

NEW IPAD-BASED TEST FOR THE DETECTION OF COLOUR VISION DEFICIENCIES

Running title: Optopad for detection of CVD

Dolores de Fez¹

M^a José Luque²

Lucía Matea¹

David P. Piñero^{1,3}

Vicente J. Camps¹

¹Department of Optics, Pharmacology and Anatomy, University of Alicante, Spain

²Department of Optics, University of Valencia, Spain

³Department of Ophthalmology (Oftalmar), Vithas Medimar International Hospital, Alicante, Spain

Corresponding author: Dolores de Fez, dolores.fez@ua.es

ACKNOWLEDGEMENTS

The authors acknowledge the support of FUNCAVIS (Foundation for the Visual Quality, Alicante, Spain) for allowing participating in their lazy eye screening campaign in different schools by including the assessment of the colour vision with the Optopad test.

ABSTRACT

Purpose: To develop and validate a new iPad-based colour vision test (Optopad).

Methods: A total of 341 student eyes were enrolled in a first comparative study between Optopad and the Ishihara tests. In a second comparative study, Optopad vs. the Farnworth-Munsell test (FM 100H), a total of 66 adult eyes were included. Besides the agreement between tests, the correlation between FM 100H and Optopad outcomes were investigated. Multiple regression analysis was used to predict the Total Error Score (TES) from contrast thresholds measured with the Optopad test.

Results: The Ishihara and Optopad tests detected the same anomalous patients. Concerning FM 100H vs. Optopad, 10 subjects were diagnosed as anomalous with both tests, 3 mild anomalous cases based on TES were classified as normal with Optopad, and 2 anomalous subjects based on Optopad test showed normal TES values. Statistically significant correlations of TES and partial error red-green (PTESRG) with thresholds measured with the red-green Optopad stimuli were found. A multiple quadratic regression model was obtained relating TES and chromatic contrast values from Optopad ($R^2=0.855$), with only 13 cases showing residuals of ≥ 25 units.

Conclusions: The design and implementation of a chromatic contrast discrimination test has been carried out, with promising clinical results. This test seems to provide comparable outcomes to those obtained with Ishihara and FM 100H tests.

Keywords: iPad, chromatic discrimination, Optopad, Farnsworth-Munsell 100 Hue Test, Ishihara plates

INTRODUCTION

It is increasingly common to perform on a tablet different tasks that could be only performed before on a computer [1-4]. Tablets allows us to work comfortably and quickly, being also portable. However, there are doubts about the reliability of these devices to perform visual tests [5]. If a spatial and colorimetric characterization of the device is not performed to ensure reliable reproduction of the stimuli, the validity of the results obtained is questionable. Likewise, to use the characterization of a device in other devices can also lead to erroneous results.

Our research group has developed protocols for the spatial and colorimetric characterization of displays, applying them to different tablet and smartphone devices [5-7]. Spatial characterization is simple to perform and basically it should be considered that the design of the visual test must be customized for the size of the display that is intended to be used. Reproductions on screens of other sizes and/or resolution are not convenient as the zoom of the device can introduce significant modifications of the spatial design. The colorimetric characterization requires the calculation of the chromatic profile of the screen, either by means of a mathematical model (linear, quadratic, exponential ...) or by 3D LUT method (lookup table).

If we have a characterized device, we can design any visual test to be reproduced in it [5]. We can implement the algorithms based on the numerical description of the spatial, temporal and colorimetric characteristics of the desired stimulus and then transform them into the switch-on instructions submitted to each pixel of the screen to reproduce such stimulus. Once the characterization of a device is obtained, it is necessary to evaluate if algorithms work correctly and reproduce with an acceptable error the desired stimulus. The color difference can be used as the quality parameter of the colorimetric characterization [8].

Apple devices, which support a 3D LUT characterization, have acceptable colour reproduction errors [5]. The average colour differences obtained are less than 1 unit using the CIEDE2000 formula [9], a value very close to the minimum distinguishable for a normal subject [5,10]. Studies with more recent devices [11, in press] have shown colour differences of less than 0.4 units. Colour

reproduction errors considerably increase if, instead of characterizing our particular device, we assume the colorimetric profile of a similar one. For instance, our study of Retina screens in three different iPad models found errors of up to 8 units, far above the minimum distinguishable difference, when the colorimetric profile of a device was used in another. The possible consequences of cross reproduction errors can be understood for instance with one Ishihara plate: colors initially lying on the same confusion line could end in regions in color space different enough to be distinguishable for the anomalous or dichromatic subject the plate was originally designed for.

Our research group has recently designed a new test for colour vision assessment, called Optopad [12], based on the measurement of color discrimination thresholds using an iPad. The test implementation consists of a device-independent mathematical design stage, and a stimulus reproduction stage, in which the information contained in the 3D LUT of each specific device is used. Thus, the algorithms can be used for any device colorimetrically characterized using the same procedure. We have developed specific software providing a series of images to be sent to the device and to be viewed through the corresponding application on the iPad.

To this date, tests, such as Ishihara plates and the Farnsworth-Munsell 100 Hue Test (FM 100H), are usually used in clinical practice to evaluate the chromatic vision of patients [13,14]. Ishihara plates are used as a screening test, as they allow the clinician to discriminate quickly between normal subjects and subjects with red-green colour vision abnormalities. The FM 100H test performs a much more comprehensive analysis, diagnosing anomalies in both red-green and blue-yellow directions, which is a great advantage in acquired colour vision defects. There are several standard diagnostic criteria, including Dain & Birch's criterion and Vingrys' analysis, that have been used in the current study [15,16]. In addition to the patient's Total Error Scores (TES), partial error scores red-green and blue-yellow mechanisms can be computed (PTESRG and PTESBY) [17]. Both TES (Total Error Score) and PTES (Partial Total Error Score) vary with age, but not so Vingrys' parameters. Although normal average values have been defined for all of these parameters, only the normative databases for TES and Vingrys have been reported [13,15-18]. In recent years, other formulas

have been developed to evaluate the diagnostic ability of the FM 100H test through the combination of PTES and TES [17,19-22]. The limitation of these new criteria is the absence of normal databases allowing a statistical classification of cases.

The objective of the current study was to describe a new colour vision test (Optopad) developed by our research group and to pre-validate it clinically by comparing the diagnosis of colour vision abnormalities obtained with such test and that provided by commercial tests. As Optopad is designed to analyze the level of chromatic discrimination in both red-green and blue-yellow directions, our first comparison option was a test evaluating both chromatic channels, the FM 100H test. The choice of another test, such as the Ishihara test, would be justified for a population with impossibility of using a more complex test. As example, in a children population the FM 100H test methodology is very difficult to apply, and a test like Ishihara can be used, although assuming that only anomalies in the red-green direction can be detected.

METHODS

First, the colorimetric design of the test will be described, which is device independent, although access to the colorimetric characterization of the screen used in the measurements is required. The design of the test and the characterization of the visualization display were performed using proprietary software in Matlab® environment (R2008a).

Optopad test design

The Optopad test consists of three plates, called PDT1, PDT2 and RGBY1 (see figure 1).

Each of the PDT plates is formed by three rows of 10 Landolt-Cs at random orientations, on an achromatic background K with average luminance 60 cd/m^2 , subtending 1.7° . The stroke of the C and its opening subtend 0.23° and the whole letter 1.16° . The colour of the letter C belonging to the protan (row P), deutan (row D) or tritan (row T) confusion lines, respectively, passing through the background K. The chromaticity of the letter in row j and column i is defined

in the cone excitation space [23] (Equation 1), with the S cones scaled to ensure that $S(498\text{nm})=L(489\text{nm})+M(489\text{nm})$ [24]:

$$\begin{pmatrix} L_{C,i,j} \\ M_{C,i,j} \\ S_{C,i,j} \end{pmatrix} = \left(1 + C_{i,j} \begin{pmatrix} \delta_{1,j} & 0 & 0 \\ 0 & \delta_{2,j} & 0 \\ 0 & 0 & \delta_{3,j} \end{pmatrix} \right) \begin{pmatrix} L_K \\ M_K \\ S_K \end{pmatrix} \quad \text{Eq. 1}$$

In the equation, indexes $j=1,2,3$ stand for P, D and T, respectively, $i=1$ to 10 is the position of the letter in the sequence, $\delta_{n,j}$ is the Kronecker delta, (L_K, M_K, S_K) are the cone excitations for the background, and $C_{i,j}$ is the cone contrast between letter and background, defined as:

$$C_{i,j} = \sqrt{\left(\frac{\Delta L_{C,i,j}}{L_K}\right)^2 + \left(\frac{\Delta M_{C,i,j}}{M_K}\right)^2 + \left(\frac{\Delta S_{C,i,j}}{S_K}\right)^2} \quad \text{Eq. 2}$$

To minimize detections by an achromatic mechanism, the (L, M, S) vector at each point (x, y) of the image was multiplied by a factor $k(x, y)$, randomly chosen among 3 possible values (0.1, 0.55 and 1).

The contrasts used in the PDT1 and PDT2 plates are plotted in Figure 2. The contrast values for P and D are the same and change at constant logarithmic steps. The contrast values for T are the same for both plates. The differences between plate PDT1 and PDT2 are due to the fact that in PDT1 we are exploring the deutan confusion line along the red direction and in PDT2 along the green one. The device's colour gamut limits the values of the cone contrasts in the second case. The contrasts for T are larger to ensure a visibility similar to that of the P and T stimuli, and, although contrast also changes in constant logarithmic steps, the changes between samples 6 and 10 are smaller than between samples 1 and 5 to allow a finer gradation of normal subjects.

In Figure 3, we simulate how a protan subject would see plates PDT1 and 2, using the corresponding pair algorithm [25]. In both plates, the protanope would be unable to read the P line, but in the PDT1 plate there is a hue difference between background (yellow) and letter (blue), whereas in PDT2 both letter and background have the same hue and different colorfulness - the attribute of a visual perception according to which the perceived colour of an area appears to

be more or less chromatic- and brightness. For deutan subjects (not shown), the background and the letter in row P have the same hue and different colorfulness and brightness in both plates.

