D, respectively). Note however, that both CORs are very large and suggest that 95% of the scores are within approximately ± 3.50 to 4.25D of the gold standard measures. Both devices underestimated the amount of hyperopia by wide margins (Mean spherical refractive error scores: PS09 = 0.53D; WASS = 0.36D; cycloplegic retinoscopy = 2.04D). As noted above, this underestimation of hyperopia is due to natural accommodation leading to the myopic shift (Bushner, Schnorbus, Grenzebach, Busse, 2004; Iuorno, Grant, Noel, 2004; Kemper et al., 2005; Kulp et al., 2007; Jost et al., 2014; Moghaddam et al., 2012). Importantly, the PS09 yielded slightly better validity as it implements a greater test distance and is therefore, less affected by natural accommodation. However, it is noteworthy that the myopic shift reported here does not preclude the use of these devices for screening purposes, as the pass/fail criteria for the tests can be adjusted to account for the myopic shift.

Both devices demonstrated better agreement with cycloplegic retinoscopy on cylindrical refractive error than on spherical refractive error. This finding is well documented and as noted above, it is likely because cylindrical refractive error is not affected by accommodation (Iurno et al., 2004). According to the COR analyses, the WASS showed better agreement with cylindrical refractive error than did the PS09 (0.87D v. 1.06D). Interestingly, both devices overestimated cylindrical refractive error significantly. Although this is a common finding (Kulp et al., 2007; Rowatt et al., 2007),

there is no clear explanation for it. Once again, this overestimation does not preclude the use of these devices in the context of vision screening if appropriate pass/fail criteria are chosen.

Several other studies have evaluated the validity of handheld autorefractors and photoscreeners such as the WASS, the PA09, the PA12, the Nikon Retinomax K-Plus, the Nikon Retinomax K Plus-2, the Nikon Retinomax K-Plus 3, and the PS08 by comparing measures obtained using these devices to cycloplegic retinoscopy or cycloplegic refraction (Demirci et al., 2014; Fogel-Levin, Doron, Wygnanski-Jaffe, Ancri, 2016; Paff, Oedesluys-Murphy, Wolterbeck, Swart-van den Berg, de Nie et al., 2010; Payerols, Eliaou, Trezeguet, Villain, & Daien, 2016; Yan, Jiao, Xu, Li Wang, 2016; Yilmaz et al., 2015). Collectively, CORs for spherical refractive error range from 1.09D to 3.70D. The present COR for the PS09 lies near the upper end of this range, whereas the present COR of the WASS is well beyond this range. For cylindrical refractive error, these studies report CORs ranging from 0.63D to 1.80D. The CORs for both the PS09 and the WASS were well within this range. It is not clear why the devices yielded relatively poor validity on spherical refractive error estimates. It does not appear to be due to tester experience as in previous studies, lay screeners and nurses have obtained superior CORs (Paff, et al., 2010;Yan, Jiao, Xu, Li & Wang, 2016), while trained physicians have obtained similarly poor CORs (Yilmaz et al., 2015). It is possible that, as noted above in the Reliability subsection, participants were timid during the test and attempted to focus on the device. Indeed, second measures of spherical refractive error obtained with the WASS showed less myopic shift. As such, one might assume that the second measures of spherical refractive error were more valid. However, this is not the case as second measures of

spherical refractive error obtained using the PS09 and the WASS were also poor in terms of validity (CORs = 3.38 and 4.08, respectively). Thus, it is unclear why the validity of these devices, particularly the WASS, is so poor on estimates of spherical refractive error.

Screening Effectiveness

To determine screening effectiveness, each participant was classified as passing or failing screening with each device based on established pass/fail criteria (Nathan & Donahue, 2010; Schmidt et al., 2004). This classification was then compared to the gold standard diagnosis, which was based on the criteria of Donahue et al. (2003). A comparison of the two devices reveals that the WASS yielded higher sensitivity than the PS09 (69% v. 46%, respectively) suggesting that it identified more participants with ametropia. In fact, the PS09 identified fewer than half of those with refractive error. This has a dangerous consequence as parents of children with undetected disorders who passed vision screening are not likely to have their children examined by optometrists or ophthalmologists. Thus, the disorder will remain undetected and may worsen leading to permanent deficits. Of the two devices, the PS09 yielded higher specificity (90% v. 54%). This means that the PS09 correctly identified over 90% of those with normal vision, whereas the WASS correctly identified only slightly more than half of those with normal vision. Thus, in a vision screening context, the WASS would lead to a high number of over referrals. Consequently, eyecare specialists would have to perform unnecessary eye exams and the failed screening would be unnecessarily worrisome to parents of children who do not have disorders.

Because the inverse relationship between sensitivity and specificity often makes it difficult to compare different devices, the screening accuracy was also calculated for each

device. The accuracy of PS09 was slightly higher than that of the WASS (66% vs. 62%). The higher accuracy scores suggest that pass/fail outcomes of the PS09 showed better agreement with the gold standard diagnosis than did the WASS outcomes. Although this indicates that the PS09 is the superior screening device, the accuracy of the PS09 was still unimpressive as it correctly classified only two-thirds of the participants.

