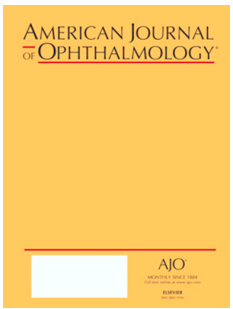
Accepted Manuscript



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. Saleh, Gabriela Czanner, Jane Ashworth

PII: S0002-9394(16)30397-X

DOI: 10.1016/j.ajo.2016.08.013

Reference: AJOPHT 9854

To appear in: *American Journal of Ophthalmology*

Received Date: 24 May 2016

Revised Date: 8 August 2016

Accepted Date: 8 August 2016

Please cite this article as: Aslam TM, Tahir HJ, Parry NRA, Murray IJ, Kwak K, Heyes R, Saleh MM, Czanner G, Ashworth J, Automated Measurement of Visual Acuity in Pediatric Ophthalmic Patients Using Principles of Game Design and Tablet Computers, *American Journal of Ophthalmology* (2016), doi: 10.1016/j.ajo.2016.08.013.

Abstract

Purpose: To report on the utility of a computer tablet based method for automated testing of visual acuity in children based upon 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: 112 sequentially recruited patients.
- Intervention: For each patient one 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 tests 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: 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 2SD (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.

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

Short Title : Automated Vision Testing in Pediatric Patients.

Tariq M Aslam[1,2,5], Humza J Tahir[3], Neil R A Parry [2,4,6], Ian J Murray[2,3], Kun Kwak[1], Richard Heyes[1] , Mahani M Saleh[1], Gabriela Czanner[7], Jane Ashworth [1,2,3].

1. Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK
2. Vision Science Centre, Manchester Royal Eye Hospital, Central Manchester University Hospitals NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester , UK
3. Faculty of Life Sciences, University of Manchester, Manchester, UK
4. Centre for Hearing and Vision Research, Institute of Human Development, University of Manchester, Manchester, U.K
5. Heriot-Watt University, Edinburgh, U.K
6. Faculty of Life Sciences, University of Bradford
7. Department of Eye and Vision Science and Biostatistics, University of Liverpool, UK

Corresponding Author; Tariq Aslam, Central Manchester Foundation Trust, Oxford Road, Manchester, M13 9WL.
Tel 0161 2761234.
Email Tariq.aslam@cmft.nhs.uk

Introduction

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 20-70 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 school children 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 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 upon an adult test (MAVERIC) we have previously reported.[7]

Methods

We conducted a reliability and validity study to assess a novel system of automated vision testing in children (MAVERIC –C) in terms of its reliability and agreement with standard methods of vision assessment as well as 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 (IRAS). Informed consent was obtained for all patients, who were recruited from a general paediatric ophthalmology clinic in a teaching hospital in England.

The structure and development of the fundamental aspects of the Mobile Assessment of Vision by intERactIve Computer (MAVERIC) system has 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 one of the four 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 subject 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 four phases. Phase one involves a screening test to derive an initial rapid and approximate threshold. Phase two involves detailed threshold detection; three out of four 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 three correct responses out of four. By the end of this challenging stage children's concentration might be waning and so an additional simple supra-threshold 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), selected for its high screen resolution. The display resolution of 1600 x 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 40cm. 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 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. In order 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 40cm near testing if they were pseudophakic. Before conducting a computerized visual assessment, 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 they 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 their hands inside the booth to provide a comfortable location from where 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 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, 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 log 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 asses agreement between two measures. We used the same procedure to assess the agreement of the MAVERIC-C score with the standard near acuity charts.

Results

112 children (52 males) were entered into the study, aged between 4 and 16 years old (mean: 10.2 ± 2.82). 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 forty-five with primary diagnoses of anterior segment disorders (twenty-seven with keratoconjunctivitis, fourteen with uveitis, four with cataract). There were twenty-seven children with strabismus and amblyopia, including eleven with convergent squint (four alternating eso) and six with divergent squint (one intermittent). There were fourteen children entered into the study with oculoplastic disorders and sixteen with neuro-ophthalmology disorders, including four with nystagmus. There were ten children with other miscellaneous diagnoses or for whom there was no abnormality found. The distribution of near visual acuity was fairly broad, displayed in Figure 1.