In the RGBY1 plate, the colour of the samples is chosen along one of the cardinal directions in the opponent modulation space [26], as defined in Eq. 3

$$\begin{pmatrix} \Delta A \\ \Delta RG \\ \Delta BY \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -\frac{L_K}{M_K} & 0 \\ 1 & 1 & -\frac{L_K+M_K}{S_K} \end{pmatrix} \begin{pmatrix} \Delta L \\ \Delta M \\ \Delta S \end{pmatrix} \quad \text{Eq. 3}$$

Rows R and G explore, respectively, the positive and negative axis along the red-green cardinal direction, whereas samples in the Y and B rows lie along the positive and negative axis, respectively, of the yellow-blue cardinal direction. Therefore, the cone excitations defining a sample at row j and column i are given by Eq. 4

$$\begin{pmatrix} L_K \\ M_K \\ S_K \end{pmatrix} + C_{i,j} \begin{pmatrix} L_K & M_K & 0 \\ M_K & -M_K & 0 \\ S_K & 0 & -S_K \end{pmatrix} \begin{pmatrix} 0 \\ (-1)^{j+1}(\delta_{1,j} + \delta_{2,j}) \frac{L_K}{\sqrt{L_K^2 + M_K^2}} \\ (-1)^j(\delta_{3,j} + \delta_{4,j}) \end{pmatrix} \quad \text{Eq. 4}$$

where $j=1$ for R, $j=2$ for G, $j=3$ for B and $j=4$ for Y. The values of the cone contrast, $C_{i,j}$, for each sample, are plotted in Figure 4. Contrast changes at constant logarithmic steps, but to allow fine grading of normal subjects, step-size is reduced after sample 4 in the G row, after sample 5 in the B row and after sample 7 in the Y row. Again, random luminance noise was introduced to minimize detections by an achromatic mechanism.

Clinical pre-validation of the Optopad test

All measurements in the clinical pre-validation study were conducted in a dark room, monocularly and at an observation distance of 50 cm. Conventional tests, Ishihara and FM 100H tests, require a controlled illuminant, and therefore a Color Viewing Light BASIC with Daylight Illuminant (D65) was used. The

Optopad test was presented in all cases on an iPad4 tablet (Apple, Cupertino, CA, USA, 2013) with Retina screen (9.7 inches, 1536x2048 pixels and 264 dpi). This screen is based on the IPS (in-plane switching) technology, a specific LCD (Liquid Crystal Display) type with an LED (light emitting diode) backlight. This device had been colorimetrically characterized by 3D LUT in previous studies [5,27]. During the measurement with the test, the patient's task was to indicate which was the direction of the opening of the Landolt rings in each row of each plate. Measurements were carried out monocularly and a single eye per patient, randomly selected, was tested.

First, a repeatability analysis of the measures obtained with the Optopad test was performed, involving 10 subjects who took three times the three plates of the test in our laboratory. To avoid possible memorization effects, the orientation of the letters was changed in each measurement. By the Shapiro-Wilk test, we verified that data did not follow the normal distribution. The non-parametric Friedman test for K related samples revealed that there were no statistically significant differences between the three measurements for any of the lines ($p > 0.05$), concluding that repeatable measures could be obtained with the test.

For the comparison of the Optopad and Ishihara plate tests (24 plates, 2002), a total of 341 eyes of students aged between 3 and 11 years (55% boys and 45% girls) were enrolled in a first comparative study after confirming that younger patients were able to read numbers and to identify correctly right, left, up and down directions with their hands. Total measurement time did not exceed 5 minutes in any case. It was considered that the patient presented an alteration of color vision if he had more than 5 failures during the test. The Ishihara plates 22-25 were used, as suggested by the test guidelines, for the characterization of the type of colour vision alteration (protan-deutan), since one of the numbers of each plate belongs to a deutan confusion line and the other to a protan.

For the comparison between the Optopad and FM 100H tests, a total of 66 eyes of adult patients (34% men and 66% women) aged between 20 and 52 years, with a median of 22 years, were enrolled. We searched first for a sufficiently large number of anomalous patients (12 + 4 suspects) and we established afterwards a population of comparison of 50 healthy subjects. For this reason,

the percentage of anomalous patients and the gender distribution differ from the commonly reported in previous studies.

The total duration of the measurement session was below 20 minutes, including breaks to avoid the fatigue of the observer if necessary. For the calculations of the outcomes of the FM 100H test, we decided to use the following formulas: a combination of the square roots of the PTES already introduced by Smith [17] (Eq. 5) and a new index resulting from the quotient between PTESRG and the total error [19] (Eq. 6)

$$- \sqrt{PTESBY} - \sqrt{PTESRG} \quad (\text{Eq. 5})$$

$$- \frac{PTESRG}{TES} \quad (\text{Eq. 6})$$

In both cases, we decided to apply the normality pattern defined for each formula by the authors that developed them [19]. Additionally we have obtained the parameters of Vingrys analysis: angle (indicates type of anomaly) and C-index (severity of colour loss) [16,28]:

The calculations for obtaining the results of the FM 100H test, the normal pattern of the Optopad test and comparisons with the standards of both tests were performed using our own software in Matlab® environment (R2008a).

The study adheres to the tenets of the Declaration of Helsinki for Research Involving Human Observers and the project was approved by the ethics committee of the University of Alicante. The informed consent was obtained from the subjects after explanation of the nature and possible consequences of the study.

Statistical analysis

In the validation analysis with the Ishihara test (341 eyes), we applied the Pearson chi-square test to compare the percentage of anomalous patients detected with the colour tests, and the McNemar Test to compare between the protan / deutan / not classifiable diagnosis of anomalous eyes. Tritan defects were excluded from the analysis, since they cannot be detected by the Ishihara plates.

In the validation analysis with the FM 100H test (66 eyes), we evaluated the normality of the sample using the Kolmogorov-Smirnov test. The bivariate correlations of TES and Optopad diagnoses were evaluated using the Spearman correlation coefficients due to the non-normality of all data distributions. Furthermore, a multiple linear regression analysis was performed by using the backward elimination method, with the purpose of obtaining a mathematical expression relating the chromatic thresholds measured by Optopad with TES. Additionally, multiple quadratic regression analysis was performed in order to check if a better fit might be obtained. For all these statistical calculations, the IBM SPSS Statistics 23 software was used.

RESULTS

Comparative study Ishihara vs. Optopad

The Ishihara test detected six patients with problems of colour vision. One child was unable to read any number, four children only read the four numbers corresponding to the Deutan-type anomaly in plates 22 to 25 and one child only read two of these four numbers.

These six patients detected with the Ishihara tests were able to read easily the T, B and Y lines of the Optopad test (corresponding to blue-yellow anomalies), but showed great difficulties in reading the stimuli of the P, D, R and G lines (red-green anomalies), failing to identify most of the optotypes. Two children identified up to optotype 4 in the PDT1 protan line, and two identified optotype 3 in the protan line of PDT1 and PDT2. In all other lines, children responses ranged from 'I do not see' to 'I can see until optotype 2'. According to these results, these six patients presented a clearly anomalous behavior with our Optopad test, when compared with the results obtained in the rest of subjects. A statistically significant dependence between tests was found in terms of percentage of cases identified as normal or anomalous ($p < 0.01$).

Normality intervals for each direction in colour space were obtained for our Optopad test with the total population, excluding the 6 anomalous eyes. An example of the results obtained for one of the anomalous patients can be seen

in Figure 5, which shows the normality region (gray area) and the discrimination thresholds of the patient (coloured lines) in the CIE1931xy space. For this subject, the results in PDT1 and PDT2 plates are out of normal limits, more evidently so along the deutan confusion line. In RGBY1, this anomalous behavior corresponded to the red-green mechanism in both directions.

The six anomalous patients showed red-green chromatic thresholds outside the normal range with all Optopad plates. Four of these eyes showed a more normal behavior on the D-line of PDT1 and PDT2. Only three of the anomalous cases were then coincident with the deutan diagnosis indicated by the Ishihara test. No statistically significant dependence between Ishihara and Optopad tests was found in terms of percentage of protan-deutan or not-classifiable cases ($p>0.05$).

Comparative study Farnsworth-Munsell 100 Hue Test vs. Optopad

This comparative study was performed with a population of 66 eyes. The median, range and confidence interval of the mean for parameters TES, PTESRG, PTESBY, $PTESBY^{1/2}$ - $PTESRG^{1/2}$, $PTESRG/TES$ and C-index are summarized in Table 1 (colour normal and colour anomalous population separately). This table contains also the Optopad results for the same populations.

In the sample evaluated, a total of 12 eyes with an anomalous TES were found. From these patients, only 7 showed an anomalous PTESRG value and only 1 an anomalous PTESBY value. The remaining 4 subjects were out of normal limits in both PTES and therefore could not be classified. In such cases, we only can state that there is a low level of chromatic discrimination. Two eyes showed an anomalous behavior for PTESRG, but not for TES. These cases were finally excluded from the normal population, increasing the number of anomalous eyes to 14. Furthermore, two cases yielding clearly anomalous results with the Optopad test but not with the FM 100H test were detected. These two subjects had been previously diagnosed as having anomalous colour vision. These cases were also excluded from the normal population. According to this, a total of 16 eyes with either clear or suspected colour vision anomaly were found.

Of the 14 eyes with non-normal TES or PTES values, 11 were diagnosed as protan and the remaining three were classified as deutan according to the angle parameter of Vingrys. In our population, no eyes classified as tritan by Vingrys analysis was detected. One single eye showed errors in TES and PTESBY, but it was surprisingly classified as protan according to the Vingrys analysis.