Given that test 1 and test 2 refractive error scores obtained with the WASS were significantly different, screening effectiveness of this device was also determined based on test 2 scores. This analysis indicated that the screening effectiveness of test 2 scores were slightly higher than those for test 1 scores, particularly in terms of specificity (sensitivity = 70% v. 69%; specificity = 63% v. 54%; accuracy = 67% v. 62%). In fact, four children who failed screening on test 1 (i.e., false positive), were classified correctly as having normal vision on test 2. Conversely, there were no false positives on test 2, who were correctly identified as having normal vision on test 1. The reason for the superior specificity of test 2 is not clear, especially given the fact that the test 1 false positives correctly identified by test 2 fell under four different classifications (the four classifications were hyperopia, anisometropia, myopia/astigmatism, astigmatism). Moreover, even though test 2 yielded higher specificity, two points must be noted. First, the scores based on test 2 are still only mediocre. In particular, the accuracy score indicates that the WASS and gold standard agree only on two-thirds of all cases. Second, given the time constraints involved in vision screening and the limited attention span of young children, a device that provides accurate measures on the first test is preferable to one that requires two or more tests to provide accuracy.

Table 9 below provides a summary of the screening effectiveness of the WASS and the PS09 from the present study and previous studies (Buchner et al., 2004; Iuorno, Grant & Noël, 2004; Jost et al., 2014; Lim, Bae & Shin, 2014; Rogers, Neely, Chapman, Plager, Sprunger, et al. 2008; Schmidt et al., 2004; Silbert, Matta & Ely, 2014; Silbert, Matta, Tian & Singman, 2014; The VIP Study Group, 2005a; 2005b), along with the performance of other PlusoptiX devices (Arthur et al., 2009; Bloomberg & Suh, 2013; Crescioni, Miller, Harvey, 2015: Demirci et al., 2014; Fogel-Levin et al., 2016; Lim, Bae & Shin, 2014; Matta, et al., 2011; Matta et al., 2010; Moghaddam et al., 2012; Paff et al., 2010; Rajavi et al., 2012; Silbert, Matta, & Ely, 2014; Silbert, Matta, Tian, & Singman, 2014; Singman et al., 2013; Rajavi et al., 2015; Ugurbas et al., 2001; Wang & Suh, 2012; Yan, Jiao, Li & Wang, 2016), an early autorefractor (i.e., the Nikon Retinomax; Barry & König, 2001; Cordonnier & Kallay, 2001; Cordonnier & Dramaix 1998; 1999; The VIP Study Group, 2011), early photoscreeners (Arnold & Armitage 2014; Cooper et al., 1996; Enzenauer et al., 2000; Freedman & Preston, 1992; Granet et al.,, 1999; Guo et al.,2000; Kennedy & Sheps, 1989; Kennedy et al., 2000; Morgan & Johnson 1987; Ottar, Scott & Holgado, 1995; Silbert, Noelle, & Matta, 2013; Tong et al., 2000; Tong, Macke, Bassin, Everett, EnkeMiyazake, et al. 2000; Wang & Suh, 2012; Watts et al., 1999; Weinand et al., 1998), and modern autorefractors and photoscreeners (Crescioni et al., 2015; Jost et al., 2014; Lavezzo, de Sousa, Kanamura & Scellini, 2010; Paff et al., 2010; Silbert & Matta, 2014; The VIP Study Group, 2011). Accuracy scores are not provided in the Table as they are rarely reported in previous studies. The sensitivity and specificity of the WASS reported here for both test 1 and test 2 are well within range of that reported by other studies investigating the screening effectiveness of this device (sensitivity = 35% to

97%; specificity = 5% to 94%). This is due in large part to the extremely broad range from the previous studies. The sensitivity of the WASS for both test 1 and test 2 also fell within range of the scores reported from studies of the Nikon Retinomax (52% to 80%), early photoscreeners (54% to 100%), PlusoptiX photoscreeners (45% to 100%), and modern autorefractors and photoscreeners (51% to 97%). Yet, the sensitivity of the WASS on test 1 and test 2 was below the range of that reported in studies using PlusoptiX autorefractors (75% to 98%). The specificity of the WASS for test 1 was very poor and was below the range reported from studies using the Nikon Retinomax (58% to 98%), those using modern photoscreeners and autorefractors (74% to 90%), and those using PlusoptiX autorefractors (68% to 97%). The specificity of the WASS for test 2 was slightly better as it was within the range reported by the Nikon Retinomax (58% to 98%), but below the range for modern photoscreeners and autorefractors (74% to 90%), and the PlusoptiX autorefractors (68% to 97%).

Table 9.

A summary of the performance of the WASS and the PS09 from the present study vs. previous studies, along with the performance of previous autorefractors, photoscreeners, and other PlusoptiX devices.

