
Article:
Dawkins, A. and Bjerre, A. (2016) Do the near computerised and non-computerised crowded Kay picture tests produce the same measure of visual acuity? British and Irish Orthoptic Journal, 13. pp. 22-28. ISSN 1743-9868

10.22599/bioj.98

Br Ir Orthopt J 2016; **13:** 22–28

Do the near computerised and non-computerised crowded Kay picture tests produce the same measure of visual acuity?

AARON J. DAWKINS MMedSci BMedSci (Hons) AND ANNE BJERRE MSc BSc (Hons)

Department of Optometry and Orthoptics, Gloucestershire Hospitals NHS Foundation Trust, Gloucestershire

Abstract

Aims: **Apps have been developed to assess visual acuity (VA) on tablet computers. The aim of this study was to compare near VA scores using crowded Kay pictures on the iSight app and the printed crowded Kay picture test in amblyopic and typically developing children to determine whether the tests are clinically interchangeable.**

Methods: **Fifty-seven participants (34 typically developing and 23 amblyopic children) aged 3–9 years had their uniocular near VA measured using computerised crowded Kay pictures on the iSight app for the iPad and the printed near crowded Kay picture test. Data were analysed to determine whether there was a significant difference between the two tests. Bland–Altman plots were constructed to assess how well the tests agreed with each other.**

Results: **There was no significant difference between the two tests in all test conditions with the exception of the left eyes of typically developing children measured using the per line termination criteria ($p = 0.01$). Bland–Altman analysis showed good agreement between the iSight app and near crowded Kay picture test.**

Conclusions: **The results of the study indicate that near Kay pictures on the iSight app are clinically interchangeable with the traditional printed Kay pictures. These results compare well with other published studies comparing computerised vision tests with their traditional counterparts.**

Key words: iSight app, Kay picture test, Visual acuity

Introduction

Kay pictures, a picture-based near and distance visual acuity (VA) test, has been designed to measure VA in children aged 2–3 years old and can also be used to measure VA in illiterate adults.[1] The near printed Kay picture test is designed to examine VA at 33 cm and can assess VA from 0.875 to 0.000 logMAR. The test is comprised of a card with crowded and uncrowded pictures. The crowded format shows four pictures at each acuity level with no repeats surrounded by a crowding bar. In clinics, Kay pictures are primarily used to assess children.

VA apps designed for smart phones and tablet computers have been developed for non-eye-care professionals to test VA in children. It has been advocated the apps could play a beneficial role in vision screening, testing patients at the bedside or even the introduction of tele-health care within orthoptics.[2,3]

The iSight app for the iPad has been produced with Kay pictures. A near and distance crowded Kay picture test is available using the iSight app. The near test is based on the same principles as the printed near crowded Kay pictures. It measures VA from 0.800 to 0.000 logMAR, with four different pictures shown on each acuity line surrounded by a crowding bar. One acuity line is shown at a time on the iPad. The test begins with the 0.800 logMAR size. Progressively smaller logMAR pictures are shown in 0.100 logMAR intervals. If the examiner inputs three or more picture optotypes that are named incorrectly on an acuity line the test is terminated and VA is recorded as the logMAR size above. The examiner enters the responses into the iPad and a score is generated by the app. The design of the distance crowded Kay picture iSight app is different to the near app. Due to the size of an iPad there is a different number of optotypes shown on each acuity line on the distance tests. Northway *et al.*[4] compared the distance crowded Kay picture book with the iSight app at 3 m. They found the two tests agreed well in a group of preschool children and concluded that the distance iSight app could make a useful screening tool. As the near and distance Kay picture tests on the iSight app are different, it remains unknown whether the near test on the iSight app is also comparable to the near printed crowded Kay pictures. The authors have found no published studies providing normative data on picture-based near VA charts. Thus there is a need for near picture-based logMAR tests to be investigated.

The aim of this study was to compare the VA scores using the near crowded Kay picture iSight app designed for the iPad and the printed near crowded Kay picture test in amblyopic children and typically developing children to determine whether the VA scores are clinically interchangeable. The study also aimed to investigate any differences in the sensitivity between the iSight app on the iPad and the near printed crowded Kay picture tests in revealing the inter-ocular VA difference.

