Automated Measurement of Visual Acuity in Pediatric Ophthalmic Patients Using Principles of Game Design and Tablet Computers

TARIQ M. ASLAM, HUMZA J. TAHIR, NEIL R.A. PARRY, IAN J. MURRAY, KUN KWAK, RICHARD HEYES, MAHANI M. SALLEH, GABRIELA CZANNER, AND JANE ASHWORTH

• PURPOSE: To report on the utility of a computer tablet–based method for automated testing of visual acuity in children based on the principles of game design. We describe the testing procedure and present repeatability as well as agreement of the score with accepted visual acuity measures.
• DESIGN: Reliability and validity study.
• METHODS: SETTING: Manchester Royal Eye Hospital Pediatric Ophthalmology Outpatients Department. PATIENT POPULATION: Total of 112 sequentially recruited patients. INTERVENTION: For each patient 1 eye was tested with the Mobile Assessment of Vision by intERactIve Computer for Children (MAVERIC-C) system, consisting of a software application running on a computer tablet, housed in a bespoke viewing chamber. The application elicited touch screen responses using a game design to encourage compliance and automatically acquire visual acuity scores of participating patients. Acuity was then assessed by an examiner with a standard chart-based near ETDRS acuity test before the MAVERIC-C assessment was repeated. MAIN OUTCOME MEASURE: Reliability of MAVERIC-C near visual acuity score and agreement of MAVERIC-C score with near ETDRS chart for visual acuity.
• RESULTS: Altogether, 106 children (95%) completed the MAVERIC-C system without assistance. The vision scores demonstrated satisfactory reliability, with test-retest VA scores having a mean difference of 0.001 (SD ±0.136) and limits of agreement of 2 SD (LOA) of ±0.267. Comparison with the near EDTRS chart showed agreement with a mean difference of −0.0879 (±0.106) with LOA of ±0.208.
• CONCLUSIONS: This study demonstrates promising utility for software using a game design to enable automated testing of acuity in children with ophthalmic disease in an objective and accurate manner. (Am J Ophthalmol 2016;170:223–227. © 2016 Elsevier Inc. All rights reserved.)

AMBLYOPIA IS THE MOST COMMON CAUSE OF VISUAL impairment in children (prevalence in childhood of 1%–4%) and the leading cause of monocular vision loss in the 20- to 70-year- old age group.[1] There are many other significant causes of central visual loss in children, including cataract, corneal opacity, and retinal disease.[2] Crucial to the management of these conditions is an accurate and reliable assessment of vision, appropriate to the precise clinical needs. This may range from testing of a broad number of detailed visual functions by specialists in hospital patients to more basic vision screening of schoolchildren in vans.[3] Recent attention has been directed to computerized vision testing, and many applications are readily available for desktop, laptop, mobile, and tablet devices, with several publications addressing this potential.[4–7] However, most of these computerized systems are for testing adults and are not validated for automatic testing of vision in children without expert assistance.

This paper presents assessment of a system of automated vision testing in children using a customized computer tablet–based acuity test. Its key features are that it is housed in its own controlled viewing environment and uses game design principles to automatically present appropriate graphical targets, eliciting responses from children without the need for external intervention. It has been developed through an extensive, iterative period of testing and redesign and is based on an adult test (Mobile Assessment of Vision by intERactIve Computer; MAVERIC) on which we have previously reported.[7]

METHODS

WE CONDUCTED A RELIABILITY AND VALIDITY STUDY TO assess a novel system of automated vision testing in

AJO.com Supplemental Material available at AJO.com.

Accepted for publication Aug 8, 2016.

From the Faculty of Medical and Human Sciences (T.M.A., K.K., R.H., M.M.S., J.A.), Faculty of Life Sciences (H.J.T., I.J.M., J.A.), and Centre for Hearing and Vision Research, Institute of Human Development (N.R.A.P.), University of Manchester, Manchester, United Kingdom; Vision Science Centre, Manchester Royal Eye Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, United Kingdom (T.M.A., N.R.A.P., I.J.M., J.A.); Heriot-Watt University, Edinburgh, United Kingdom (T.M.A.); Faculty of Life Sciences, University of Bradford, Bradford, United Kingdom (N.R.A.P.); and Department of Eye and Vision Science and Biostatistics, University of Liverpool, Liverpool, United Kingdom (G.C.).