Test	No. of Studies	Mean Sensitivity	Mean Specificity
WASS Present Study (Test 1)	1	**68.8**	**53.7**
WASS Present Study (Test 2)	1	**70**	**63**
PS09 Present Study	1	**46**	**90**
Photorefraction	23	81 (54 - 100)	85 (52-99)
Autorefraction	12	70 (52-80)	88 (58-98)
WASS	9	70 (35-97)	75 (5-94)
PS04/PS08/PS12	10	81 (45-100)	83 (39-100)
PS09	1	88	96
PlusoptiX Autorefractors	6	88 (75-98)	81 (68-97)
Modern Autorefractors and Photoscreeners	6	77 (51-97)	84 (74-90)

These discrepancies are not likely due to the pass/fail criteria as the implementation of more lenient criteria in the present study would increase specificity, but would also reduce sensitivity. Furthermore, the criteria implemented here were the same as those utilized by the VIP Study group who reported higher sensitivity and specificity (Schmidt et al., 2004), albeit, with a much larger sample (N = 1452 v. 89). This latter point raises the possibility that the sample size may have contributed to the

poor screening effectiveness. Specifically, whereas the VIP Study Group conducted large population-based studies, the present study used an “enriched sample”, i.e., a small sample of outpatients attending a clinic, many of whom have ametropia. Still, this explanation is unlikely because whereas the studies that have investigated the screening effectiveness of the WASS using small enriched samples do often report low specificity (mean = 63%; Buchner et al., 2004; Iurno et al., 2004; Jost et al., 2014; Rogers et al., 2008; Silbert, Matta, & Ely, 2014; Silbert, Matta, Tian, & Signman, 2014), they tend to report at least moderate sensitivity (mean = 77%).

The effectiveness of the PS09 in the present study was poor compared to that reported by Lim et al (2014; sensitivity = 46% v. 88%, respectively; specificity = 90% vs. 96%, respectively). Overall, the specificity reported here was at the high end of the range reported from early photoscreeners (52% to 99%), from other PlusoptiX photoscreeners (39% to 100), from modern autorefractors and photoscreeners (74% to 90%), from PlusoptiX autorefractors (68% to 97%), from the Nikon Retinomax (58% to 98%), and from other studies using the WASS (5% to 94%). On the other hand, the sensitivity reported here was below the range reported from early photoscreeners (54% to 100%), from modern from autorefractors and photoscreeners (51% to 97%), from PlusoptiX autorefractors (75% to 98%), and from the Nikon Retinomax (52% to 80%). The poor sensitivity is not likely due to the small sample size and/or the use of enriched populations, as most studies evaluating the effectiveness of the PS09 and other PlusoptiX devices also used small enriched samples. A possible explanation for the poor sensitivity of the PS09 is that it was particularly ineffective at detecting hyperopia. In fact, in the present study, the PS09 detected only 5/28 participants with hyperopia (sensitivity =

18%). Indeed, PlusoptiX devices may be relatively ineffective at detecting hyperopia in general. This is difficult to determine, as few studies provide sensitivity scores to specific types of amblyogenic factors. However, Rajavi et al. (2012) reported that the PlusoptiX S04 yielded a sensitivity of only 45% for hyperopia. Moreover, creating more conservative pass/fail criteria to detect hyperopia, such as reducing the current pass/fail criteria of the PS09 for spherical refractive error (e.g. from ≥ 3.50D to ≥ 3.00D) would have detected no additional cases of hyperopia.

Limitations

While this study is the first to compare the PS09 and the WASS, there are limitations that must be addressed. First, the testability of the PS09 was poor as 32 participants were unable to complete testing and therefore, could not be included in the analyses. This is detrimental as it suggests that a relatively large percentage of children screened with this device would not be able to complete screening. If these children possess vision disorders, they would remain undiagnosed. Importantly, the poor completion rate of the PS09 is not necessarily due to an onerous testing procedure, instead, it is because it assesses refractive error binocularly. If the eyes are misaligned due to strabismus, the device is unable to obtain a measure. In the present study, 17/32 participants who could not complete testing with the PS09 were documented cases of strabismus. Considering this finding, it would be wise to refer children who are unable to complete testing with the PS09 as it is quite possible that they have strabismus. Alternatively, one could also use a test of stereoacuity along with the PS09 as such tests are relatively sensitive to strabismus (Ciner, et al., 2014). The testability of the WASS was better as a total of 11 participants were unable to complete testing. The majority of

these participants (n = 6) had hyperopia. As with the PS09, this result suggests that it might be necessary to refer children who are unable to complete testing with the WASS. In addition, the WASS can be used in conjunction with other tests such as a test of visual acuity.

A second limitation of the present study is that the sample constituted an enriched sample. Specifically, 44% of the participants had ametropia. Thus, the sample is not representative of a typical vision screening population in which approximately 9% have ametropia (Drover, Kean, Courage, and Adams, 2008). Considering this point, the screening effectiveness results provided here may not be generalizable. Nevertheless, the use of enriched samples is common in assessing the screening effectiveness of vision tests due to their convenience, and the fact that a smaller sample can be tested to determine whether the test detects disorders (Arici et al. 2012; Buchner et al., 2004; Cooper et al., 1996; Enzenauer et al., 2000; Funarunart et al., 2009; Granet et al., 1999; Iurno et al., 2004; Kennedy & Sheps, 1989; Kennedy & Thomas, 2000; Matta et al., 2011; Rajavi et al., 2012; Rogers et al., 2008; Silbert et al., 2014; Singman et al., 2013; Tong et al., 2000; Ugurbas et al., 2001; Watts et al., 1999; Weinand et al., 1998). Furthermore, the use of an enriched sample does not preclude a comparison of the two devices from the present study. Still, there is a possibility that this comparison would yield different results if this were a population-based study.