The Bland-Altman plot for the repeatability measurement of the MAVERIC-C system is presented in Figure 2 (left plot), 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. Limits of agreement (LOA) of 2SD were +/-0.267 (95% CI for the upper LOA was +0.268 to +0.267 and lower LOA was -0.265 to -0.266).
Figure 2 (right plot) 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 2SD 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 due 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 upon algorithms and user interfaces that drive users through the robust threshold assessments. Ruamvibasoon, Beck and Moke 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 limits of agreement (LOA) of 2SD +/-0.267. Results for the foundational MAVERIC test in adults were mean 0.003 and limits of agreement 2SD +/- 0.17 for high contrast testing and mean -0.03 and limits of agreement of +/-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% limits of agreement ±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 worst 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 were ±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 standard measures. The differences were approximately normally distributed with a mean of -0.0879, and limits of agreement of 2SD were +/-0.208. The comparator tests were chosen as they were the most similar available tests that had accepted validity. However, they represent different psychophysical tasks to 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 acuity determined with Landolt C chart is significantly lower than that determined by ETDRS chart, possibly due 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 two gaps equated to VAs of -0.08 and 0.22. The use of a more modern tablet, the Galaxy tab 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 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] (Dae Jong, 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 to near in various types of strabismus, nystagmus as well as in pseudophakes without appropriate correction. In addition, some reports have found that visual acuity at near differs compared to 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 due 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 to be used. It would also necessitate the test to be set to the correct testing distance, at least 3m away. Control over illumination, to minimise 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 setup 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 into other public environments or patient's homes. Self-testing of vision at patient's 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 dependence 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 more precise pathologies.

Acknowledgements / Disclosures

a. Funding/Support: The study was supported by a grant from Central Manchester Foundation Trust. R03034

b. Financial Disclosures:
Tariq Aslam ; received fees from Bayer, Novartis, Bausch and Lomb, Thea
Neil Parry ; No financial disclosures
Ian Murray ; No financial disclosures
Humza Tahir ; No financial disclosures
Kun Kwak; No financial disclosures
Mahani Saleh; No financial disclosures
Richard Heyes; No financial disclosures
Gabriela Czanner; No financial disclosures
Jane Ashworth : Travel grant and expenses from Biomarin Ltd