Inquiries to Tariq M. Aslam, Central Manchester Foundation Trust, Oxford Rd, Manchester, M13 9WL UK; e-mail: Tariq.aslam@cmft.nhs.uk

0002-9394/$36.00
http://dx.doi.org/10.1016/j.ajo.2016.08.013

children (Mobile Assessment of Vision by intERactIve Computer for Children; MAVERIC-C) in terms of its reliability and agreement with standard methods of vision assessment, as well as its acceptability to pediatric hospital patients. The research adhered to the tenets of the Declaration of Helsinki and ethics committee approval for the full testing protocol was obtained prior to the start of the study through the U.K. independent research approval system. Informed consent was obtained for all patients, who were recruited from a general pediatric ophthalmology clinic in a teaching hospital in England.

The structure and development of the fundamental aspects of the MAVERIC system have been previously described in detail in its application to adults.[7,8] In essence, it consists of a computer tablet loaded with specifically designed software housed in a custom-made viewing booth. The software generates a diminishing-sized square array resembling a Landolt "C" in the center of the screen and requires the user to detect the location of the gap. The subject responds by pressing 1 of the 4 surrounding buttons that corresponds to the gap location. The button provides audible and visible feedback and the software is programmed to give verbal encouragement to the subjects if they fail to respond to a target within a time limit. The game design element of the software is implemented by custom-made graphics and specific animations when the correct or incorrect response is given. For the system used in testing children (MAVERIC-C) the animation graphics and game design were enhanced to make the process more appealing to children. The onscreen graphics present either a mouse going into a hole, a pig into a pen, or a sheep into a pen when the correct responses are given, with multiple additional animations and sounds to encourage a child's participation. The MAVERIC vision testing strategy involves 4 phases. Phase 1 involves a screening test to derive an initial rapid and approximate threshold. Phase 2 involves detailed threshold detection; 3 out of 4 correct responses are required to progress to the next reduced target size. If this is failed, the same target size is repeated; and if failed twice in a row, the phase is ended and the threshold taken as the last sequence of 3 correct responses out of 4. By the end of this challenging stage children's concentration might be waning, so an additional simple suprathreshold test with new graphics is incorporated as phase 3 before a final repeated detailed threshold-level test. This principle of using multiple tests concurs with other established vision-testing algorithms.[9,10]

The tablet used in the current study was the Galaxy Tab Pro 8.4 SMT 320 (Samsung Electronics, South Korea), selected for its high screen resolution. The display resolution of 1600 × 2560 pixels over an 8.4-inch-diagonal screen size allowed for the smallest letter size to be −0.21 logMAR (Snellen equivalent 6/3.7 or 20/12.3) at the testing distance of 40 cm. The next step was 0.09 logMAR (6/7.4 or 20/24.7), as this was the next possible size according to the screen resolution. For this study the maximum testing size used was 1.22 logMAR (6/100 or 20/333.3) and the device was calibrated with a photometer such that the central target luminance was black (0.57 cd/m^2) while the surrounding luminance was set to the maximum of 397.6 cd/m^2. Overall contrast was therefore 99%.

One hundred and twelve children were enrolled for the study from a typical pediatric ophthalmology outpatient clinic. Patients were excluded only if they had a physical disability that excluded use of a tablet computer. To incorporate a significant range of visual acuities into the study, the eye chosen for the trial was the eye with worst visual acuity, up to 1.22 logMAR (6/100). If the vision was the same in both eyes, the right eye was chosen. All children wore habitual correction for all vision testing, with additional correction for 40 cm near testing if they were pseudophakic.