Future Research

Future research evaluating automated devices that measure refractive error should focus on three areas. First, instead of utilizing enriched samples, researchers should conduct population-based vision screening studies so that the findings can be generalized.

Second, future studies should include more modern autorefractors and photoscreeners such as the PS12, the Spot Photoscreener, the Pediatric Vision Screener, the Palm Autorefractor, the PA12, etc. These devices are relatively new and have not yet been evaluated fully. Indeed, they may be superior to both the WASS and the PS09. Finally, researchers should determine the cost-effectiveness of the devices from a third party payer perspective. Specifically, the cost of identifying each new patient with ametropia should be determined. This is critical as these devices are expensive compared to traditional tests (e.g., visual acuity), and thus, potential third-party payers will have to consider the purchase of an autorefractor/photoscreener very carefully.

Conclusions

The findings of the present study suggest that the devices are moderately reliable and that neither possesses high validity in terms of spherical refractive error. In addition, both devices demonstrated poor potential for early vision screening given the screening effectiveness results. Moreover, the PS09 yielded poor testability (i.e., the PS09). Given their disappointing performance and the fact that these devices are costly compared to traditional vision screening tests, one might question whether they should be used in a vision screening program. Yet, it must be reiterated that measurement of refractive error is considered by some to be the most effective way of detecting amblyopia and amblyogenic factors (Rotsos et al., 2009). These devices provide automatic estimates of refractive errors and as such, can be used by lay screeners in population-based vision screening programs. Also, many studies have reported far more positive results with these or similar tests suggesting that they do indeed possess screening potential. In light of the results of the present study, it would perhaps be wise to use these or similar devices in a

screening program with traditional vision screening tests (e.g., visual acuity, stereoacuity), while following lenient pass/fail criteria to ensure high specificity.

References

Adams, R.J., Dalton, S.M., Murphy, A.M., Hall, H.L., Courage M.L. (2002). *Ophthalmic and Physiological Optics*. Testing young infants with the Welch Allyn suresight non-cycloplegic autorefractor. 22(6):546-51. DOI: 10.5336/ophthal.2015-48416

Arnold R.W., Armitage M.D. 2014. Performance of four new photoscreeners on pediatric patients with high risk amblyopia. *Journal Pediatric Ophthalmology of Strabismus*. 51(1):46-52. doi: 10.3928/01913913-2

Arthur, B. W., Riyaz, R., Rodriguez, S., & Wong, J. (2009). Field testing of the PlusoptiX S04 photoscreener. *Journal of the Americal Association of Pediatric Ophthalmology and Strabismus, 13*, 51-57. doi: http://dx.doi.org/10.1016/j.jaapos.2008.08.016

Büchner, T. F., Schnorbus, U., Grenzebach, U. H., & Busse H. (2004). Examination of preschool children for ametropia: first experiences using a new hand-held autorefractor. *Strabismu, 12,* 111-117. http://dx.doi.org/10.1080/09273970490517854

Cavallini, A., Fazzi, E., Viviani, V., Astori, M. G., Zaviero, S., Bianchi, P. E., & Lanzi, G. (2002). Visual acuity in the first two years of life in healthy term newborns: An experience with the Teller Acuity Cards. *Functional Neurology*, *17*, 87-89

Chat, S. W., & Edwards, M. H. (2001). Clinical evaluation of the Shin-Nippon SRW-5000 autorefractor in children. *Ophthalmic Physiology of Optometry, 21*, 87-100. doi: 10.1016/S0275-5408(00)00049-1

Ciner, E. B., Ying, G., Kulp, M. T., Maguire, M. G., Quinn, G. E., Orel-Bixler, D., VIP Study Group. (2014). Stereoacuity of Preschool Children with and without Vision Disorders. *Optometry and Vision Science, 91,* 351–358. http://doi.org/10.1097/OPX.0000000000000165

Ciuffreda, K. J., Levi, D. M., & Selenow, A. (1991). Amblyopia, Basic and Clinical Aspects. *Butterworth-Heinemann: Stoneham.*

Cooper, C. D., Bowling, F. G., Hall, J. E., Colville, D. J. , Dortmans, R. J., & Munch J, Gole G. A. (1996). Evaluation of photoscreener instruments in a childhood population. 1. Otago photoscreener and Dortmans videophotorefractor. *Australian and New Zealand Journal of Ophthalmology, 24,* 347-355. doi: 10.1111/j.1442-9071.1996.tb01606.x

Cordonnier M., & Kallay O. (2001). Non-cycloplegic screening for refractive errors in children with the hand-held autorefractor Retinomax: final results and comparison with non- cycloplegic photoscreening. *Strabismus, 9,* 59-70

Cordonnier, M., & Dramaix, M. (1999) Screening for refractive errors in children: accuracy of the hand held refractor Retinomax to screen for astigmatism. *British Journal of Ophthalmology, 83,*157–161.