The following results were obtained when comparing the FM 100H and Optopad tests, calculating the normal pattern for our Optopad test with the total population, excluding the 16 anomalous eyes:

- In the 12 eyes with anomalous TES, 6 presented a total agreement in the diagnosis obtained with the FM 100H and Optopad tests (four with errors in both RG and BY directions, two only in the red-green direction), 3 showed a partial agreement (red-green error in FM100 H whereas in Optopad test errors in both directions) and the remaining 3 were classified as normal by the Optopad test (two with alterations in the red-green mechanism according to the FM100H and the other with alterations in the blue-yellow mechanism).

- For the 2 eyes with normal TES but anomalous PTESRG, Optopad test classified one case as normal, whereas the other was classified as anomalous in both the RG and BY directions. The agreement is partial in both cases.

- The Optopad test classified two eyes with normal TES and PTES as anomalous, one red-green and the other in both directions.

In summary, 11 subjects were diagnosed similarly with both tests, 3 anomalous cases according to TES were classified as normal with the Optopad test and 2 anomalous subjects according to the Optopad test showed normal TES values.

Correlation between outcomes with Farnsworth-Munsell 100 Hue and Optopad

Statistically significant but poor correlations of TES with thresholds along red-green lines of the Optopad test (P1, D1, P2, D2, R1, G1) were found ($p < 0.40$, $p \leq 0.01$). Furthermore, weak but significant correlations of TES with the thresholds along the blue-yellow lines of the Optopad test (T1, T2, B1, Y1) were observed ($p < 0.40$, $p \leq 0.04$).

Besides this, the correlations of PTESRG with the thresholds along the red-green lines of the Optopad test were moderate and statistically significant ($-0.453 \leq \rho \leq 0.585$, $p < 0.01$), whereas the correlations of this parameter with the thresholds along blue-yellow lines were weak ($\rho < 0.40$, $p \leq 0.02$) (not significant with T1, $\rho: 0.240$, $p: 0.053$). Likewise, poor but statistically significant correlations of PTESBY with the thresholds along blue-yellow lines of the Optopad test ($\rho < 0.40$, $p \leq 0.04$) as well as with the red-green lines were found ($\rho < 0.40$, $p \leq 0.01$), except with P1 ($\rho: 0.205$, $p: 0.09$) and D1 ($\rho: -0.232$, $p: 0.06$).

Concerning C, it was found to be correlated significantly with all lines of the Optopad test ($-0.386 \leq \rho \leq 0.871$, $p \leq 0.04$).

Poor correlations of $PTESBY^{1/2}$ - $PTESRG^{1/2}$ and $PTESRG/TES$ with P1 and D1 were obtained ($\rho < 0.40$).

The results obtained in this first analysis suggested that could be a complex relationship between the diagnosis obtained with FM 100H and our Optopad test. This leads us to consider in the following section a multiple regression analysis.

Multiple regression analysis

A multiple regression analysis was performed to find the appropriate mathematical expression relating the TES value (only, to simplify) with the chromatic contrast values of the Optopad test. This study was performed again with a population of 66 eyes. The purpose was to find a model with the ability of predicting the TES value from the discrimination values obtained in each of the three plates of the Optopad test. The Appendix shows the results for a linear simple model (L) and a more complex model including quadratic terms (QL). We also show the predictability value R^2 and the cases of residues lower than 25 units.

In order to verify the level of agreement between the real TES and predicted TES^m in the two models developed, a Bland-Altman analysis was done for each model, representing the difference $TES - TES^m$ against the average value (figure 6). As shown, limits of agreement were similar in the two models, although slightly smaller with QL (<50 units). Furthermore, the dispersion was higher for

low values of TES, becoming minimal for mean and high values in the model including quadratic terms.

DISCUSSION

In color vision testing, patients must perform a task based on the response of the visual mechanism presumed to be damaged or missing. If we restrict ourselves to those tests that grade damage in the chromatic mechanisms, two different design strategies can be observed: either to explore all directions of color space and determine the region or regions where the patient suffers greater sensitivity loss, as in the Cambridge Color Test (CCT) [29] or in the Color Assessment and Diagnosis (CAD) test [30] or to evaluate the patient's capabilities along a few directions, assumed to be those isolating the visual mechanisms of interest. Although these directions are usually those isolating the response of a particular cone type, as in the anomaloscope, the Trivector version of the CCT [29] or in the Arden test [31,32], for acquired color defects it is more efficient to explore the directions isolating the red-green and the blue-yellow mechanisms [33]. In both design strategies, there is a risk of unwanted responses from the achromatic mechanisms, which, even if the equiluminance condition for each observer could be determined, is usually minimized by introducing static [29] or dynamic [30,32] achromatic noise, although also tests relying on the isoluminant condition derived for an average observer have been proposed [34]. In the mechanism-isolating strategy, the problem of how to determine the directions isolating the cones or the opponent mechanisms in spite of inter-observer variability must be solved. Although it is usual to assume directions computed from average cone-sensitivities, other strategies, such as exploring a range of directions around the assumed average directions to account for observer variability, have been used [35].

Our design follows the mechanism-isolating strategy, with plates for congenital and acquired color defects. The stimuli assumed to favour each particular mechanism have been computed with color vision models (the average cone sensitivities determined by Smith and Pokorny and the red-green and blue-

yellow sensitivities derived from the opponent modulation space [26]). Our study shows that testing subjects along these directions allows detection of colour vision defects revealed by other techniques.

In the first part of the validation of the Optopad test, we found that it is able to detect the same anomalous cases as the Ishihara test, with 100% of agreement between tests in terms of classification as normal/non-normal. The agreement between tests in terms of the diagnosis of protan/deutan type anomaly was 50%. Birch et al. [36] found that the diagnostic plates 22-25 of the Ishihara test were more efficient for the deutan classification than the rest in a comparative study with the Nagel anomaloscope. In our series, Ishihara and Optopad agreed in the diagnosis of three deutan subjects, suggesting the validity of Optopad to detect of deutan anomalies.

In the second part of the validation of the Optopad test, we compared the outcomes of our test with those obtained with the FM 100H test, obtaining a lower level of agreement in both diagnosis and classification. Of the 16 suspect eyes, 3 classified as anomalous according to TES were classified as normal by the Optopad and 2 anomalous according to Optopad were classified as normal according to TES. As in other comparative studies of colour vision tests, some level of discrepancy was observed [16,36-38]. Vingrys [16] found in a comparative study between Nagel anomaloscope and FM 100H tests that the FM 100H test was limited in its ability of diagnose the type and severity of a colour vision defect, especially in congenital cases, with only 26% of deuteranomals being correctly classified and 38% being misclassified as protan. Birch et al. [37] found that half of anomalous trichromats obtained error scores of less than 100 with the FM 100H, without an axis of confusion. These subjects could be mistakenly identified as having normal colour vision. In a study with the Ishihara plates [36], Birch and coauthors found that the hidden digit designs (18-21 plates) were only able to detect approximately 50% of color-deficient subjects. Seshadri et al. [38] did not find a perfect agreement between the Nagel anomaloscope, used as a gold standard, and some colour vision tests: sensitivity was 96% and 100% with Ishihara and FM 100H tests, respectively, specificity was 100% and 83%, and the coefficient of agreement was 0.96 and 0.83, respectively. For these reasons, authors recommended a convenient

battery of tests for a correct diagnosis of colour vision defects, a recommendation that we support.

The efficacy of the Optopad test in diagnosing colour vision defects compared to the FM 100H test (using only TES as gold standard) in the total adult population was 75% (9/12) and the efficacy in detecting cases with normal colour vision was 94.4% (51/54). These results are, of course, strongly dependent on the normality region considered for the Optopad test. Future studies with very large samples should be conducted to define more consistent patterns of normality for the Optopad as well as for the most commonly used colour vision tests. In addition, the main weakness of the current validation of the Optopad test in our adult population may be the small number of anomalous subjects who participated in the study.

To provide a different approach in addition to the direct comparison of the diagnosis provided by the TES according to the normative database defined by Kinnear and Sahraie [15] and that provided by the Optopad test according to our own normative database (calculated according to the results of 50 subjects), the calculation of a consistent model of prediction of TES (TES^m) from the quantitative measure of color discrimination values of the Optopad test was investigated. Specifically, a quadratic + linear model was the option providing the best adjustment ($R^2 = 0.855$), and therefore reproducing more accurately the TES values provided by the FM 100H test, with a smaller number of cases showing prediction errors of more than 25 units (13 cases = 20% of population). Furthermore, according to Bland & Altman analysis, the best agreement between TES^m and TES was also obtained for this model combining quadratic and linear terms.

Finally, there are other advantages of the Optopad tests over the Ishihara and FM 100H tests that should be mentioned. The test is quick and easy to perform at all ages, without the requirement of a controlled light cabinet to perform the test accurately. However, it should be considered that the display used to show the stimuli of the Optopad test must be colorimetrically characterized in order to ensure a reliable colour vision assessment. On the other hand, the use of a tablet to show the stimuli can be attractive and motivating for most of subjects, especially for children. Children with difficulty in reading the numbers of the

Ishihara plates can benefit from using a test only based on the identification of the directions of the openings of Cs presented on the tablet (they can use their own hands to indicate the orientation). Likewise, younger patients are not able to perform the FM 100H test because it has more complicated instructions and requires a high level of attention. In addition, the long time needed to sort the four boxes induces a fast drop in the patient's level of concentration. Our experience with this test has shown that not all age groups can take the FM 100H test with precision. Simpler and faster tests could have been more adequate for use with children, such as D15 or Roth 28 Hue Color Test, both reduced versions of the FM 100H. A numerical parameter (the total error score) could still be used to assess the results, but the colour difference between samples is larger than in the FM 100H, reducing sensitivity and accuracy. For these reasons, we aimed at developing a new test based on simpler instructions than reading numbers. In addition, the fact that each plate is formed by independent lines in which a specific hue is used avoids the presence of artefacts, such as those reported by Viliunas et al. [39] that originated from separation of the 85 correlative parts of FM 100H test into four independent boxes to be ordered between the first and last piece of each box. This fact may not affect subjects with good colour discrimination, but it can affect large TES.