Correspondence and offprint requests to: Aaron Dawkins. e-mail: a_dawkins@hotmail.co.uk

Methods

The study prospectively collected data using a cross-sectional convenient sample in a repeated measures design. Ethics approval was gained from the London City and East Research Ethics Committee prior to data collection. All parents and participants were given participant and parent information sheets. Consent forms were signed by parents of all participants prior to inclusion in the study. Children aged 7 years and over completed an assent form.

All participants were examined by the same examiner to ensure standardised testing. To minimise order effect, counterbalancing using a Latin square was used to decide the order of vision test and which eye was tested first. The order of the pictures is different on the reverse side of the near Kay picture card. The reverse side was used for the second eye tested to prevent participants memorising the test. The iSight app automatically randomises the order pictures are shown each time the test is restarted. To enable VA to be assessed to threshold, and thus avoid ceiling and floor effects, the testing distance was altered. If the participant was unable to identify the largest size pictures (0.800 logMAR) the test was moved closer. If the participant was able to see all the smallest size pictures (0.000 logMAR) the testing distance was increased. The altered testing distances were calculated by multiplying or dividing the test distance by the logarithmic progression ratio 1.2589.[5,6] Each time the testing distance was altered the printed Kay picture card was reversed and the iSight app was restarted to randomise the picture order.

Termination criteria

The individual optotype termination criteria (referred to as the per picture termination criteria) and the iSight app termination criteria (referred to as the per line termination criteria) were used for both the iSight app and the printed Kay picture tests. For the per picture termination criteria each participant was asked to name every picture optotype from the largest size until an entire line of optotypes were read incorrectly. The response for each optotype shown was included in the VA score. The iSight app does not use a per picture termination criteria scoring method. If three or more optotypes on one acuity line are not read correctly the test is terminated and VA is scored at the logMAR acuity level above. As the examiner enters the responses the app calculates the VA score using the above-described per line termination criteria. The researcher recorded the per line termination criteria first and manually continued the test until the per picture termination criteria were met.

Participants

Two participant groups (amblyopic paediatric patients and typically developing children) were chosen to determine whether measures of VA threshold and inter-ocular VA differences obtained using the near printed crowded Kay picture test versus the iSight app were different. A convenient sample of 34 typically developing participants were recruited from one primary school and 23 amblyopic participants were recruited from patients attending outpatient orthoptic clinics at Cheltenham General and Gloucestershire Royal Hospitals. Patients aged 3–9 years attending the hospital eye clinic diagnosed with strabismic, anisometropic, meridional, or combined amblyopia were included. Amblyopia was defined as a best corrected VA of 0.200 logMAR or worse in the amblyopic eye and an inter-ocular VA difference of no less than 0.100 logMAR as described by Elliot and Firth.[7] Participants were required to wear up-to-date spectacles issued within the last 12 months if prescribed. Typically developing children aged between 3 and 9 years were included in the study but those aged between 4 and 5 years who did not pass their school vision screening test were excluded. Children with known intraocular pathology, stimulus deprivation amblyopia, attention deficit hyperactivity disorder (ADHD) and ptosis were excluded. Non-English-speaking participants were excluded from the study as there was no funding for the use of an interpreter.

Procedures

Both VA tests were held 33 cm from the participant's eye in a slightly depressed position. A tape measure was used to ensure VA was tested at the correct distance. Parents of participants held the tape measure in place for amblyopic participants recruited from hospital clinics. The testing distance for typically developing participants was re-measured each time a participant moved their head as no third person was available to hold a tape measure in place. VA was measured uniocularly using the printed near crowded Kay picture test, and the computerised near crowded Kay pictures using the iSight app on a third-generation iPad with retina display.