Before a computerized visual assessment was conducted, children were shown the tablet computer outside its booth and given a few minutes to familiarize themselves with the MAVERIC-C game. When the patient showed that he or she understood and could perform the basic test, the tablet was placed inside the booth and the patient invited to look through the viewing aperture, placing his or her hands inside the booth to provide a comfortable location from which to operate the tablet. For all children the fellow eye was occluded using a patch. The patient began the visual testing by pressing a large central green "start" button on the tablet's screen. Once the test started, no further external input was given while the test ran through levels automatically, giving programmed audio encouragement where required. When the final acuity was determined a cheer sounded to signify the end of the examination. Masked to the MAVERIC-C vision result, the examiner then tested the near visual acuity of the patient using an ETDRS logMAR chart (Precision Vision, Lasalle, United States, SKU 2112). The examiner positioned the near visual acuity chart at the 40 cm test distance from the bridge of the child's nose using the attached cord. The acuity score was recorded as a logarithm of the minimum angle of resolution (logMAR) value for the last line in which the subject identified 3 or more optotypes on that line, plus a value of −0.02 log unit for each optotype that was identified correctly beyond that line.

Finally, approximately 20 minutes after the first test, a second MAVERIC-C test was initiated.

The test-retest reliability of the MAVERIC-C system was assessed using the Bland-Altman limits of agreement (LOA) method[11] to assess agreement between 2 measures. We used the same procedure to assess the agreement of the MAVERIC-C score with the standard near acuity charts.

RESULTS

A TOTAL OF 112 CHILDREN (52 MALE) WERE ENTERED into the study, aged between 4 and 16 years (mean:

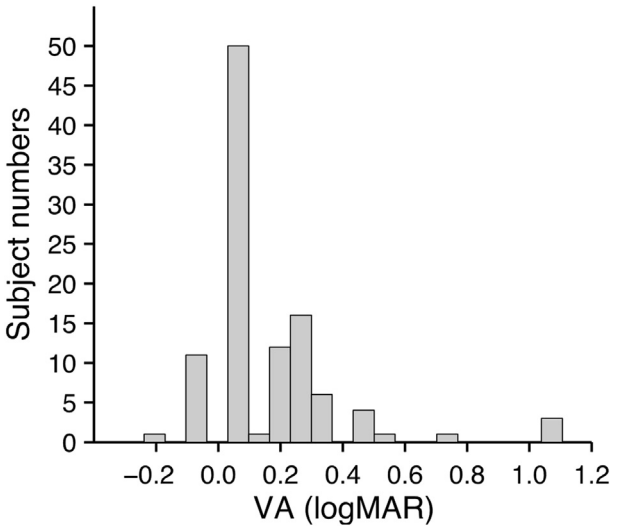

FIGURE 1. Distribution of near visual acuity (VA) scores across subjects as measured by the MAVERIC-C system. Measurable acuity ranges from −0.21 logMAR (Snellen equivalent 6/3.7 or 20/12.3) to 1.22 logMAR (6/100 or 20/333.3) for this study.

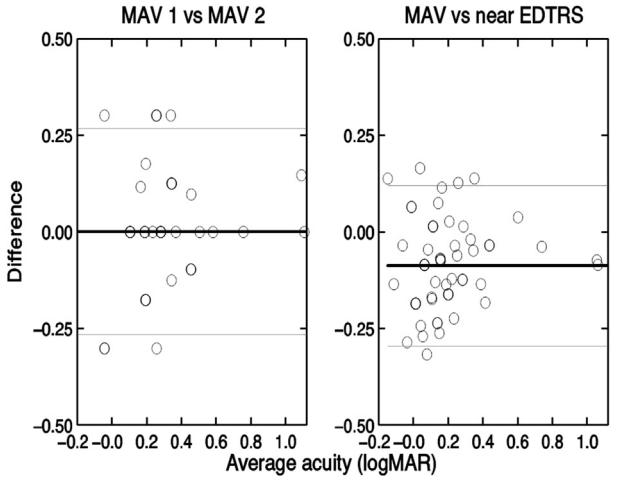

FIGURE 2. Bland-Altman plot of repeatability of MAVERIC-C test measures (Left) and MAVERIC-C vs near ETDRS test (Right). Thick black line shows mean difference (second test − first test), thin black lines show ±1.96 SD, and dashed lines show 95% confidence interval for the upper and lower limits of agreement.