In conclusion, a new fast and easy to perform test for assessing colour vision has been developed, called Optopad. According to our clinical pre-validation, this test allows detecting colour vision anomalies, especially in the red-green direction, in acceptable agreement with other conventional tests, such as Ishihara or FM 100H tests. Only some discrepancies were found between Optopad and FM 100H tests in some cases near the limit of the range defined by the literature for classifying a case as normal or anomalous. This should be investigated in future studies including larger samples of patients with colour anomalies comparing both tests Optopad vs. FM 100H. In addition, the results achievable with the FM 100H test can be predicted considering those obtained with the Optopad test within a clinically reasonable range. Future studies should confirm the consistency of this prediction model with larger sample sizes including more cases of colour vision anomalies, especially blue-yellow defects.

Funding: No funding was received for this research.

Conflict of Interest: All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Ethical approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

REFERENCES

- 1: Black JM, Jacobs RJ, Phillips G, Chen L, Tan E, Tran A, Thompson B (2013) An assessment of the iPad as a testing platform for distance visual acuity in adults. *BMJ Open* 3(6) pii: e002730.
- 2 : Kollbaum PS, Jansen ME, Kollbaum EJ, Bullimore MA (2014) Validation of an iPad test of letter contrast sensitivity. *Optom Vis Sci* 91:291-296.
- 3: Obradovic IV, Cappelli R, Priluck JC, Chalam KV, Grover S (2012) Comparison of color vision testing by standard Ishihara Color Plates versus iPad version. *Invest Ophthalmol Vis Sci* 53:E-Abstract 6394.
- 4: LC Nguyen, E Yi-Luen Do, A Chia, Y Wang, H Benn-Lird Duh (2014) DoDo Game, a color vision deficiency screening test for young children. Paper presented at the ACM CHI Conference on Human Factors in Computing Systems, Toronto, 26 April-1 May 2014.
- 5: de Fez D, Luque MJ, Garcia-Domene MC, Camps V, Piñero D (2016) Colorimetric characterization of mobile devices for vision applications. *Optom Vis Sci* 93:85-93.
- 6: Garcia-Domene MC (2013) Design and testing of a campimeter of incremental threshold for projection. Dissertation. University of Alicante, Spain <http://rua.ua.es/dspace/handle/10045/35676>. Accessed 22 March 2018.
- 7: García-Domene MC, Luque MJ, de Fez MD (2015) Software [Design of a campimeter of incremental threshold for projection]. Property registration 09 / 2015 / 447; University of Alicante, Spain.
- 8: Stokes M, Fairchild MD, Berns RS (1992) Precision requirements for digital color reproduction. *ACM Trans Graph* 11:406-422.
- 9: Luo MR, Cui G, Rigg B (2001) The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Res Appl* 26:340–350.
- 10: Day EA, Taplin L, Berns RS (2004) Colorimetric characterization of a computer-controlled liquid crystal display. *Color Res Appl* 29:365Y73.

- 11: de Fez MD, Luque MJ, García-Domene MC, Caballero MT, Camps VJ (2018) Can applications designed to evaluate visual function be used in different iPads? *Optom Vis Sci* (in press).
- 12: de Fez MD, Luque MJ (2017) Optopad Software. <http://hdl.handle.net/10045/69698>. Accessed 22 March 2018.
- 13: Dain SJ (2004) Clinical colour vision test. *Clin Exp Optom* 87:276-293.
- 14: Birch J (1991) Colour Vision Tests: General classification. Vol 7 in *Vision and visual dysfunction*. General editor Cronly-Dillon JR, MacMillan Press, Boca Raton, FL.
- 15: Kinnear PR, Sahraie A (2002) New Farnsworth-Munsell 100 hue test norms of normal observers for each year of age 5-22 and for age decades 30-70. *Br J Ophthalmol* 86:1408-1411.
- 16: Vingrys, AJ, Atchison, DA, Bowman, KJ (1992) The use of colour difference vectors in diagnosing congenital colour vision deficiencies with the Farnsworth-Munsell 100-hue test. *Ophthalmic Physiol Opt* 12:38-45.
- 17: Smith VC, Pokorny J, Pass AS (1985) Color-axis determination on the Farnsworth-Munsell 100-hue test. *Am J Ophthalmol* 100:176-182.
- 18: Mäntyjärvi M (2001) Normal test scores in the Farnsworth–Munsell 100 hue test. *Doc Ophthalmol* 102:73-80.
- 19: García-Domene MC, Díez Ajenjo A, de Fez Saiz D, Luque Cobija MJ (2016) Base de datos normativa para las puntuaciones parciales rojo-verde y azul-amarillo del test Farnsworth-Munsell 100Hue. Paper presented at the XI Congreso Nacional de Color, Universidad de Vigo, Spain, 19-22 July 2016 http://color2016.laserphotonics.org/site/web/libro_color.pdf. Accessed 22 March 2018.
- 20: Kitahara K, Kandatsu A (1985) Determining a standard for the FM 100-hue test in cases where there may or may not be a clear orientation axis of the hue errors. *Nippon Ganka Gakkai Zasshi* 89:1089-1093.
- 21: Dain SJ, Birch J (1987) An averaging method for the interpretation of the farnsworth-munsell 100-hue test—I. Congenital colour vision defects. *Ophthalmic Physiol Opt* 7:267-280.

- 22: Barton FB, Fong DS, Knatterud GL (2004) Classification of Farnsworth-Munsell 100-hue test results in the early treatment diabetic retinopathy study. *Am J Ophthalmol* 138:119-124.
- 23: Smith VC, Pokorny J (1996) The design and use of a cone chromaticity space: a tutorial. *Color Res Appl* 21:375–383.
- 24: Boynton RM (1986) A system of photometry and colorimetry based on cone excitations. *Color Res Appl* 11: 244–252.
- 25: Capilla P, Díez-Ajenjo MA, Luque MJ, Malo J (2004) Corresponding-pair procedure: a new approach to simulation of dichromatic color perception. *J Opt Soc Am A Opt Image Sci Vis* 21:176-186.
- 26: Derrington A M, Krauskopf J, Lennie P (1984) Chromatic mechanisms in lateral geniculate nucleus of macaque. *J Physiol* 357:241–65.
- 27: de Fez MD, Luque MJ (2017) 3DLUT for iPad 4 (cubic interpolation) <http://rua.ua.es/dspace/handle/10045/69708>. Accessed 22 March 2018.
- 28: Vingrys AJ, King-Smith PE (1988) A quantitative scoring technique for panel tests of color vision. *Invest Ophthalmol Vis Sci* 29:50-63.
- 29: Regan BC, Reffin JP, Mollon JD (1994) Luminance noise and the rapid determination of discrimination ellipses in colour deficiency. *Vis Res* 34:1279–1299.
- 30: Barbur JL, Rodriguez-Carmona M, Harlow A (2006) Establishing the statistical limits of “normal” chromatic sensitivity. Paper presented at the CIE Expert Symposium, National Research Council of Canada, Ottawa, 16-17 May 2006.
- 31: Arden GB, Gunduz K, Perry S (1988) Colour vision testing with a computer graphics system. *Clin Vision Sci* 2:303–320.
- 32: Arden GB, Wolf JE (2004) Colour vision testing as an aid diagnosis and management of age related maculopathy. *Br J Ophthalmol* 88:1180–1185.
- 33: O’Neill-Biba M, Sivaprasad S, Rodriguez-Carmona M, Wolf JE, Barbur JL (2010) Loss of chromatic sensitivity in AMD and diabetes: A comparative study. *Ophthalmic Physiol Opt* 30:705–716.
- 34: Rabin J, Gooch J, Ivan D (2011) Rapid quantification of color vision: The cone contrast test. *Investig Ophthalmol Vis Sci* 52:816–820.

- 35: Rodríguez-Carmona M, Sharpe LT, Harlow JA, Barbur JL (2008) Sex-related differences in chromatic sensitivity. *Vis Neurosci* 25:433–40.
- 36: Birch J (1997) Efficiency of the Ishihara test for identifying red-green colour deficiency. *Ophthalmic Physiol Opt* 17:403-408.
- 37: Birch J (1989) Use of the Farnsworth-Munsell 100 Hue test in the examination of congenital colour vision defects. *Physiol Opt* 9:156-162.
- 38: Seshadri J, Christensen J, Lakshminarayanan V, Bassi CJ (2005) Evaluation of the new web-based “Colour Assessment and Diagnosis” test. *Optom Vis Sci* 82: 882-885.
- 39: Viliunas V, Lukauskiene R, Svegzda A, Zukauskas A (2006) End-box scoring artefact evaluation of the Farnsworth-Munsell 100-Hue colour vision test. *Ophthalmic Physiol Opt* 26:580–586.

FIGURE LEGENDS

Figure 1. Optopad test plates. 1a, upper: PDT1, 1b, center: PDT2, 1c, lower: RGBY1. See the text for a description of the specific characteristics of the test, allowing for colour reproduction errors.

Figure 2. Cone contrast between sample and background in the two PDT plates of the Optopad test. Note that the colour contrast for the tritan rows is identical in the two plates.

Figure 3. Part of the PDT1 (left) and PDT2 (right) tests as seen by a protanope according to a simulation using the corresponding pair algorithm by Capilla and co-workers [25], allowing for colour reproduction errors.

Figure 4. Cone contrast between sample and background in the RGBY1 plate of the Optopad test.

Figure 5. Chromatic diagram CIE1931xy showing the results for the Optopad test for one of the six children classified as anomalous with the Ishihara test. Grey: normality region. Coloured lines: chromatic thresholds of the subject (left: PDT1, center: PDT2, right: RGBY1).