Statistical analysis

Data recorded using the per picture termination criteria and per line termination criteria were recorded and analysed separately. Data for the two participant groups were also analysed separately to enable an analysis of any differences in inter-ocular VA between the two tests, as it is possible for there to be a difference in the inter-ocular difference in VA in the absence of a significant difference in monocular VA. The data were analysed to assess for normal distribution using the Shapiro–Wilks test for normality. As the data were normally distributed parametric tests were performed for further analysis. A two-tailed paired *t*-test assessed for significant difference in VA scores between printed and computerised crowded Kay pictures at near using the iSight app, and also between the per picture and per line termination criteria. The Bonferroni method was used to adjust for multiple comparisons. The traditional significance level of 0.05 has been divided by the number of comparisons made. The *p* values are therefore significant if equal to or less than 0.01. Descriptive statistics using methods outlined by Bland and Altman[8] were used to assess how well near computerised crowded Kay pictures using the iSight app for the iPad agreed with the printed near crowded Kay picture card. IBM SPSS Statistics version 21 was used to analyse the data.

BIOJ ::: ::: MS No. 7 {gb} {F2}
Produced on Wednesday 24th August 2016
File: O:/BRITORTH/VOL 73-2016/MAKE-UP/7 - DAWKINS.3D

Table 1. Mean visual acuity, given in logMAR, with standard deviation for the typically developing (group 1) and amblyopic participants (group 2) employing the per line termination criteria and the per picture termination criteria

Termination criteria	Test	Group 1		Group 2	
		Right eye	Left eye	Amblyopic eye	Non-amblyopic eye
Per line	iSight app	0.018 ± 0.145	0.021 ± 0.127	0.287 ± 0.184	0.035 ± 0.192
	Printed Kay pictures	−0.003 ± 0.145	−0.021 ± 0.127	0.304 ± 0.155	0.043 ± 0.170
Per picture	iSight app	0.005 ± 0.151	−0.015 ± 0.134	0.284 ± 0.183	0.016 ± 0.184
	Printed Kay pictures	−0.022 ± 0.143	−0.048 ± 0.194	0.285 ± 0.172	0.012 ± 0.165

Group 1, typically developing children; group 2, amblyopic children.

Table 2. The mean difference ± standard deviation and significance values (two-tailed paired *t*-test) between the iSight app and the near printed crowded Kay picture test for typically developing participants (group 1) and amblyopic participants (group 2) using the per line and per picture termination criteria

Termination criteria	Value	Group 1		Group 2	
		Right eye	Left eye	Amblyopic eye	Non-amblyopic eye
Per line	Mean difference ± SD	0.015 ± 0.097	0.042 ± 0.090	−0.017 ± 0.089	−0.009 ± 0.124
	p values	0.38	0.01	0.36	0.74
Per picture	Mean difference ± SD	0.027 ± 0.109	0.027 ± 0.097	−0.030 ± 0.030	0.004 ± 0.102
	p values	0.16	0.13	0.03	0.84

Group 1, typically developing children; group 2, amblyopic children.

Table 3. The mean difference and significance values (two-tailed paired *t*-test) between the per line and per picture termination criteria in typically developing participants (group 1) and amblyopic participants (group 2)

Test	Value	Group 1		Group 2	
		Right eye	Left eye	Amblyopic eye	Non-amblyopic eye
iSight app	Mean difference ± SD	0.013 ± 0.047	0.023 ± 0.054	0.033 ± 0.054	0.018 ± 0.064
	p values	0.13	0.02	0.00	0.18
Near printed crowded Kay picture test	Mean difference ± SD	0.025 ± 0.055	0.027 ± 0.128	0.019 ± 0.060	0.032 ± 0.061
	p values	0.01	0.23	0.13	0.02

Group 1, typically developing children; group 2, amblyopic children.

Results

A total of 57 participants aged between 3 years 2 months and 9 years 8 months (mean 6 years 3 months ± 1 year 7 months) participated in the study. Thirty-four participants were recruited to the typically developing group (mean age 6 years 5 months ± 1 year 8 months, range 3 years 2 months to 9 years 8 months) and 23 participants (mean age 6 years 1 month ± 1 year 4 months, range 3 years 4 months to 9 years 2 months) to the amblyopic group. In the amblyopic group, 13 had strabismic, 4 anisometropic and 6 a combination of strabismic and anisometropic amblyopia.

