Testing a Popular Smartphone Application for Colour Vision Assessment in Healthy Volunteer Subjects

Michael J Flitsos^a, Jiawei Zhao^a, Tejus Pradeep^a, Mehrnaz Ighani^b, and Allen O Eghrari^a

^aThe Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA; ^bCollege Park, University of Maryland, College Park, Maryland, USA

ABSTRACT

We aimed to compare the Ishihara pseudoisochromatic colour vision test with a colour vision test from a popular smartphone application (EyeHandBook [EHB]) using digital image processing to simulate colour vision deficiencies. Three digital versions of the Ishihara and EHB slides were created: full colour; 32 bit- greyscale (removing all colour information); and blue channel (to simulate red-green colour vision deficiencies). Twenty healthy volunteers were shown each colour-edited plate. The answers they reported were compared with what would be expected for that colour-simulation scenario based on the answer key provided in the Ishihara booklet (“expected” answer). There were nine plates that had comparable patterns between the EHB and Ishihara test. We found no significant difference in the overall proportion of “expected” answers for the full colour ($p = .35$), 32 bit-greyscale ($p = .39$) and blue channel ($p = .22$) conditions. There were significant differences between the proportion of “expected” answers among six individual colour-edited plates ($p < .05$ for each). Colour vision assessment from the EHB is distinct from comparable Ishihara plates. Clinical scenarios that require serial assessment of colour vision may benefit from using the same modality consistently rather than exchanging between the two tests with the assumption of equivalence. Refinement of digital colour editing techniques beyond 32-bit greyscale and RGB channel splitting is necessary in order to accurately simulate colour vision deficiency.

ARTICLE HISTORY

Received 10 July 2020
Accepted 9 August 2020

KEYWORDS

Colour vision testing;
Ishihara; mHealth;
smartphone applications;
mobile applications

Introduction

The mobile-health (mHealth) industry has reshaped the practice and delivery of medicine, with most physicians using their smartphones to fulfil professional obligations¹ and nearly 500 million smartphone users worldwide have mHealth applications.² The “Eye Handbook” (EHB) mobile application is an illustrative example of the reach of mHealth in eye care. There have been 2.3 million downloads of this application, with 25,000 active global users.³ The application provides basic screening tests for users including Snellen charts for visual acuity, Pelli-Robson charts for contrast sensitivity, and Amsler grids. The colour vision test of the EHB is of similar design to conventional Ishihara pseudoisochromatic colour vision plates. Recent studies have highlighted potential inconsistencies in colour vision testing between physical and digital Ishihara replications.⁴ These differences may be in part due to the colour gamut differences between various smartphone screens.⁵ Differences between colour vision

testing features of different mobile applications on the same device have also been described.⁶

While there are clear design similarities between the EHB and Ishihara plates, it is unknown whether the plates were designed by the application creators themselves, or if they are digital replications of the Ishihara plates. Additionally, there is no interpretation guide in the app, presenting a challenge when attempting to correlate results with the physical Ishihara test. While general agreement between the EHB and Ishihara has been reported in a previous investigation,⁷ a pilot study we conducted demonstrated differences in reported answer choices to the EHB versus Ishihara when subject to changes in colour information.⁸ In the present study, we aimed to investigate EHB’s colour vision testing functionality in healthy human volunteers under scenarios of normal and simulated colour vision deficit and compare with digital replications of Ishihara plates under the same simulations. We then characterised how participants’

answer choices differed between the different slide sets.

Materials and methods

All study procedures were approved by the Johns Hopkins University School of Medicine Institutional Review Board. Participants who were free of colour vision deficit by self-report were recruited and informed consent was obtained. Inclusion criteria included normal best-corrected Snellen visual acuity, normal contrast sensitivity assessed with a Pelli-Robson chart, and normal colour vision via Ishihara pseudoisochromatic colour vision testing. None of the research participants were excluded from the study based on these criteria.

Images from 16 plates in the colour vision test of the EHB application were obtained by screenshots from an iPhone 6S Plus (Apple Inc., Cupertino, CA, USA). Photographs were taken of slides from a booklet of 17 Ishihara pseudoisochromatic plates (1997 edition; Kanehara & Co., Ltd., Tokyo, Japan) at a distance of 6 inches with a Canon T5i digital single-lens reflex camera (Canon Inc., Ōta, Tokyo, Japan). The Ishihara pseudoisochromatic plates we photographed were obtained specifically for the purposes of this study and were stored closed in a drawer in a room with no external light exposure. Next, ImageJ software (version 1.46 r; Wayne Rasband, National Institutes of Health, Bethesda, MD, USA) was used to process images to 32 bit-greyscale (removing all colour information). Next, split red-blue-green (RGB) channels were used to separate the blue, red, and green colour information from the full-colour images using a methodology that has been described previously.⁹ The 32 bit-greyscale images represent lack of all colour vision, while the blue channel images simulate removal of red and green colour information. Examples of colour-manipulated images for the Ishihara plates are demonstrated in Figure 1 and for EHB in Figure 2 and Supplemental Figure 1.