Figure 6. Bland-Altman plots comparing FM 100H total error score (TES) with modelled error scores (TES^m) based on the Optopad test results for a) linear, and b) linear + quadratic models.

Table 1: Median, range and confidence interval of mean for FM 100H parameters and each line of the Optopad test. A: anomalous subjects classified according to TES, N: 50 normal subjects.

	median A	range A	conf interv	median N	range N	conf interv
TES	94	[300,80]	40.33	32	[76,8]	5.20
PTESRG	64	[192,30]	28.32	16	[40,4]	2.33
PTESBY	47	[108,14]	15.54	18	[48,0]	3.59
BY^{1/2}- RG^{1/2}	-0.90	[2.40,-5.09]	1.29	0	[3.47,-4.00]	0.39
RG/TES	0.57	[0.85,0.33]	0.08	0.50	[1.00,0.20]	0.05
C	1.67	[3.2,1.47]	0.30	1.26	[1.68,1.06]	0.04
P1	0.071	[0.307, 0.027]	0.055	0.035	[0.071,0.016]	0.004
D1	0.116	[0.027,0.815]	0.174	0.027	[0.027,0.071]	0.003
T1	0.577	[2.5,0.118]	0.377	0.207	[0.940,0.118]	0.035
P2	0,072	[0.305,0.021]	0.058	0.021	[0.046,0.015]	0.002
D2	0,175	[0.305,0.021]	0.080	0.021	[0.046,0.015]	0.002
T2	0.465	[4.077,0.118]	0.627	0.245	[0.577,0.118]	0.026
R1	0.151	[0.375,0.029]	0.092	0.029	[0.042,0.014]	0.002
G1	0.090	[0.201,0.019]	0.043	0.019	[0.030,0.012]	0.001
B1	0.374	[1.738,0.064]	0.331	0.224	[0.625,0.064]	0.030
Y1	0.198	[1.111,0.042]	0.190	0.148	[0.351,0.042]	0.016

Appendix

Multiple regression analysis relating the TES value with the chromatic contrast values of the Optopad test (PDT1, PDT2, RGBY1). Equation of linear model (L) and a linear + quadratic model (QL). Value of R^2 and the cases of residues lower than 25 units.

Model	Equation	R^2	Res<25
Linear	$TES_L^m = 2.6 + 109.8 * D1 + 64.3 * T1 - 322.7 * D2$ $- 74.4 * T2 + 825.2 * R1 + 183.6 * Y1$	0.777	68%
Linear + Quadratic	$TES_{QL}^m = 34.1 + 660.8 * P1 - 6451.6 * P1^2 + 708.3$ $* D1 + 955.6 * D1^2 + 9.9 * T1 + 67.0$ $* T1^2 - 435.9 * P2 + 4767.0 * P2^2$ $+ 460.3 * D2 - 3068.1 * D2^2 - 29.5$ $* T2 - 75.8 * T2^2 - 229.5 * R1$ $+ 2545.6 * R1^2 - 1169.5 * G1$ $+ 1167.0 * G1^2 + 27.3 * B1 - 8.8 * B1^2$ $- 107.0 * Y1 + 404.6 * Y1^2$	0.855	80%

Figure 1a

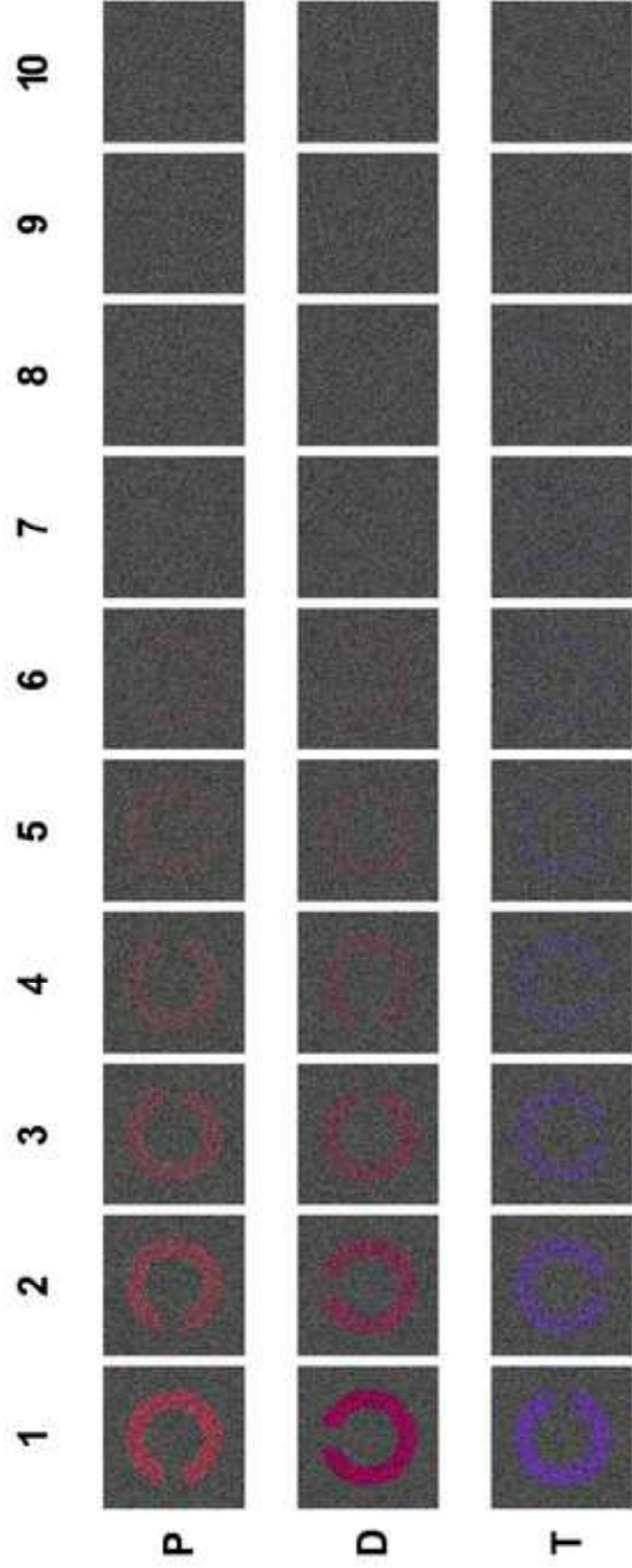
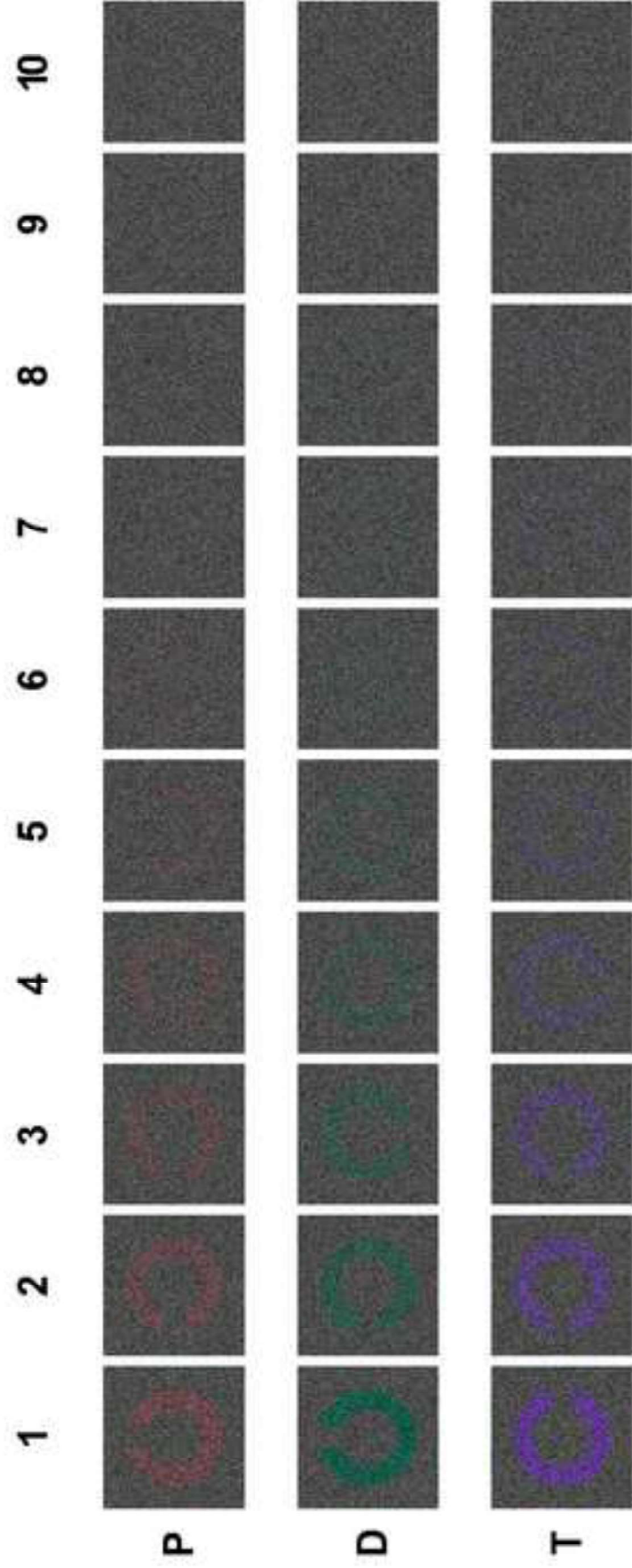