Kolmonogorov–Smirnov ($p > 0.05$) and Shapiro–Wilks tests ($p > 0.05$) revealed the data to be normally distributed.

The mean VA ± standard deviation (SD) measured using the per line termination criteria and per picture termination criteria data for each test condition are shown in Table 1.

The differences in VA scores obtained between the two tests are displayed in Table 2.

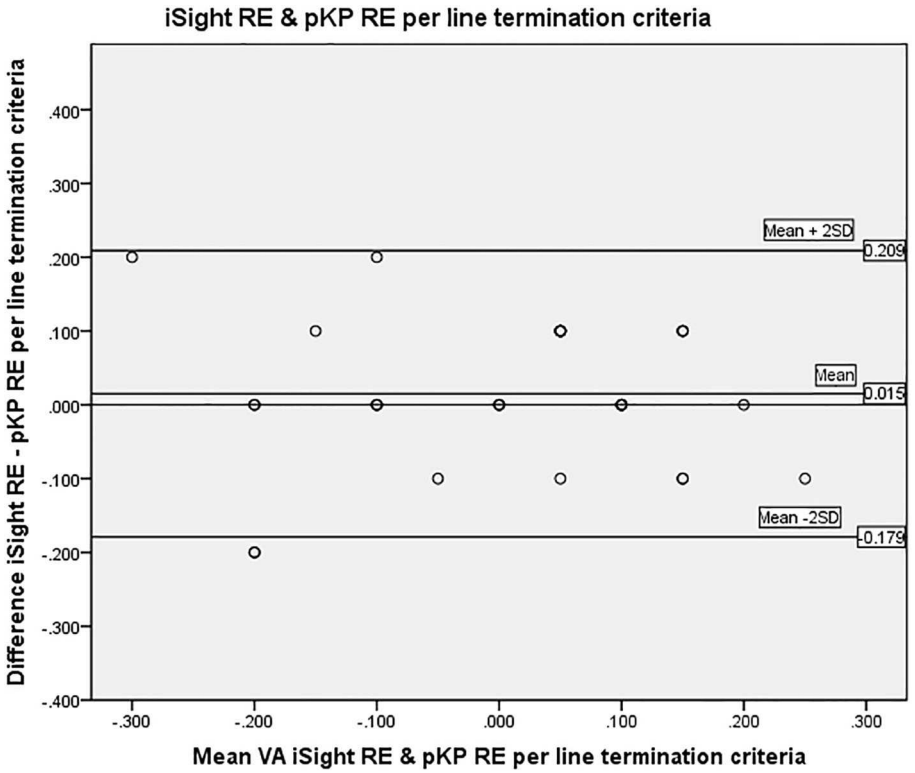
The impact of the two different termination criteria on the VA scores obtained for each test was analysed using a two-tailed paired *t*-test. Table 3 shows the mean difference and significance values between the per line and per picture termination criteria.

The mean differences between the inter-ocular VA difference for the iSight app and near printed crowded Kay pictures using the per line and per picture termination criteria in typically developing participants were 0.003 ± 0.004 and 0.005 ± 0.002 logMAR respectively. In amblyopic participants the mean differences were 0.009 ± 0.156 logMAR and −0.058 ± 0.214 logMAR for the per line and per picture termination criteria respectively. There was no statistically significant difference between the inter-ocular differences using the two tests in typically developing participants when assessed with the per line ($t = 1.999$, $p = 0.85$) and the per picture termination criteria ($t = 1.999$, $p = 0.65$), or in amblyopic participants assessed with the per line ($t = 0.267$, $p = 0.79$) and the per picture termination criteria ($t = -1.314$, $p = 0.20$).

Bland–Altman analysis was performed to assess the level of agreement between the iSight app and near printed crowded Kay pictures. Fig. 1 shows the best and least agreement in typically developing participants and Fig. 2 the equivalent data for amblyopic participants. The mean differences in VA, SD of the difference, 95% upper and lower limits of agreement and coefficient of agreement (1.96× SD of the difference) for the two groups of participants are shown in Table 4. The coefficient of agreement represents the amount of variation in VA that can be expected between the two tests.