10.2 ± 2.82 years). Of the 112 subjects recruited, 106 were able to complete the test without any further assistance (95%). Testing algorithms dictated that the time to completion of the whole test from pressing start to the finishing cheers was between 3 and 6 minutes and average time was approximately 5 minutes. Those who could not complete the test had difficulty with understanding or willingness to play and were excluded from further analysis.

The range of pathologies of the children included 45 with primary diagnoses of anterior segment disorders (27 with keratoconjunctivitis, 14 with uveitis, 4 with cataract). There were 27 children with strabismus and amblyopia, including 11 with convergent squint (4 alternating eso) and 6 with divergent squint (1 intermittent). There were 14 children entered into the study with oculoplastic disorders and 16 with neuroophthalmology disorders, including 4 with nystagmus. There were 10 children with other miscellaneous diagnoses or for whom there was no abnormality found. The distribution of near visual acuity (displayed in Figure 1) was fairly broad.

The Bland-Altman plot for the repeatability measurement of the MAVERIC-C system is presented in Figure 2 (Left), with differences randomly scattered around the mean. The differences were approximately normally distributed, with a mean of 0.001 and a standard deviation of ±0.136. LOA of 2 SD were ±0.267 (95% confidence interval [CI] for the upper LOA was +0.268 to +0.267 and lower LOA was −0.265 to −0.266).

Figure 2 (Right) shows the Bland-Altman plot for the average MAVERIC-C acuity scores and the near EDTRS scores. The differences were approximately normally distributed, with a mean of −0.0879 and a standard deviation of 0.106. Limits of agreement of 2 SD were ±0.208 (95% CI for the upper LOA was +0.120 to +0.120 and lower LOA was −0.295 to −0.296).

DISCUSSION

THE MAVERIC-C TEST DEMONSTRATED A HIGH DEGREE OF acceptability and capability of automatically testing children's vision. The children were recruited from a pediatric ophthalmology clinic and none were ultimately excluded owing to physical inabilities. Of the 112 children recruited, all but 6 were able to complete the test (95%). The children were recruited from a hospital setting and therefore had previous practice with visual acuity testing, but had not been previously exposed to this tablet-based MAVERIC-C in any form. This acceptability of the MAVERIC-C system is based on algorithms and user interfaces that drive users through the robust threshold assessments. Ruamviboonsuk and associates, Moke and associates, and Beck and associates, in particular, recognized the importance of well-designed algorithms to achieve good vision measures,[5,9,10] and our system built upon the concepts used in those adult studies, adding features such as game designs, animations, voice feedback, and individualized timed responses.

The visual acuity measurement was of satisfactory repeatability, considering difficulties of automated testing of vision in this group and comparison with other published studies; in this study, the mean difference in the repeated scores was 0.001 and LOA was 2 SD

± 0.267. Results for the foundational MAVERIC test in adults were mean 0.003 and LOA 2 SD ± 0.17 for high-contrast testing and mean −0.03 and LOA 2 SD ± 0.31 for low-contrast testing. This disparity may be partly explained by particular challenges around vision testing in children. A recent paper using game design principles[12] used an interactive video game to evaluate vision and demonstrated reliability indices superior to those of our system (reliability 95% LOA ± 0.18 logMAR). However, that study required the investigators to have direct input into testing, with resulting potential for bias. There was no time interval between tests, meaning a reduced likelihood of fatigue but also potentially greater possibility of memorizing cues to eventual outcome. Their system involved distance rather than near testing, and children were screened to include those who had good dexterity with use of a computer mouse. In addition, in our study, the worse-seeing eyes were recruited in order to fully test the system by including significant numbers of eyes with poor vision. However, this would also inevitably have led to greater numbers of amblyopic eyes that might demonstrate greater variability. Finally, other factors, such as learning to play the game and fatigue, could of course have significant impact in this new technique tested on children who may have already had a significant wait for their outpatient appointment. These differences in protocol may contribute to apparent differences in repeatability. Overall, the repeatability of the MAVERIC-C compares favorably to other acuity tests such as the peekaboo test,[13] where the LOA was ±0.33, and to reports of repeatability of gold-standard pediatric acuity tests (mean 0.01, LOA ±0.35).[13]