References

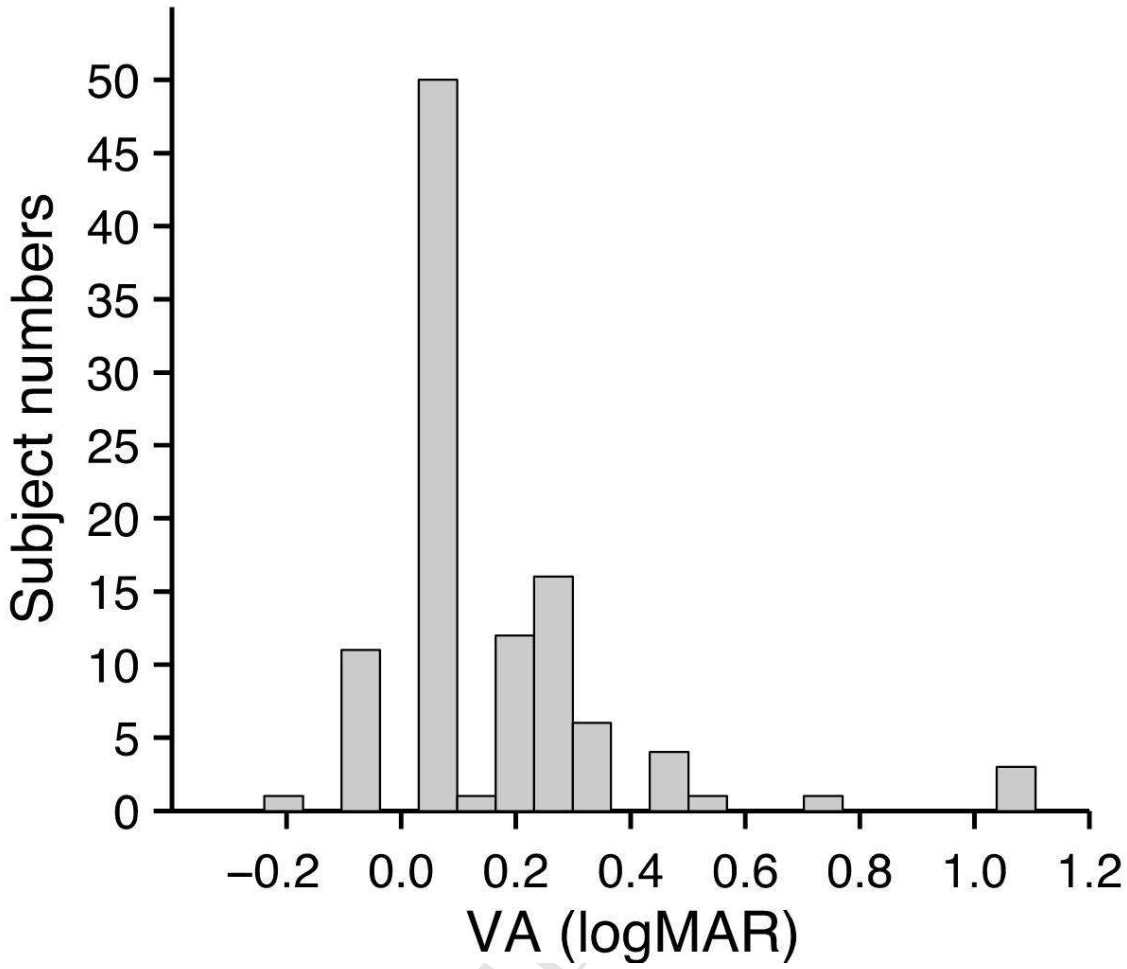
1. Attebo K, Mitchell P, Cumming R, Smith W, Jolly N, Sparkes R. Prevalence and causes of amblyopia in an adult population. *Ophthalmology* 1998;105(1):154-9.
2. De Paula CH, Vasconcelos GC, Nehemy MB, Granet D. Causes of visual impairment in children seen at a university-based hospital low vision service in Brazil. *J AAPOS* 2015;19(3):252-6.
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;163:108-114.e1.
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:S109-16.
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):662-9.
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).pii
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):891-9
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):903-9.
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):135-60.
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):195-201.e1.

13. Livingstone IA, Lok AS, Tarbert C. New mobile technologies and visual acuity. *Conf Proc IEEE Eng Med Biol Soc* 2014;2014:2189-92.
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):601-5.
15. Christoff A, Repka MX, Kaminski BM, Holmes JM, Pediatric Eye Disease Investigator G. Distance versus near visual acuity in amblyopia. J AAPOS 2011;15(4):342-4.
16. Catford GV. Amblyopia: a comparison between distance and near vision. *Br J Ophthalmol* 1956;40(10):633-5.
17. Hokoda SC, Ciuffreda KJ. Measurement of accommodative amplitude in amblyopia. *Ophthalmic Physiol Opt* 1982;2(3):205-12.

Figure Captions

Figure 1. Distribution of near visual acuity scores across subjects as measured by MAVERIC-C system. Measureable 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 plot) and MAVERIC-C vs near ETDRS test (right plot). Thick black line shows mean difference (second test- first test), thin black lines show +/-1.96 SD and dashed red lines show 95%CI for the upper and lower LOA.


Subject numbers
50
45
40
35
30
25
20
15
10
5
0
–0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
VA (logMAR)