BIOJ ::: ::: MS No. 7 {gb} {F2}
Produced on Wednesday 24th August 2016
File: O:/BRITORTH/VOL 73-2016/MAKE-UP/7 - DAWKINS.3D

(*a*)



(*b*)

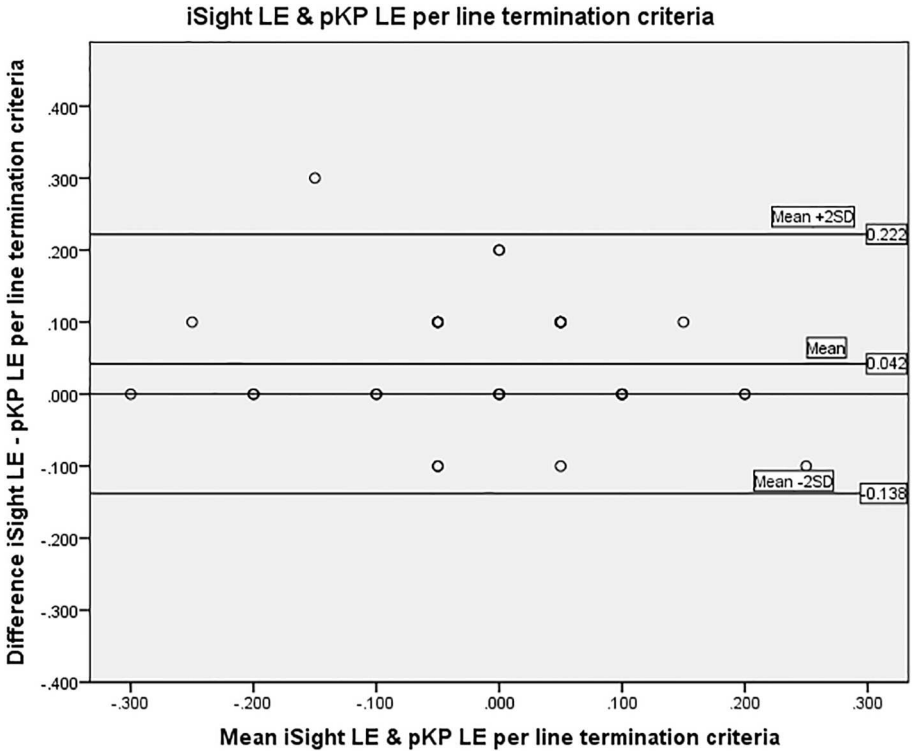
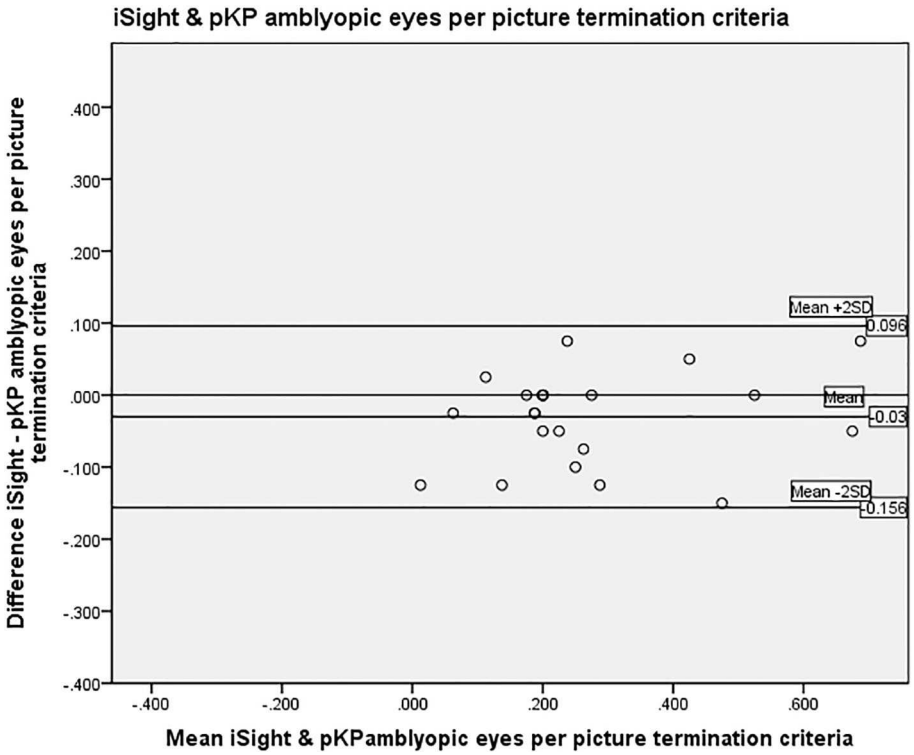