Cordonnier, M., & Dramaix M. (1998). Screening for abnormal levels of hyperopia in children: a non-cycloplegic method with a hand held refractor. *British Journal of Ophthalmology, 82,* 1260–1264. doi: 10.1136/bjo.82.11.1260

Demirci, G., Arslan, B., Özsütçü, M., Eliaçık, M.,& Gulkilik G. (2014). Comparison of photorefraction, autorefractometry and retinoscopy in children. *International Ophthalmology, 34,* 739-746. doi:10.1007/s10792-013-9864-x

Donahue, S. P., Arnold, R. W., Ruben, J. B., & AAPOS Vision Screening Committee. (2003*).* Preschool vision screening: what should we be detecting and how should we report it? Uniform guidelines for reporting results of preschool vision screening studies. *Journal of the American Association for Pediatric Ophthalmology and Strabismus, 7,* 314-316. doi: 10.1016/S1091853103001824

Drover, J. R., Felius, J., Cheng, C. S., Morale, S. E., Wyatt, L., & Birch, E. E. (2008) Normative pediatric visual acuity using single surrounded HOTV optotypes on the Electronic Visual Acuity Tester following the Amblyopia TreatmentStudy protocol. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 12*,145-149. doi: http://dx.doi.org/10.1016/j.jaapos.2007.08.014

Durr, N. J., Dave, S. R., Vera-Diaz, F. A., Lim, D., Dorronsoro, C., Marcos, S., Lage, E. (2015). Design and Clinical Evaluation of a Handheld Wavefront Autorefractor. *Optometry and Vision Science, 92,* 1140-1147. doi: 10.1097/OPX.0000000000000732

Economides, J. R., Adams, D. L., & Horton, J. C. (2012). Perception via the deviated eye in strabismus. *The Journal of Neuroscience, 32,* 10286-10295. doi: 10.1523/JNEUROSCI.1435-12.2012

Elliot, D. B., & Wilkes, R. D. (1989). A clinical Evaluation of the Topcon RM6000 Autorefractor. *Clinical and Experimental Optometery, 72,* 150–153. doi: 10.1111/j.1444-0938.1989.tb03076.x

Enzenauer, R. W., Freeman, H. L., Larson, M. R., & Williams, T. L. (2000). Photoscreening for amblyogenic factors by public health personnel: The Eyecor Camera System. *Ophthalmic Epidemiology, 7,* 1-12. http://dx.doi.org/10.1076/0928-6586(200003)711-2FT001

Eva, P. R., Pascoe, P. T., & Vaughan, D. G. (1982). Refractive change in hyperglycaemia: hyperopia, not myopia. *The British Journal of Ophthalmology, 66,* 500–505.

Ferhood, Q. K. (2012). Cycloplegic Refraction in Children with Cyclopentolate versus Atropine. *Clinical and Experimental Ophthalmology, 3.* doi:10.4172/2155-9570.1000239

Fogel-Levin, M., Doron, R., Wygnanski-Jaffe, T., Ancri, O., & Ben Zion, I. (2016). A comparison of plusoptiX A12 measurements with cycloplegic refraction. *Journal of Pediatric Ophthalmology Strabismus, 20,* 310-314. doi: 10.1016/j.jaapos.2016.04.006 doi:10.1016/j.jaapos.2016.04.006.

Freedman, H. L., & Preston, K. L. (1992). Polaroid photoscreening for amblyogenic factors. An improved methodology. *Ophthalmology, 99,* 1785-1795. https://doi.org/10.1016/S0161-6420(92)31722-1

Friedburg, D., & Klöppel, K. P. (1996). Early correction of hyperopia and astigmatism in children leads to better development of visual acuity. *Klinische Monatsblätter für Augenheilkunde, 209,* 21-24.

Funarunart P., Tengtrisorn S., Sangsupawanich P, & Siangyai P. (2009). Accuracy of Noncycloplegic Refraction in Primary School Children in Southern Thailand. *The Journal of Medical Association of Thailand, 92,* 806-812

Granet, D. B., Hoover, A., Smith, A. R., Brown, S. I., Bartsch, D. U., & Brody, B. (1999). A new objective digital computerized vision screening system. *Journal of Pediatric Ophthalmology and Strabismus, 36,* 251-256. doi: 10.3928/0191-3913-19990901-05

Guo, X., Jia, X., & Guo, L. (2000). Comparison of computer-photoscreening with noncycloplegic retinoscopy for amblyopiogenic risk factors in children. *Chinese Medical Journal, 113,* 1007-1010.

Hanneman, S. K. (2008). Design, analysis, and interpretation of method comparison studies. *AACN Advanced Critical Care, 19,* 223-224. doi: 10.4037/15597768-2008-2015

Harvey, E. M., Dobson, V., Miller, J. M., Clifford-Donaldson, C. E., Green, T. K., Messer, D. H., & Garvey KA. (2009). Accuracy of the Welch Allyn SureSight for measurement of magnitude of astigmatism in 3- to 7-year-old children. *Journal of American Association for Pediatric Ophthalmology and Strabismu13,* 466-471. doi: 10.1016/j.jaapos.2009.08.013

Hsiao, C., Park, J., & Orel-Bixler, D. (1999). Validity of Retinomax Autorefraction in Young Children. (Abstract). *American Academy of Optometry*.