Figure 1b



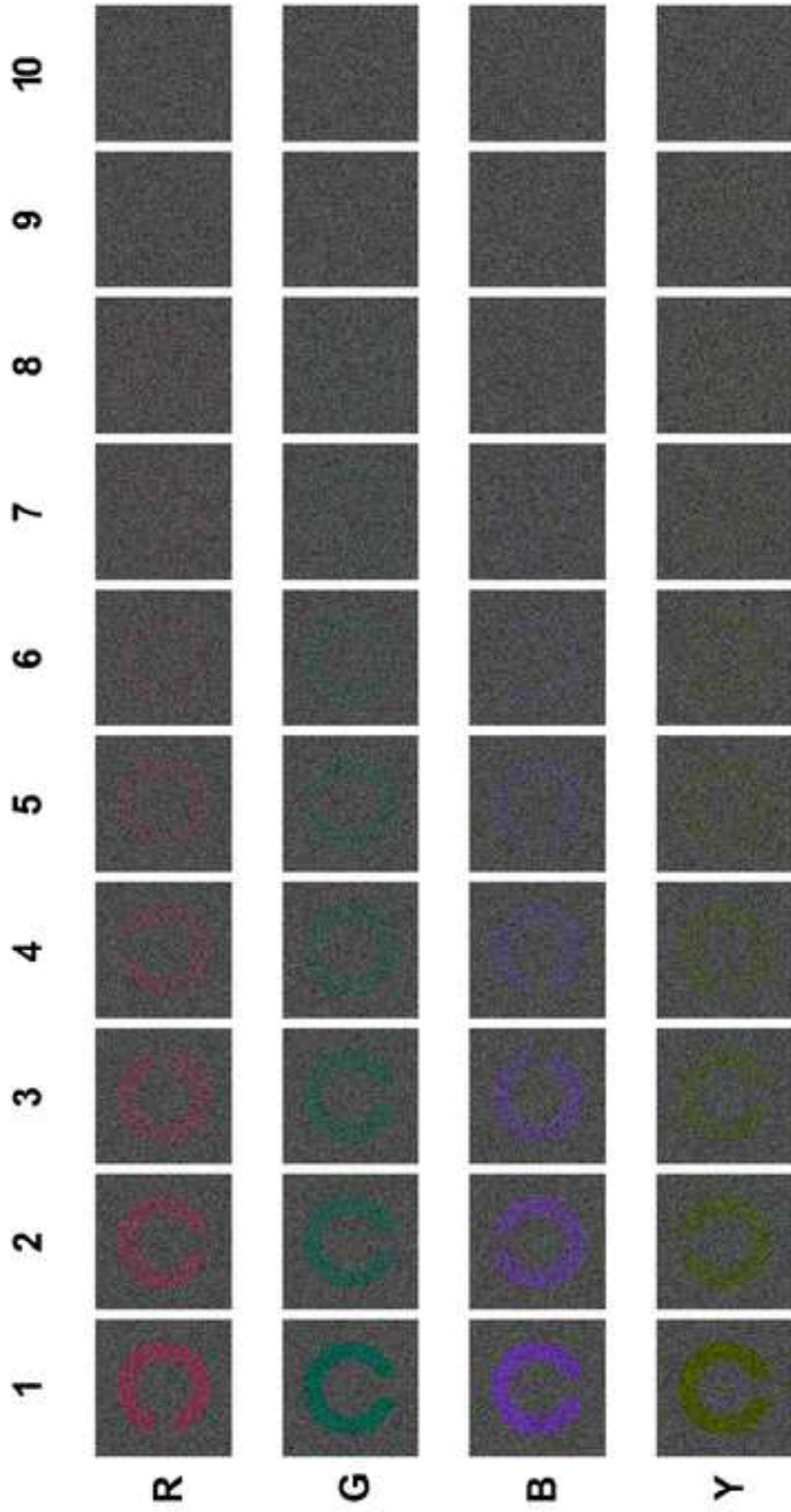
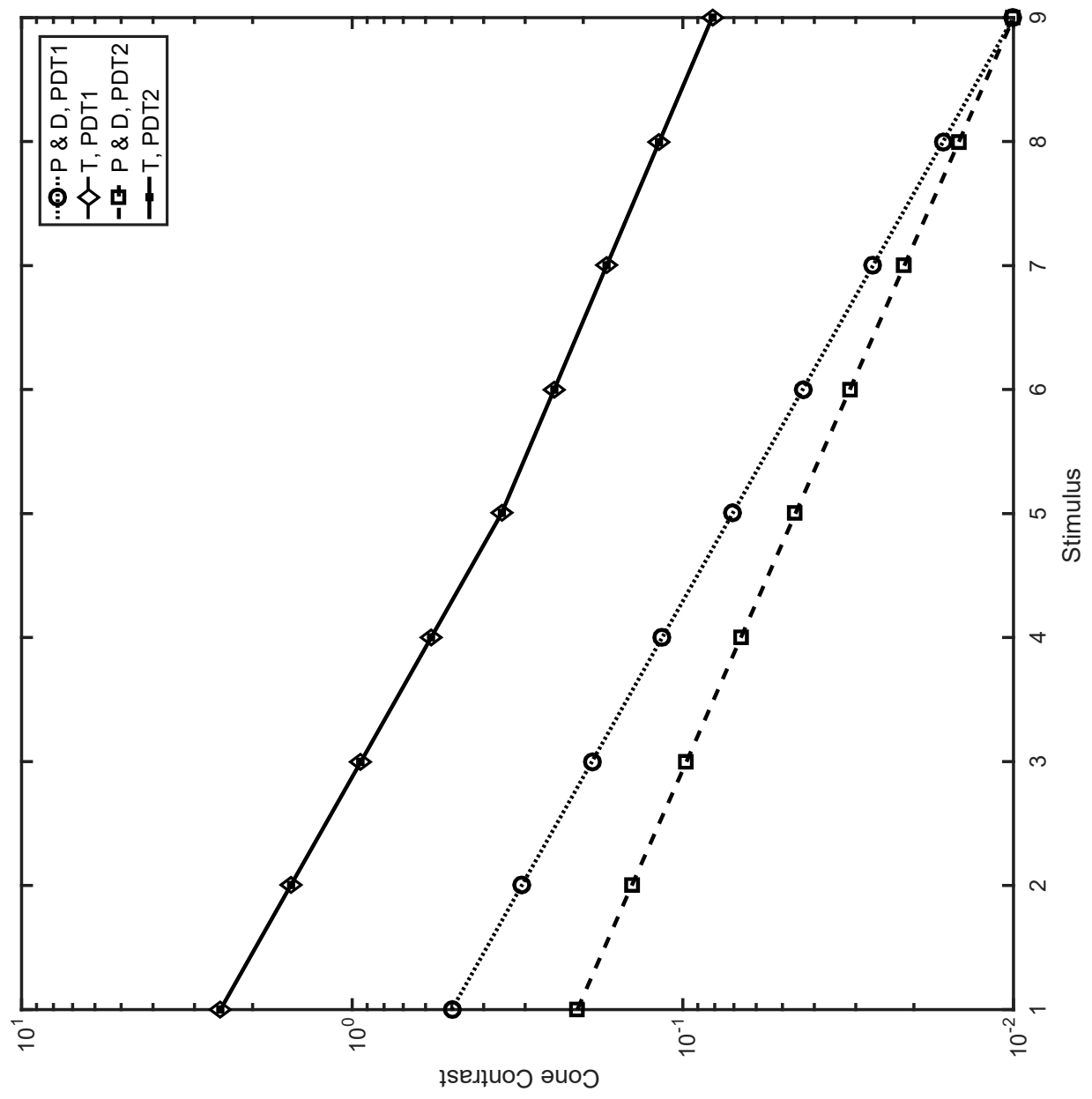