Fig. 1. Bland–Altman plot showing the best and least agreement respectively between the iSight app and near printed crowded Kay picture test (pKP) in typically developing participants (group 1): (*a*) in the right eyes using the per line termination criteria; (*b*) in the left eyes using the per line termination criteria.

BIOJ ::: ::: MS No. 7 {gb} {F2}
Produced on Wednesday 24th August 2016
File: O:/BRITORTH/VOL 73-2016/MAKE-UP/7 - DAWKINS.3D

(*a*)



(*b*)

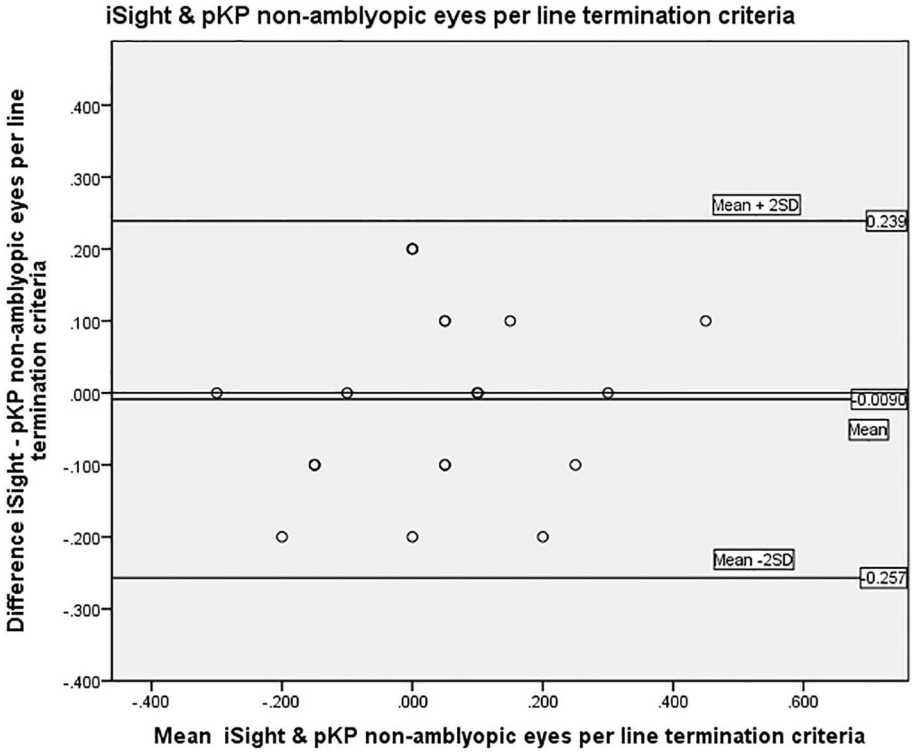


Fig. 2. Bland–Altman plot showing the best and least agreement respectively between the iSight app and near printed crowded Kay picture test (pKP) in amblyopic participants (group 2): (*a*) in amblyopic eyes using the per picture termination criteria; (*b*) in non-amblyopic eyes using the per line termination criteria.