The full colour, 32 bit-greyscale, and blue channel images as described above were presented to participants in random order on a 9.7-inch iPad Pro (2016 edition, Apple Inc., Cupertino, CA, USA) at full brightness. The viewing distance at which these images were presented was proportional relative to the viewing distance of the Ishihara

Figure 1. Visual manipulation of plate 1 of the Ishihara plates. The number 12 is visible in the full colour (A) slide. It is not visible in the 32 bit-greyscale (B) slide. It is visible in the blue channel (C), green channel (D), and red channel (E).

pseudoisochromatic plates when compared with the size of the iPad screen. The Ishihara images were viewed at a distance of 30 inches from participants' line of sight, while the EHB images were shown at a distance of 35 inches. Participants were asked to state whether they saw a specific numerical answer, nothing (coded as "n"), or a squiggly line they could trace (coded as "s").

There is no colour vision interpretation guide in the version of EHB that we tested. Thus, we carried the assumption many clinicians may hold that the pattern (i.e. number) represented on the EHB slide corresponds to the same slide containing that number in the Ishihara plates. When a number was reported as an answer choice by our study participants, it was compared with what was expected for that specific plate based on the answers expected for the Ishihara plates based on the answer key in the version of Ishihara we tested (Supplemental Figure 2). Expected answers for the 32-bit greyscale plates were correlated with the Ishihara answer key column "Person with total colour

Figure 2. Visual manipulation of plate 1 of the EHB plates. The number 12 is visible in the full colour (A) plate. It is faintly visible in the 32 bit-greyscale (B) plate, as well as blue channel (C), green channel (D), and red channel (E).

blindness and weakness”, while the expected answers for the blue channel condition were correlated to the “Person with red-green (R-G) deficiencies” column.

Statistical analysis for the study was performed using Stata 15 (StataCorp, College Station, TX, USA). A paired-samples t-test was used to determine whether there was a statistically significant mean difference between correct answers reported by subjects using matched Ishihara and EHB plates for each colour condition, both when taken as a whole and per each individual colour-manipulated slide. The association between the subjects’ ability to report expected answer choices based on the type of test (Ishihara versus EHB) stratified by colour condition was examined using Fisher’s exact tests.

Results

The characteristics of our 20 study participants are outlined in Table 1. Our cohort had a mean age of 23.8 years, was mostly female (60%), and was diverse from the perspective of race/ethnicity. In plates 2 through 15 of the Ishihara plates, the majority of participants correctly identified the expected pattern in the full-colour plates (mean 100% accuracy) and all 32 bit-greyscale plates (mean 85.88% accuracy). In contrast, the mean accuracy of the blue channel was 65.58%.

For EHB, plates 8, 9, 10, 12, 13, 14, and 16 were excluded from comparison with the percentage expected responses, as there was no correlating Ishihara plate design to which these plates could be compared. In the selected EHB slides, a majority of participants identified the expected number of full-colour slides (mean expected answer rate of 99.44%), most 2 bit-greyscale slides (mean 67.22%), and most blue channel slides (mean 68.89%).

Next, the rate of expected answer choices based on the Ishihara answer key was compared between Ishihara and EHB for the plates that had the same central number pattern (Table 2). There were no statistically significant differences in subject performance between the Ishihara and EHB for the full colour, 32 bit-greyscale, or blue channel conditions ($p = .35$, $p = .39$, $p = .22$, respectively). There were statistically significant differences in proportion of given expected answer choices for six individual plates using Fisher’s exact tests (Table 2; p -values < 0.01).

There were also qualitative differences in the range of answer choices provided for select plates and colour conditions (Table 2). There were no statistically significant differences observed when accounting for the

Table 1. Demographic characteristics of the study participants.