Hu, Y. Y., Wu, J. F., Lu, T. L., Wu, H., Sun, W., Wang XR, … Jonas, J. B. (2015) Effect of cycloplegia on the Refractive Status of Children: The Shandong Children Eye Study. doi:10.1371/journal.pone.0117482

Huang, J., Maguire, M. G., Ciner, E., Kulp, M. T., Quinn, G. E., & Orel-Bixler, D. (2013). Inter-tester Agreement in Refractive Error Measurements. *Optometry and Vision Science, 90,* 1128–1137. doi: 10.1097/OPX.0000000000000025

Isenberg, S. J., Del Signore, M., & Madani-Becker, G. (2001). Use of the HARK autorefractor in children. *American Journal of Ophthalmology, 131,* 438-441. https://doi.org/10.1016/S0002-9394(00)00861-8.

Iuorno, J. D., Grant, W. D., & Noël, L. P. (2004). Clinical comparison of the Welch-Allyn SureSight handheld autorefractor versus cycloplegic autorefraction and retinoscopic refraction. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 8,* 123-127. doi: 10.1111/j.1463-5224.2012.01075.x

Joly, O., & Frankó, E. (2014). Neuroimaging of amblyopia and binocular vision: a review. *Frontiers in Integrative Neuroscience, 8.* http://doi.org/10.3389/fnint.2014.00062

Jost, R. M., Yanni, S. E., Beauchamp, C. L., Stager, D.R. Sr, Stager, D. Jr, Dao, L., & Birch, E. E. (2014). Beyond screening for risk factors: objective detection of strabismus and amblyopia. *JAMA Ophthalmology, 132,* 814-820. doi:10.1001/jamaophthalmol.2014.424

Kaur, G., Koshy, J., Thomas, S., Kapoor, H., Zachariah, J. G., & Bedi, S. (2016). Vision screening of school children by teachers as a community based strategy to address the challenges of childhood blindness. *Journal of Clinical Diagnostic Research, 10,* doi:10.7860/JCDR/2016/18939.7628.

Kemper, A.R., Keating, L. M., Jackson, J. L., & Levin EM. (2005). Comparison of Monocular autorefraction to comprehensive eye examinations in preschool-aged and younger children. *Archives of Pediatric Adolescence of Medicine, 159,* 435-439. doi: 10.1001/archpedi.159.5.435.

Kennedy, R., Sheps, S., & Bagaric D. (1995). Field trial of the Otago photoscreener. *Canadian Journal of Ophthalmology, 30,* 193–197.

Kennedy, R. A., & Sheps, S. B. (1989). A comparison of photoscreening techniques for amblyopic factors in children. *Canadian Journal of Ophthalmology, 24,* 259–264.

Kennedy, R. A., & Thomas, D. (2000). Evaluation of the iScreen digital screening system for amblyogenic factors. *Canadian Journal of Ophthalmology, 35,* 258-262. https://doi.org/10.1016/S0008-4182(00)80075-7

Kulp, M. T., Ying, G. S., Huang, J., Maguire, M., Quinn, G., Ciner EB, … Moore, B. D. (2014). Accuracy of noncycloplegic retinoscopy, retinomax autorefractor, and SureSight vision screener for detecting significant refractive errors. *Investigative Ophthalmology of Vision Science, 55,* 1378-1385. doi: 10.1167/iovs.13-13433

Kulp, M.T., Ying, G.S., Huang, J., Maguire, M.G., Quinn, G.E., Ciner, E., Vision in Preschoolers Study Group. (2011). Validity of Noncycloplegic Retinoscopy, Retinomax Autorefractor and SureSight Vision Screener for Detecting Significant Refractive Errors. (Abstract). *Investigative Ophthalmology and Visual Science, 52,* 1601

Lim, J. H., Bae, G. H., & Shin, S. J. (2014). The reliability and usefulness of refractive measurements by Plusiotix S09 in children. *Journal of Korean Ophthalmological Society, 55,* 1071-1076. https://doi.org/10.3341/jkos.2014.55.7.1071

Lowery, E. A., Wang, W., & Nyong'o O. (2014). Objective vision screening in 3-year-old children at a multispecialty practice. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 19,* 16-20. doi: 10.1016/j.jaapos.2014.09.008.

Lowery, J. P., Joachim, A., Olson, R., Peel, J., & Pearce, N. N. (2005). Autorefraction vs. Retinoscopy: A comparison of non-cycloplegic measures in a pediatric sample. *Journal of Behavioral Optometry*, *16*, 3-8.

Lim, J. H., Bae, G. H., & Shin, S. J. (2014). Reliability and Usefulness of Refractive Measurements by PlusoptiX S09 in Children. *The Korean Ophthalmological Society, 55,* 1071-1076. doi: 10.3341/jkos.2014.55.7.1071

Matta, N. S., Arnold, R. W., Singman E. L., & Silbert, D. I. (2011) Can a photoscreener help us remotely evaluate and manage amblyopia? *American Orthoptic Journal, 61,* 124-127. doi: 10.3368/aoj.61.1.124.