BIOJ ::: ::: MS No. 7 {gb} {F2}
Produced on Wednesday 24th August 2016
File: O:/BRITORTH/VOL 73-2016/MAKE-UP/7 - DAWKINS.3D

Table 4. Mean difference in VA between the iSight app and near printed crowded Kay picture test, with standard deviation (SD), 95% upper and lower limits and the coefficient of agreement for typically developing participants (group 1) and amblyopic participants (group 2)

Termination criteria	Group 1				Group 2			
	Right eye		Left eye		Amblyopic eye		Non-amblyopic eye	
	Per line	Per picture	Per line	Per picture	Per line	Per picture	Per line	Per picture
Mean difference	0.015	0.027	0.042	0.027	−0.017	−0.030	−0.009	0.004
SD of the difference	0.097	0.109	0.090	0.097	0.089	0.063	0.124	0.102
95% upper limits	0.209	0.245	0.222	0.221	0.161	0.096	0.239	0.208
95% lower limits	−0.179	−0.191	−0.138	−0.167	−0.195	−0.156	−0.010	−0.204
Coefficient of agreement (1.96× SD)	0.190	0.215	0.176	0.190	0.174	0.123	0.243	0.200

Group 1, typically developing children; group 2, amblyopic children.

Discussion

The computerised near crowded Kay pictures using the iSight app showed similar results to the near printed crowded Kay picture test for both amblyopic and typically developing children. Using a Bonferonni correction, a statistically significant difference in the mean VA scores between the two tests was found only in the left eyes of the typically developing children tested using the per line termination criteria. The difference was <2 optotypes and would not be considered clinically significant.

There was no statistically significant difference between the inter-ocular VA differences for each of the two tests in either the typically developing or the amblyopic participants. This indicates that the two tests are comparable in their ability to detect amblyopia, as one test does not produce more of a crowding effect in either group of children.

The use of computerised VA tests in eye clinics is becoming increasingly popular. A number of studies have validated and compared computerised tests with the traditional printed versions and found comparable results.[9–11] Shah *et al.*[11] reported a mean difference of −0.01 logMAR in a paediatric amblyopic group when the printed crowded Kay picture test was compared with computerised Kay pictures on the COMPlog system – a similar finding to the −0.017 ± 0.089 logMAR mean difference found in the amblyopic eyes tested with the per line termination criteria in this study.

Bland–Altman analysis comparing the agreement between the two tests showed the coefficient of agreement to range between −0.176 and −0.215 logMAR for typically developing children and between 0.123 and −0.243 logMAR for amblyopic children. The coefficient of agreement shows there to be a range of up to two logMAR lines variation between the two tests. The iSight app and the near printed crowded Kay picture test can be considered as clinically interchangeable, as there was no significant difference between the mean VA scores obtained with the two tests and the differences in the mean VA scores shown on the Bland–Altman plots were mostly within 2 SD of the mean, indicating that the two tests agree well. These results compare well with other studies in which computerised tests have been compared with printed VA tests.[9–11] A difference of >1 line of optotypes is considered clinically significant by some studies;[7,12] however, other studies report that changes of up to two logMAR lines in either direction can occur as a result of test–retest variability.[9,11,13] This study only assessed agreement between tests. Further research is needed to assess the expected measurement error when retesting the same patients with the same technique.

A statistically significant difference was found between the per line and per picture termination criteria in two of the test conditions. In the test conditions where a significant difference was found, the mean difference was approximately one optotype. The slightly worse VA scores were measured when the per line termination criteria were used, but the difference would not be considered clinically significant. The per line termination criteria used by the iSight app would be suitable for use by non-eye-care professionals, with any recommendation advising of the potential difference that may be found.

The sample size was limited to 23 amblyopic participants. Using the mean VA of 0.158 logMAR (SD 0.176) found in the amblyopic group, and an alpha value of 0.01 looking for an effect size of 0.075 logMAR, the study had a power of 38%. For the study to have a power of 80%, 30 more participants would have been needed in this group.

Other factors affecting the results include the behaviour and concentration of children. Behaviour is a variable which cannot be controlled and may have affected results to a greater extent with younger participants. In a larger study an analysis of the results in different age groups may prove useful. The same examiner performed both tests to ensure standardisation of the testing procedure; however, this may have resulted in examiner bias. The study has assumed that light levels were similar for all participants. As light levels were not measured it is possible different light levels affected the results.