	Participants (N = 20)
Mean Age in Years (SD)	23.8 (2.19)
Sex, N (%)	
Female	12 (60%)
Male	8 (40%)
Race, N (%)	
White	7 (35%)
African American	4 (20%)
Asian	9 (45%)
Hispanic	1 (5%)
Corrective Lens Usage, N (%)	
Yes	16 (80%)
No	4 (20%)

Table 2. Side by side comparison of comparable EHB and Ishihara plates. The frequency with which different numerals were reported by study participants is shown for full colour, blue channel, and 32 bit-greyscale-manipulated images. The “expected answer” for each attempt was defined as the response given in the Ishihara colour plate key for each condition as outlined in Supplemental Figure 2. “np” represents that no pattern was detected, while “s” represents a squiggly line that is traceable from one side of the circle to the other. The “expected answer” is bolded in each of the participant-reported columns for which more than one answer choice was recorded. If the correct answer was not reported by study participants, it was included as a placeholder, bolded, and marked with an (*).

number on plate	expected answer	Ishihara		EHB		Ishihara		EHB		Ishihara		EHB	
		full colour	12 (100%)	full colour	12 (100%)	expected answer	32 bit-greyscale	32 bit-greyscale	32 bit-greyscale	expected answer	blue channel	blue channel	blue channel
12	12					12	np (100%)	np (65%)	12 (35%)	12	12 (100%)	12 (100%)	12 (100%)
8	8	8 (100%)	8 (100%)			np	np (100%)	np (100%)	np (100%)	3	n (70%) 8 (25%)	n (100%) 3 (0%)	n (100%) 3 (0%)
5	5	5 (100%)	5 (100%)			np	np (100%)	np (100%)	np (100%)	2	n (100%) 2 (0%)	8 (55%) n (40%) s (5%)	8 (55%) n (40%) s (5%)
29	29	29 (100%)	29 (95%) 23 (5%)			np	np (100%)	np (65%) 29 (10%) 20 (5%) 25 (5%) 28 (5%) 73 (5%) s (5%)	np (100%)	70	n (95%) 40 (5%) 70 (0%)	20 (45%) 29 (35%) 70 (10%) 73 (5%) 79 (5%)	2 (0%) 20 (45%) 29 (35%) 70 (10%) 73 (5%) 79 (5%)
74	74	74 (100%)	74 (100%)			np	np (90%) s (5%) 11 (5%)	np (100%)	np (100%)	21	71 (70%) 74 (20%) 21 (5%)	21 (100%)	21 (100%)
7	7	7 (100%)	7 (100%)			np	np (90%) s (10%)	7 (100%) np (0%)	np (0%)	np	np (100%)	np (90%) 7 (5%) s (5%)	np (90%) 7 (5%) s (5%)
45	45	45 (100%)	45 (100%)			np	np (70%) s (30%)	np (100%)	np (100%)	np	np (100%)	np (85%) s (15%)	np (85%) s (15%)
16	16	16 (100%)	16 (100%)			np	np (95%) s (5%)	16 (100%) np (0%)	16 (100%) np (0%)	np	np (95%) s (5%)	np (95%) 22 (5%)	np (95%) 22 (5%)
26	26	26 (100%)	26 (100%)			np	np (90%) s (5%) 16 (5%)	np (90%) 16 (5%) 26 (5%)	np (90%) 16 (5%) 26 (5%)	np	np (80%) 2 (10%) 28 (5%) 8 (5%)	np (100%) 2 (10%) 28 (5%) 8 (5%)	np (100%) 2 (10%) 28 (5%) 8 (5%)

number of different answer choices provided for each colour condition ($p = .34$, 0.78 , and 1.00 for full colour, 32-bit greyscale, and blue channel, respectively) or when comparing all Ishihara and EHB responses taken as a whole ($p = .79$).

Discussion

In this study, we characterised how subjects with normal colour vision responded to digital replications and colour manipulations of tools available to vision care providers in assessment of colour vision. Our study design enabled us to compare as closely as possible the EHB and Ishihara plates under the same visual manipulation scenarios, answering a specific question about the colour information contained in each slide and how that information was perceived by participants under different colour manipulation scenarios.

A study by Ozgur et al. conducted with healthy controls and subjects with ocular pathology found general agreement in colour vision testing results ascertained by EHB and physical Ishihara colour plates.⁷ Our study results reinforce and add to the findings of this previous study in distinct ways. When analysed as a whole, the differences in the proportion of expected answer choices given were not found to be statistically significant, meaning that the EHB is likely providing an overall assessment that is generally consistent with the traditional Ishihara plates. However, the differences we found between specific Ishihara and EHB plates in our testing highlights the possibility that certain individual plates with the same number may not be equivalent in assessment. This provides evidence for the premise that clinicians utilising the EHB should not do so with the expectation that the two modalities are exact replications of the Ishihara plates reporting the same information. Rather, use of the EHB should be considered a tool that provides a measure of colour information that is likely distinct from the Ishihara plates. Therefore, clinical or surgical decision-making that requires the assessment of a sudden change in colour vision may benefit from using the same modality consistently rather than using them interchangeably.