In addition to repeatability, we determined agreement between the MAVERIC-C acuity and standard chart-based measures. The differences were approximately normally distributed, with a mean of −0.0879, and LOA of 2 SD were ±0.208. The comparator tests were chosen because they were the most similar available tests that had accepted validity. However, they represent different psychophysical tasks than the MAVERIC-C test and we would not expect exact agreement. In some respects the computerized tests may be superior (greater objectivity in recording responses, use of timing, more standardized instruction) and in some ways inferior (limited range of acuities/contrast levels/ability to tailor encouragement and nature of test to particular child). The slightly higher mean scores we found for the ETDRS concur with previous clinical studies that demonstrate that acuity determined with Landolt C chart is significantly lower than that determined by ETDRS chart, possibly owing to complex letter shapes facilitating the recognition task.[14]

In previous studies we found pixel size limitations to be a significant restriction in testing higher acuities, where the smallest 2 gaps equated to VAs of −0.08 and 0.22. The use of a more modern tablet, the Galaxy Tab Pro 8.4, enabled us to improve this initial step size from −0.22 logMAR to 0.09 logMAR, allowing more precise measurement at the higher visual acuity levels. We anticipate that this issue will become decreasingly relevant in the future as screen technologies continue to advance. Considering these fundamental differences, the agreement between the MAVERIC-C test and ETDRS chart testing was satisfactory and compares favorably with other computerized tests, such as the peekaboo (mean 0.07, LOA ±0.33)[13] and specific computerized pediatric tests[12] (mean 0.05 LOA ± 0.27).

Although most hospital measures are based on distance acuity, we chose to develop a near rather than distance acuity test. In terms of basic geometrical optics, visual acuity should be the same for distance and near.[15] However, we acknowledge that these functions are not clinically interchangeable and near visual acuity results might in practice be different from distance acuity. Distance visual acuity is different from near in various types of strabismus and nystagmus, as well as in pseudophakic subjects without appropriate correction. In addition, some reports have found that visual acuity at near differs compared with visual acuity at distance for amblyopic eyes[16] and that accommodation is reduced in amblyopia.[17] In contrast, a recent study assessed children with amblyopia and concluded that individual differences between distance and near visual acuity are likely owing to test-retest variability.[15]

With the caveat that near and distance acuity are not necessarily interchangeable, there are distinct advantages in practicality and objectivity that the near test affords. A distance VA test would not allow for a direct touch screen response and would have led to greater dependence on an examiner or a remote device being used. It would also necessitate that the test be set to the correct testing distance, at least 3 m away. Control over illumination, to minimize glare sources and reflections, would be more difficult and the test would most likely have to be conducted in a dark room. These practical implications would render the device more difficult to set up correctly and use at home as a self-testing device, and these considerations led us to develop a near VA test in a self-contained, portable unit. This feature should have positive implications for the potential uses of the MAVERIC-C system away from the controlled environment of a clinical setting and in other public environments or patients' homes. Self-testing of vision at patients' homes might allow for safer monitoring of chronic conditions such as uveitis or cataract. It would be relatively simple to modify the system to alert the hospital or parents if unexpected vision loss occurred, for example in a patient with orbital disease. Future studies will assess the utility of the device when used in such ways.

In summary, this study demonstrates a novel, self-contained computerized unit for automated assessment of visual acuity in children. It is highly acceptable to children and demonstrates repeatability and a high level of agreement with gold-standard tests. Its design features allow it

to acquire values for visual acuity without dependence on external instructors or on external environment control, enhancing its potential use outside of a controlled hospital setting. Future studies will assess its use in different settings, in separate age groups, and in more precise pathologies.