Matta, N. S., Singman, E. L., & Silbert, D. I. (2010). Performance of the PlusoptiX S04 photoscreener for the detection of amblyopia risk factors in children aged 3 to 5. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 4,* 147-149. doi: 10.1016/j.jaapos.2010.01.006

Maurer, D., & Lewis, T. L. (2001). Visual acuity: the role of visual input in inducing postnatal change. *Clinical Neuroscience Research, 1*, 239-247. doi: 10.1016/S1566-2772(01)00010-X

Maurer, D., Lewis, T. L. & Brent, H. P. (1989). The effects of deprivation on human development: studies of children treated for cataracts. In F. J. Morrision, C. Lord, & D. P. Keating (eds.). *Applied Developmental Psychology. 3*, 39-227. doi: 10.1097/OPX.0b013e3181a7296b.

McCurry, T. C., Lawrence, L. M., Wilson, M. E., & Mayo, L. (2013). The PlusoptiX S08 photoscreener as a vision screening tool for children with autism. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 17*, 374-377. doi: 10.1016/j.jaapos.2013.05.006

Morgan, K. S., & Johnson, W. D. (1987). Clinical evaluation of a commercial photorefractor. *Archives Ophthalmology, 105*, 1528-1531. doi:10.1001/archopht.1987.01060110074035

Moghaddam, A., Kargozar A., Zarel-Ghanavati, M., Najjaran, M., Nozari, V., & Shakeri, T. (2012). Screening for Amblyopia Risk Factors in Pre-verbal Children using the Plusoptix photoscreener. *The British Journal of Ophthalmology, 96*, 83-86. http://dx.doi.org/10.1136/bjo.2010.190405

Odom, J. V., Hoyt, C. & Marg, E. (1981). Effect of natural deprivation and unilateral eye patching on visual acuity of infants and children. *Archives of Ophthalmology, 99,* 1412-1416. doi: 10.1001/archopht.1981.03930020286018

Ogbuehi, K. C., Almaliki, W. H., AlQarni, A., & Osuagwu, U. L.(2015). Reliability and reproducibility of a handheld videorefractor. *Ophthalmology of Vision Science, 92*, 632-641. doi: 10.1097/OPX.0000000000000566.

Ottar, W. L., Scott, W. E., & Holgado, S. (1995). Photoscreening for amblyogenic factors. *Journal of Pediatric Strabismus, 32,* 289-295. doi: http://dx.doi.org/10.1016/S0161-6420(92)31722-1

Paff, T., Oudesluys-Murphy, A.M., Wolterbeek, R., Swart-van den Berg, M., de Nie, J. M,, Tijssen E., & Schalij-Delfos, N. E.(2010). Screening for refractive errors in children: the plusoptiX S08 and the Retinomax K-plus2 performed by a lay screener compared to cycloplegic retinoscopy. *Journal of Pediatric Ophthalmology Strabismus, 14,* 478-483. doi: 10.1016/j.jaapos.2010.09.015.

Pan, Y., Tarczy-Hornoch K., Cotter, S. A., Wen G., Borchert M. S., & Azen, S. P., Varma, R., the Multi-Ethnic Pediatric Eye Disease Study (MEPEDS) Group. (2009). Visual acuity norms in preschool children: the multi-ethnic pediatric eye disease study. *Optometry and Vision Science, 86,* 607–612. oi: 10.1097/OPX.0b013e3181a76e55

Payerols, A., Eliaou, C., Trezeguet, V., Villain, M., & Daien V. (2016). Accuracy of PlusOptix A09 distance refraction in pediatric myopia and hyperopia. *BMC Ophthalmology, 16,* 2-7. doi: 10.1186/s12886-016-0247-8.

Pediatric Eye Disease Investigator Group (2005). Early detection vs. late treatment of amblyopia. *Archives of Ophthalmology, 123,* 437-447

Peng, M. Y., Matta, N., Silbert, D., Tian, J., & Singman, E. (2014). Accuracy of the Retinomax K-plus3 in measuring refractive error in a pediatric population. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 18,* 327-31. doi:10.1016/j.jaapos.2014.02.017.

Peterseim, M. M., Papa, C. E., Wilson, M. E., Cheeseman, E. W., Wolf, B. J., Davidson, J. D., & Trivedi, R. H. (2014). Photoscreeners in the pediatric eye office: compared testability and refractions on high-risk children. *American Journal of Ophthalmology, 158,* 932-938. doi: 10.1016/j.ajo.2014.07.041

Rajavi, K., Parsafar, H., Ramezani, A., & Yaseri M., (2012). Is noncycloplegic photorefraction applicable for screening refractive amblyopia risk factors? *Journal of Ophthalmic & Vision Research, 7,* 3-10.

Rajavi, Z., Sabbaghi, H., Baghini, A. S., Yaseri, M., Sheibani, K., & Norouzi, G. (2015). Accuracy and repeatability of refractive error measurements by photorefractometry. *Journal of Ophthalmic and Vision Research, 10*, 221–228. http://doi.org/10.4103/2008-322X.170360

Rogers, D. L., Neely, D. E., Chapman, J. B., Plager, D. A., Sprunger, D. T., Sondhi, N., Ofner, S. (2014). Comparison of the MTI Photoscreener and the Welch-Allyn SureSight autorefractor in a tertiary care center. *Journal of Pediatric Ophthalmology and Strabismus, 12*, 77-82. doi: 10.1016/j.jaapos.2007.09.004

Rotsos, T., Grigoriou, D., Kokkolaki, A., & Manios, N. (2009). A comparison of manifest refractions, cycloplegic refractions and retinoscopy on the RMA-3000autorefractometer in children aged 3 to 15 years. *Clinical Ophthalmology, 3,* 429-431. doi: 10.2147/OPTH.S5145

Rowatt, A. J., Donahue, S. P., Crosby, C., Hudson, A. C., Simon, S., & Emmons K. (2007). Field evaluation of the Welch Allyn SureSight vision screener: incorporating the vision in preschoolers study recommendations. *Journal of Pediatric Ophthalmology and Strabismus, 11,* 243-248. doi: http://dx.doi.org/10.1016/j.jaapos.2006.09.008

Sabel, B. A. (2008). Plasticity and restoration of vision after visual system damage: An update. *Restorative Neurology And Neuroscience, 26,* 243-247.