Figure 1c

Figure 2



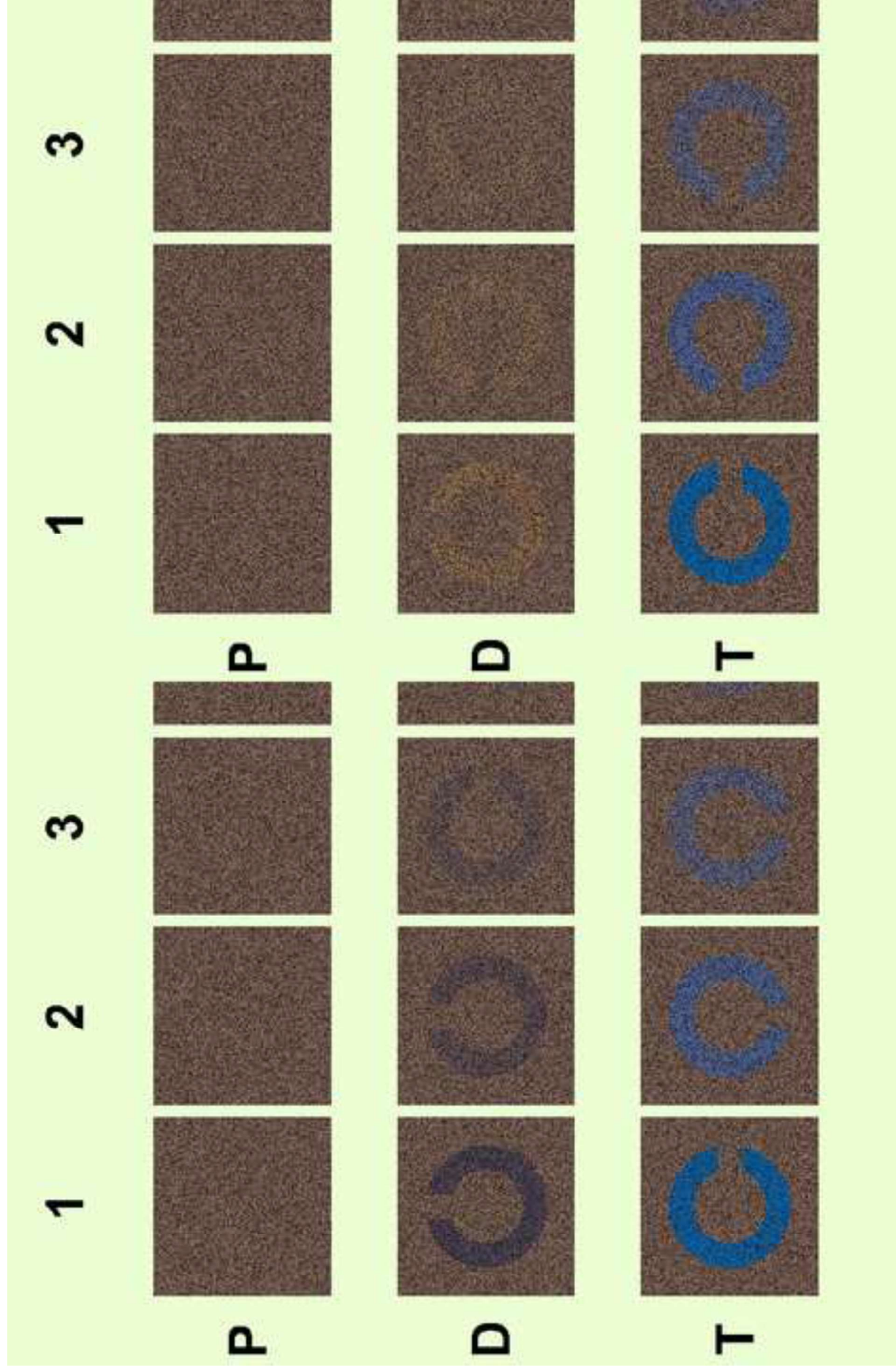


Figure 4

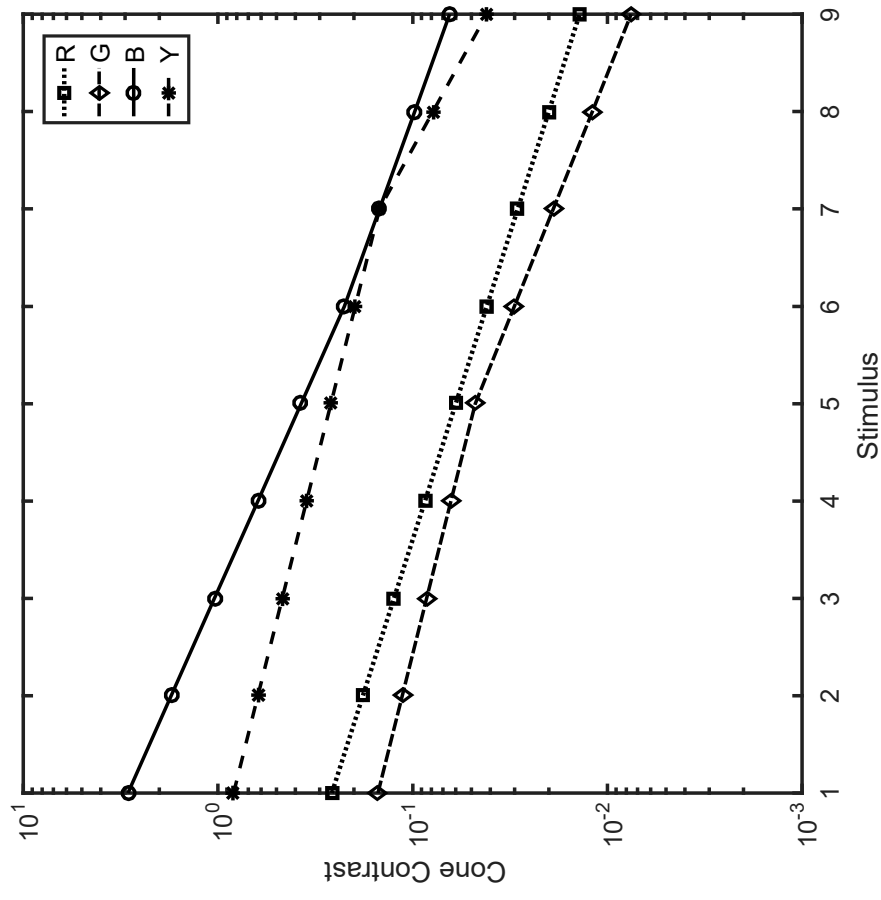


Figure 5a

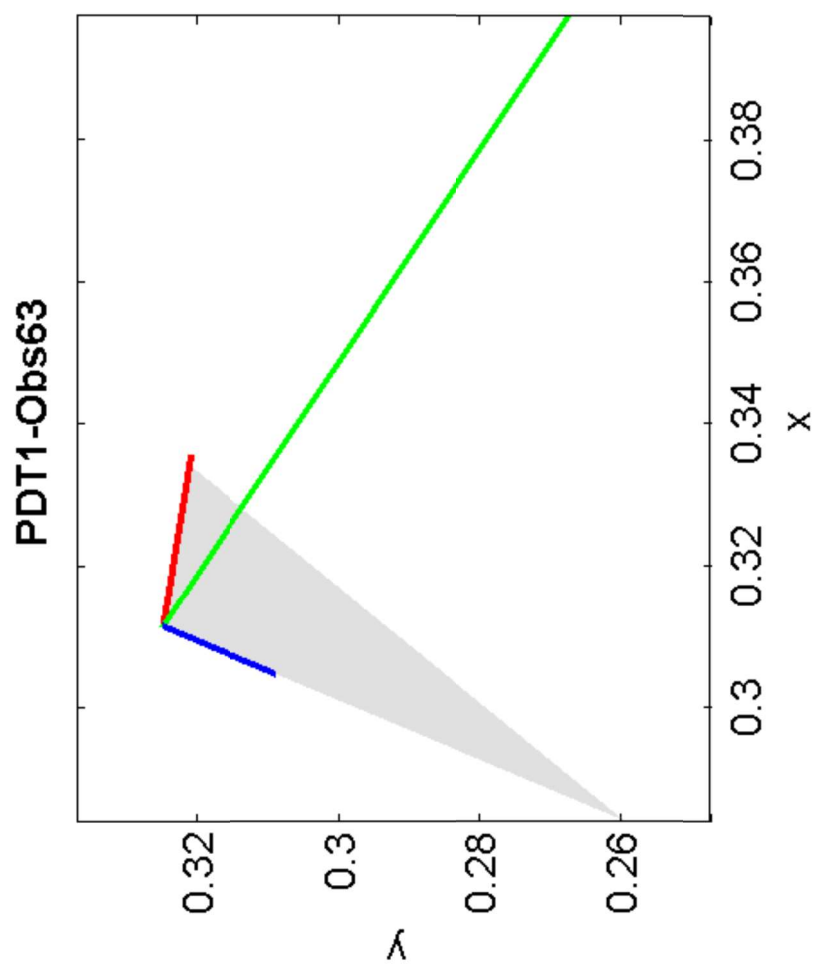


Figure 5b

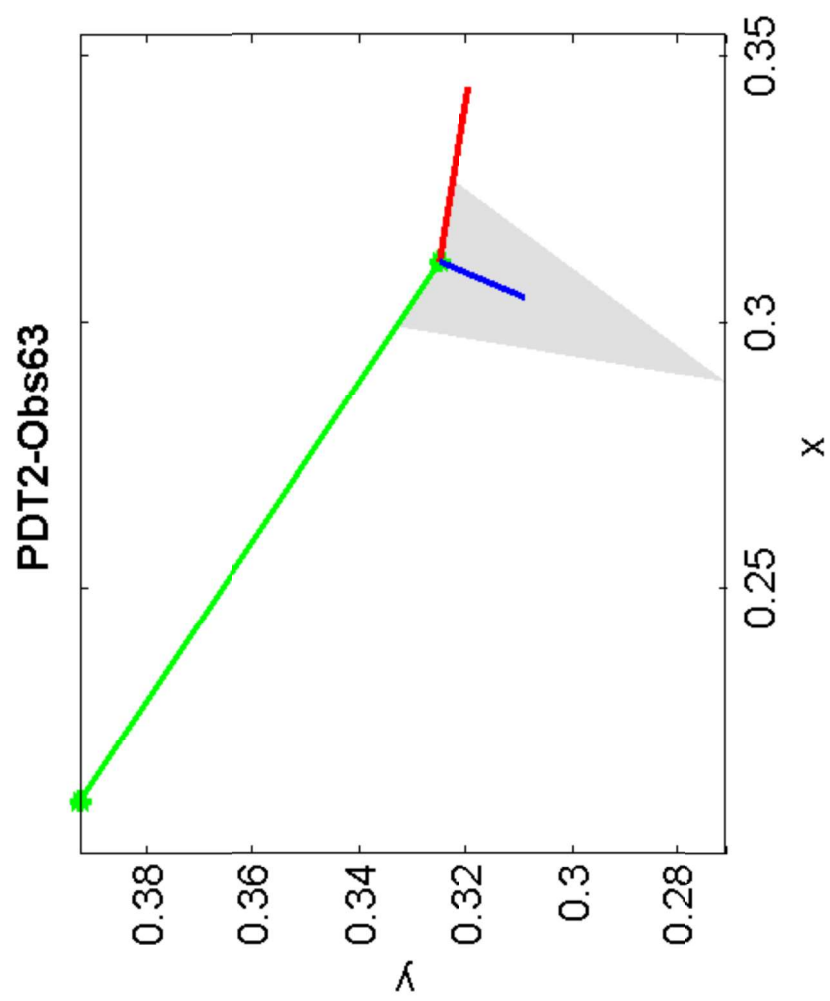
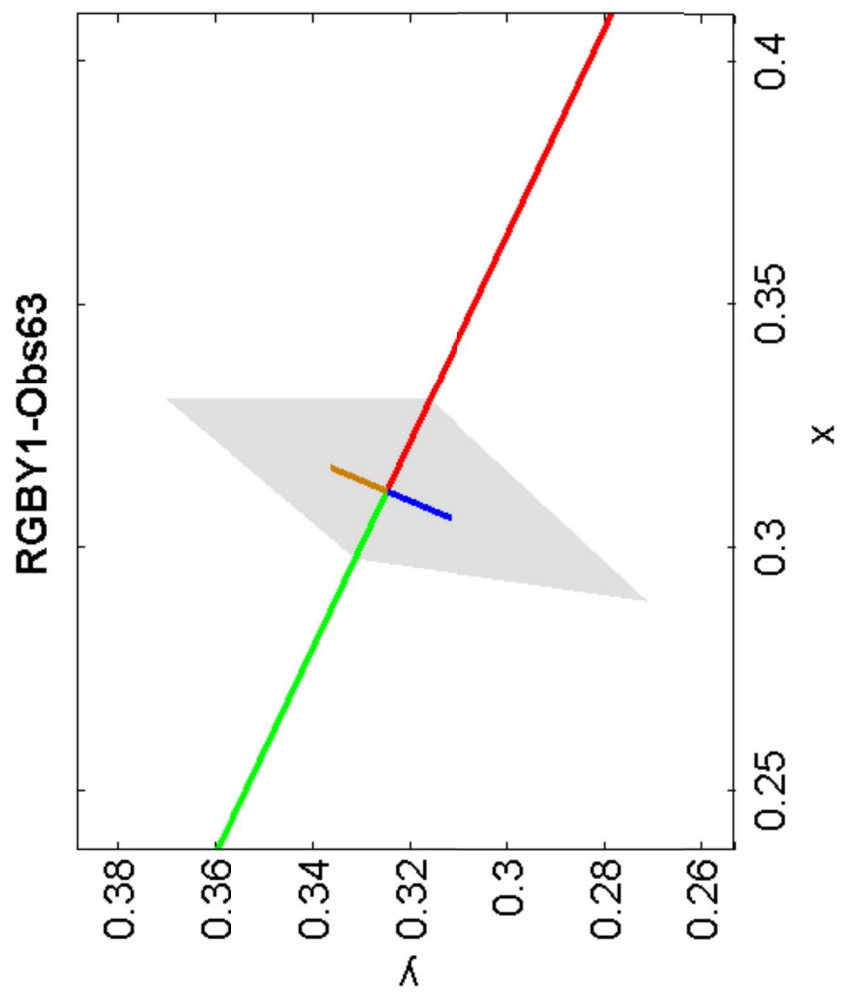