There is a clinical impression that near VA scores may improve faster than distance VA scores in amblyopic eyes undergoing amblyopia treatment. The authors have found no published research to substantiate this claim. Christoff *et al.*[14] found no statistically significant difference between the near and distance VA scores in a group of amblyopic children. These results were from data recorded using single surround HOTV optotypes. Further research is needed to investigate the relationship between the near and distance VA in normal and amblyopic participants using crowded vision tests, as this will inform the relevance for near VA testing.

BIOJ ::: ::: MS No. 7 {gb} {F2}
Produced on Wednesday 24th August 2016
File: O:/BRITORTH/VOL 73-2016/MAKE-UP/7 - DAWKINS.3D

Conclusion

This study found that results of the near crowded Kay pictures using the iSight app for the iPad agreed well with those of the printed near crowded Kay pictures in amblyopic and typically developing children. No significant difference was found between the two tests or between the inter-ocular VA differences with each test, which suggests the two tests compare well in terms of VA measures and their ability to detect amblyopia. To further validate the iSight app a larger sample size and an assessment of its repeatability is required.

The authors declare they have no competing interests.

References

1. Kay H. New method of assessing visual acuity with pictures. *Br J Ophthalmol* 1983; **67:** 131–133.
2. Armfield NR, Gray LC, Smith LC. Clinical use of Skype: a review of the evidence base. *J Telemed Telecare* 2012; **18:** 125–127.
3. Fonda SJ, Bursell SE, Lewis DG, Garren J, Hock K, Cavallerano J. The relationship of a diabetes Telehealth eye care program to standard eye care and change in diabetes health outcomes. *Telemed J E Health* 2007; **13:** 635–644.
4. Northway N, Panesar G, McCulloch D, McKay R. A validation of iSight app in pre-school vision screening Presented at the XIV Biennial Meeting of the Child Vision Research Society, 17–19 June 2013, Ontario, Canada. Available at: http://cvrsoc.org/docs/CVRS2013-web.pdf pg 34 [accessed 8 September 2013].
5. Bailey IL, Lovie-Kitchen JE. Visual acuity testing. From the laboratory to the clinic. *Vis Res* 2013; **90:** 2–9.
6. Oduntan AO. A practical logMAR near reference table for low vision practitioners: design and applications. *S Afr Optom* 2006; **65:** 157–162.
7. Elliot MC, Firth AY. The logMAR Kay picture test and the logMAR acuity test: a comparative study. *Eye* 2009; **23:** 85–88.
8. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1:** 307–310.
9. Laidlaw DAH, Tailor V, Shah N, Harcourt C. Validation of a computerised logMAR visual acuity measurement system (COMPlog): comparison with ETDRS and the electronic ETDRS testing algorithm in adults and amblyopic children. *Br J Ophthalmol* 2008; **92:** 241–244.
10. Ehrmann K, Fedtke C, Radić A. Assessment of computer generated vision charts. *Contact Lens Ant Eye* 2009; **32:** 133–140.
11. Shah N, Laidlaw DAH, Rashid S, Hysi P. Validation of printed and computerised crowded Kay picture logMAR tests against gold standard ETDRS acuity test chart measurements in adult and amblyopic paediatric subjects. *Eye* 2012; **26:** 593–600.
12. Jones D, Westall C, Averbeck K, Abdolell M. Visual acuity assessment: a comparison of two tests for measuring children's vision. *Ophthalmic Physiol Opt* 2003; **23:** 541–546.
13. Shah N, Laidlaw DAH, Brown G, Robson C. Effect of letter separation on computerised visual acuity measurements: comparison with the gold standard Early Treatment Diabetic Retinopathy Study (ETDRS) chart. *Ophthalmic Physiol Opt* 2010; **30:** 200–203.
14. Christoff A, Repka MX, Kaminski BM, Holmes JM. Distance versus near visual acuity in amblyopia. *J AAPOS* 2011; **15:** 342–344.

BIOJ ::: ::: MS No. 7 {gb} {F2}
Produced on Wednesday 24th August 2016
File: O:/BRITORTH/VOL 73-2016/MAKE-UP/7 - DAWKINS.3D