FUNDING/SUPPORT: THE STUDY WAS SUPPORTED BY A GRANT FROM CENTRAL MANCHESTER FOUNDATION TRUST (Manchester, UK, R03034). Financial disclosures: Tariq M. Aslam received fees from Bayer, Novartis, Bausch and Lomb, and Thea. Jane Ashworth received a travel grant and expenses from Biomarin Ltd. The following authors have no financial disclosures: Humza J. Tahir, Neil R.A. Parry, Ian J. Murray, Kun Kwak, Richard Heyes, Mahani M. Saleh, and Gabriela Czanner. All authors attest that they meet the current ICMJE criteria for authorship.

REFERENCES

1. Attebo K, Mitchell P, Cumming R, Smith W, Jolly N, Sparkes R. Prevalence and causes of amblyopia in an adult population. *Ophthalmology* 1998;105(1):154–159.
2. De Paula CH, Vasconcelos GC, Nehemy MB, Granet D. Causes of visual impairment in children seen at a university-based hospital low vision service in Brazil. *J AAPOS* 2015;19(3):252–256.
3. Griffith JF, Wilson R, Cimino HC, Patthoff M, Martin DF, Traboulsi EI. The use of a mobile van for school vision screening: results of 63 841 evaluations. *Am J Ophthalmol* 2016;163:108–114e1.
4. Ruamviboonsuk P, Sudsakorn N, Somkijrungroj T, Engkagul C, Tiensuwan M. Reliability of visual acuity measurements taken with a notebook and a tablet computer in participants who were illiterate to Roman characters. *J Med Assoc Thai* 2012;95(Suppl 3): S109–S116.
5. Ruamviboonsuk P, Tiensuwan M, Kunawut C, Masayaanon P. Repeatability of an automated Landolt C test, compared with the early treatment of diabetic retinopathy study (ETDRS) chart testing. *Am J Ophthalmol* 2003; 136(4):662–669.
6. Black JM, Jacobs RJ, Phillips G, et al. An assessment of the iPad as a testing platform for distance visual acuity in adults. *BMJ Open* 2013;3(6):ii.
7. Aslam TM, Parry NR, Murray IJ, et al. Development and testing of an automated computer tablet-based method for self-testing of high and low contrast near visual acuity in ophthalmic patients. *Graefes Arch Clin Exp Ophthalmol* 2016;254(5):891–899.
8. Tahir HJ, Murray IJ, Parry NR, Aslam TM. Optimisation and assessment of three modern touch screen tablet computers for clinical vision testing. *PLoS One* 2014;9(4):e95074.
9. Moke PS, Turpin AH, Beck RW, et al. Computerized method of visual acuity testing: adaptation of the amblyopia treatment study visual acuity testing protocol. *Am J Ophthalmol* 2001;132(6):903–909.
10. Beck RW, Moke PS, Turpin AH, et al. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol* 2003;135(2):194–205.
11. Bland JM, Altman DG. Measuring agreement in method comparison studies. *Stat Methods Med Res* 1999;8(2):135–160.
12. Ma DJ, Yang HK, Hwang JM. Reliability and validity of an automated computerized visual acuity and stereoacuity test in children using an interactive video game. *Am J Ophthalmol* 2013;156(1):195–201e1.
13. Livingstone IA, Lok AS, Tarbert C. New mobile technologies and visual acuity. *Conf Proc IEEE Eng Med Biol Soc* 2014;2014:2189–2192.
14. Kuo HK, Kuo MT, Tiong IS, Wu PC, Chen YJ, Chen CH. Visual acuity as measured with Landolt C chart and Early Treatment of Diabetic Retinopathy Study (ETDRS) chart. *Graefes Arch Clin Exp Ophthalmol* 2011;249(4):601–605.
15. Christoff A, Repka MX, Kaminski BM, Holmes JM. Pediatric Eye Disease Investigator Group. Distance versus near visual acuity in amblyopia. *J AAPOS* 2011;15(4):342–344.
16. Catford GV. Amblyopia: a comparison between distance and near vision. *Br J Ophthalmol* 1956;40(10):633–635.
17. Hokoda SC, Ciuffreda KJ. Measurement of accommodative amplitude in amblyopia. *Ophthalmic Physiol Opt* 1982;2(3): 205–212.