Figure 5c



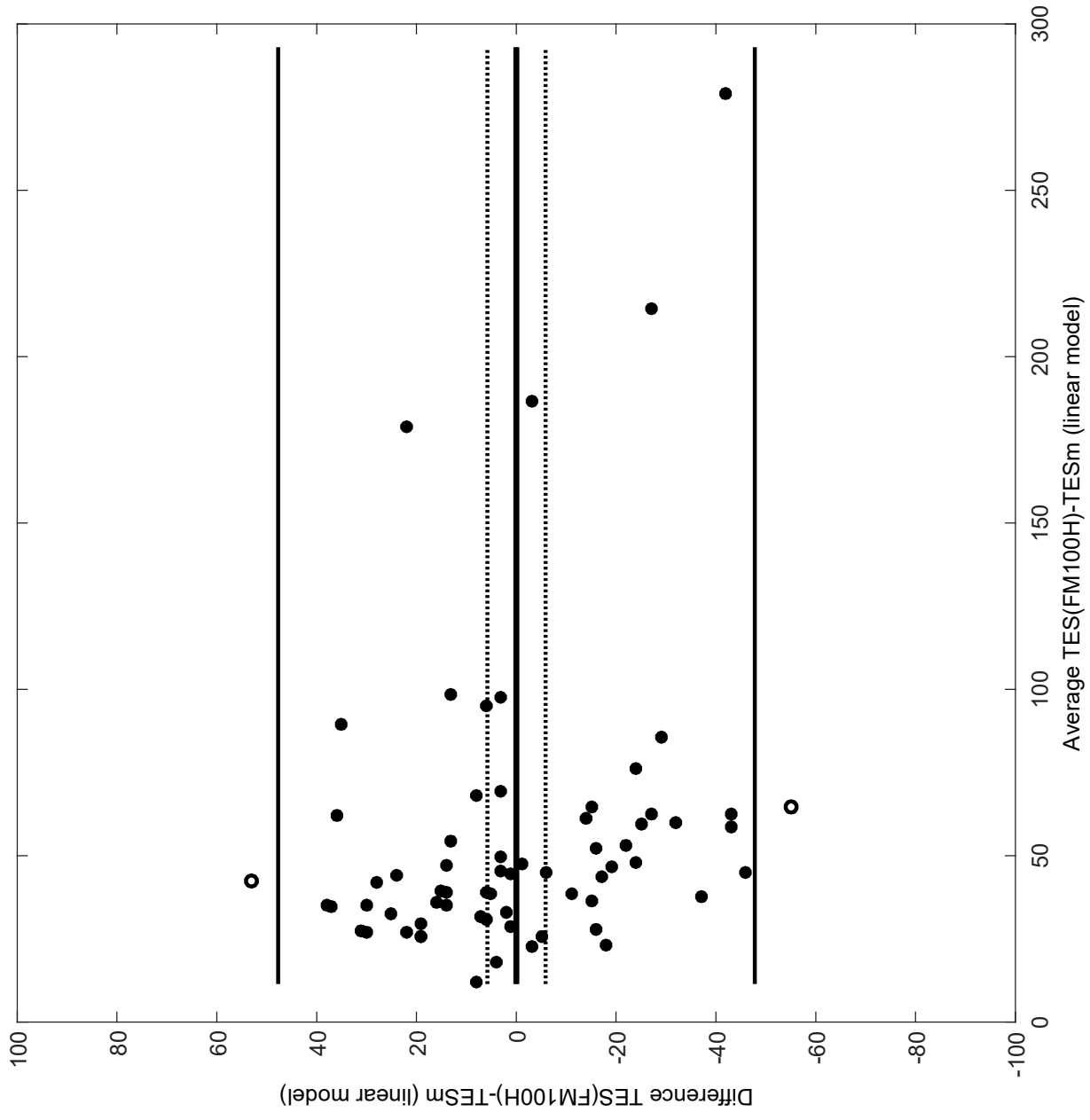


Figure 6a

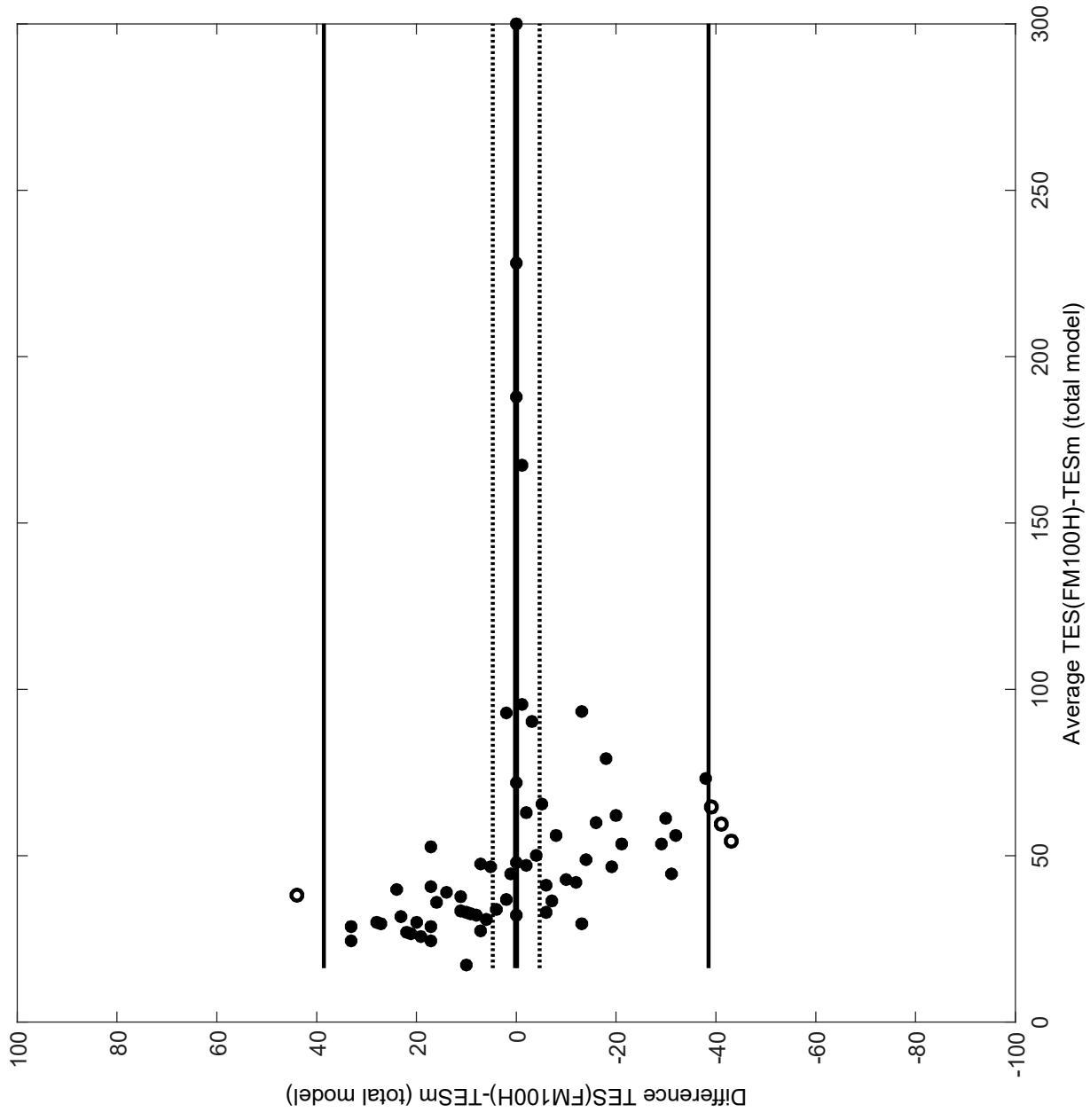


Figure 6b