There are important limitations to consider when interpreting results from our study. Our image editing methodology was derived from a calibration study

regarding the use of RGB colour manipulation to simulate colour vision changes.⁹ We employed those image editing techniques in the present study as well as our initial pilot study of EHB slides conducted with two independent raters.⁸ In that study, we described qualitative differences between the Ishihara and EHB plates subject to colour manipulations. In this study, we have added to our previous findings by characterising perception of colour information by a set of 20 human subjects using an iPad Pro tablet computing device as opposed to an iPhone 6S as in our pilot study. Differences in screen size, brightness capability, and scale of images viewed are different between these two devices and thus may have biased our study results. It is also possible that the image processing methods that we employed do not precisely simulate the colour vision deficiency scenarios as assessed by the Ishihara plates, thus contributing to bias in our assessment. Further study is necessary to determine the suitability of these image processing modalities in simulation of colour vision, both in patients with and without colour vision deficiency.

It has also been established previously that there are contrast-related effects that should be taken into account when conducting pseudoisochromatic plate testing.¹⁰ For example, it is possible that pseudoisochromatic plates lose some of the colour information when the physical plate is photographed on a DSLR camera, subjected to digital image manipulation, and viewed on a high-resolution tablet screen. There is also the possibility that in patients who demonstrate R-G colour vision loss, there is a different representation of R-G colour information as opposed to the complete absence of that information that our image set conveyed. Future studies could aim to potentially investigate how to best simulate the R-G colour-blindness condition by testing gradations of maintaining colour information different than the present study.

Finally, one of the major limitations of our study was the fact that there was no answer key published for the EHB slides. Thus, we correlated the EHB slide to a specific slide in the Ishihara plates based on the number or pattern seen in the centre. As a result, there were some EHB slides that were tested in the study population that could not be compared to a corresponding Ishihara plate and thus were not included in the final analysis.

Our study results highlight the need for vigorous, evidence-based creation and validation of mobile applications intended for clinical use by eye care providers.

Declaration of interest statement

Allen O Eghari has ownership interest in Treyetech and LuckyVision, LLC. The remaining authors have no conflicts of interest to report.

Funding

Allen O Eghari is supported by the Research to Prevent Blindness Sybil B. Harrington Special Scholar Award and the Tolsma family. The remaining authors have no sources of funding.

References

1. Jackson & Coker Research Associates Special Report. App, doctors, and digital devices. Jackson Coker Ind Rep. 2011;4(7):1–3.
2. Research2guidance. *Mobile Health Market Trends and Figures 2013–2017, the Commercialization of mHealth Applications*. Vol. 3. 2013; Berlin, Germany: Research2Guidance. <https://research2guidance.com/wp-content/uploads/2015/08/Mobile-Health-Trends-and-Figures-2013-17-Preview.pdf>. Accessed April 1, 2020.
3. Eye Handbook (Ophthalmology For Your Smartphone). 2020. <http://eyehandbook.com/EHBweb>. Accessed April 1, 2020.
4. Oli A, Kumar S, Joshi D Is smart phone based Ishihara testing an alternative to ishihara booklet for color vision assessment? All India Ophthalmological Society Proceedings 2017. Jaipur, India. 16 February 2017.
5. Dain SJ, Kwan B, Wong L. Consistency of color representation in smart phones. *J Optical Soc Am*. 2016;33(3): A300–305. doi:10.1364/JOSAA.33.00A300.
6. Dain SJ, Al-Mderdef A. Colorimetric evaluation of iPhone apps for colour vision tests based on the Ishihara test. *Clin Exp Optom*. 2018;99:264–273. doi:10.1111/cxo.12370.
7. Ozgur OK, Emborgo TS, Vieyra MB, Huselid RF, Banik R. Validity and acceptance of color vision testing on smartphones. *J Neuroophthalmol*. 2018;38(1):13–16. doi:10.1097/WNO.0000000000000637.
8. Zhao J, Fliotsos M, Ighani M, Eghrari AO. Comparison of a smartphone application with Ishihara pseudoisochromatic plate for testing colour vision. *Neuro-Ophthalmology*. 2019;43(4):235–239. doi:10.1080/01658107.2018.1529187.
9. Pouw A, Karanija R, Sadun A. A method for identifying color vision deficiency malingering. *Graefes Arch Clin Exp Ophthalmol*. 2017;255(3):613–618. doi:10.1007/s00417-016-3570-0.
10. Annadanam A, Zhao J, Wang J, Eghrari AO. Effects of contrast sensitivity on color vision testing. *J Neuroophthalmol*. 2017;41(4):182–186. doi:10.1080/01658107.2017.1295273.