Sanchez, I., Ortiz-Toquero, S., Martin, R., & de Juan, V. (2016). Advantages, limitations, and diagnostic accuracy of photoscreeners in early detection of amblyopia: a review. *Clinical Ophthalmology, 10,* 1365-1373. doi: 10.2147/OPTH.S93714

Schmidt, P., Maguire, M., Dobson, V., Quinn, G., Ciner. E., Cyert, L., …Vision in Preschoolers Study Group. (2004). Comparison of preschool vision screening tests as administered by licensed eye care professionals in the Vision In Preschoolers Study. *Ophthalmology, 111,* 637-650. doi: http://dx.doi.org/10.1016/j.ophtha.2004.01.022

Silbert, D. I., Matta, N. S., & Ely, A. L. (2014). Comparison of SureSight autorefractor and PlusoptiX A09 photoscreener for vision screening in rural Honduras. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 18,* 42-44 . doi: http://dx.doi.org/10.1016/j.jaapos.2013.09.006

Silbert, D., Matta, N., Tian, J., & Singman, E. (2014). Comparing the SureSight autorefractor and the plusoptiX photoscreener for pediatric vision screening. *Strabismus, 22,* 64-67. doi: 10.3109/09273972.2014.904896

Silverstein, E., Lorenz, S., Emmons, K., & Donahue S. P. (2009). Limits on improving the positive predictive value of the Welch Allyn SureSight for preschool vision screening. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 13,* 45-50. doi: 10.3109/09273972.2013.811606

Singman, E., Matta, N., Tian, J., & Silbert, D. (2013). A comparison of referral criteria used by the PlusoptiX photoscreener. *Strabismus, 21,* 190-194. doi: 10.3109/09273972.2013.811606

Smith, E. L., Ramamirtham, R., Qiao-Grider, Y., Hung, L.-F., Huang, J., Kee, C., & Paysse, E. (2007). Effects of foveal ablation on emmetropization and form-deprivation myopia. *Investigative Ophthalmology and Visual Science*, *48,* 3914–3922. doi:10.1167/iovs.06-1264

Tailor, V., Bossi, M., Greenwood, J. A., & Dahlmann-Noor, A. (2016). Childhood amblyopia: current management and new trends. *British Medical Bulletin, 119,* 75-86. doi: 10.1093/bmb/ldw030

Ugurbas, S. C., Alpay, A., Tutar, H, Sagdik, H. M., & Ugurbas, S. H. (2001) Validation of PlusoptiX S04 photoscreener as a vision screening tool in children with intellectual disability. *Journal of American Association for Pediatric Ophthalmology and Strabismus, 15,* 476-479. doi: 10.1016/j.jaapos.2011.05.023

Vision in Preschoolers (VIP) Study Group, Ciner, E., Carter, A., Ying, G. S., Maguire, M., & Kulp MT. (2011). Comparison of the Retinomax and Palm-AR auto-refractors: a pilot study. *Optometry and Vision Science, 88,* 830-836. doi: 10.1097/OPX.0b013e3182192658.

Watts P., Walker K., & Beck, L. (1999).Photoscreening for refractive errors in children and young adults with severe learning disabilities using the MTI photoscreener. *Eye (London) 13,* 363-368. doi: 10.1038/eye.1999.92

Weinand, F., Graf, M., & Demming, K. (1998). Sensitivity of the MTI photoscreener for amblyogenic factors in infancy and early childhood. *Graefe's Archive for Clinical and Experimental Ophthalmology, 236,* 801-805.

Wood, I. (2002). Reliability and validity of two handheld autorefractors. *American Academy of Optometry, 79,* 229.

Wood, I. C. (1987). A review of autorefractors. *Eye (London), 1,* 529-535.

Yilmaz, I., Ozkaya, A., Alkin, Z., Ozbengi, S., Yazici, A.T., & Demirok, A. (2015) Comparison of the Plusoptix A09 and Retinomax K-Plus 3 with retinoscopy in children. *Journal of Pediatric Ophthalmology and Strabismus, 52,* 37-42. doi: 10.3928/01913913-20141230-06

Ying, G. S., Kulp, M. T., Maguire, M., Ciner, E., Cyert, L., & Schmidt P. (2005).Vision in Preschoolers Study Group. Sensitivity of screening tests for detecting vision in preschoolers-targeted vision disorders when specificity is 94%. *Optometry and Vision Science, 82,* 432-438. doi: 10.1097/OPX.0b013e3181a59bf5