Screening for congenital colour vision defects
A comparison between the Ohkuma and Ishihara plates

R Littlewood, FRACO
F Hyde, DOBA

Abstract
A prospective comparison between the Ohkuma\(^1\) and Ishihara\(^2\) pseudoisochromatic (PIC) plates was carried out in a group of 400 patients attending a general ophthalmology practice. The sensitivity of the Ohkuma test was compared to the Ishihara test, and the specificity of both was determined by reference to anomaloscopy as a gold standard.

Both tests correctly identified the same group of 24 patients as having a red/green confusion axis, and the Ohkuma test was equally as sensitive as the Ishihara. The grading plates in both tests are unreliable, but the Ohkuma test is quicker, easier to administer, gives less ambiguous responses and has a clearer cut-off score for abnormality. On the basis of this experience the Ohkuma test is recommended as more appropriate for routine colour vision screening than either the 24 or 38 plate Ishihara tests.

Key words: Colour vision defects, Ishihara plates, Ohkuma plates, pseudoisochromatic plates.

The Ishihara plates constitute one of the most widely used screening tests for colour vision defects in ophthalmic practice. Although the pigments were chosen empirically, the test has been so successful that it is recommended as a screening tool by the International Civil Aviation Organisation. The sensitivity and specificity depends on the population under test and the cut-off level chosen. Hill and Aspinall\(^3\) report the probability of correctly classifying normal colour vision is 99% in both males and females, but the probability of correctly diagnosing an abnormal observer from a failed Ishihara test is 82% in males and 19% in females if the recommended cut-off is used. A lower cut-off improves the sensitivity, while a higher cut-off improves the specificity. These data suggest that if the Ishihara test is used with a cut-off at one error, the sensitivity will be over 99% but all positive results require a second highly specific test, such as anomaloscopy, to remove the inevitable false positives.

The underlying principle of anomaloscopy is that if an observer is asked to mix a standard red with a standard green in the proportions necessary to match a test light of standard yellow, then the proportions of red and green accepted by the observer as a match for the yellow are informative about that observer’s relative sensitivity for the red and green stimuli. The relationship between this test result and the absorption spectra of the anomalous photoreceptor pigments is complex, and at this stage the relationship to the underlying molecular genetics is uncertain.\(^4\) Nevertheless, the anomaloscope remains the gold standard for colour vision testing, with 100% specificity by definition.

The variation in the red/green ratio accepted as a match is known as the matching range, which covers the full dynamic range of the instrument in a true dichromat. Normal trichromats accept very small variations in the red/green ratio around a mean value known as the mid-match point. Anomalous trichromats have a widened matching range and a shift in the mid-match point which is diagnostic of the colour defect present. Accurate results require a staircase method to determine the end point at each end of the matching range, and usually an average of three such determinations is used. The test is time consuming and requires operator

---

From The Mount Medical Centre, 32/146 Mounts Bay Road, Perth, Western Australia 6000.
Reprints: Dr R Littlewood.
training, so is not practical for routine screening. When applied only to those who first fail one plate from an edited 38-plate Ishihara test, there is a practical marriage between the convenience and sensitivity of the Ishihara test and the specificity of the anomaloscope.

The Pickford-Nicolson anomaloscope used in this study was modified to improve its reliability. A modified viewing aperture allows free viewing of a 2° test field inside a neutral grey surround of 5°. Because estimates of colour difference are unreliable in the presence of luminance differences, we use a mechanical shutter to produce variable flicker in either half of the test field while the other side is obscured. By allowing each half of the field to be matched for flicker fusion frequency, luminance differences can be minimised. This modification permits the observer with a wide matching range to make colour matching judgements without luminance clues, which is essential for the differentiation of extreme anomalous trichromats from true dichromats. The instrument has been fully calibrated spectoradiometrically, and the calculated chromaticity coordinates of the stimuli are shown in Figure 1. The dynamic range is 0 to 82 arbitrary units, with the 95th percentile range for the mid-match point being 26 to 32 units in this instrument, and the matching range less than eight units as determined in over 100 normal observers.

Lakowski\textsuperscript{6} performed a colorimetric analysis of the tenth edition of the Ishihara 38-plate series, and clarified the importance of the different plates. While some are seldom misread, others are frequently misread by normal observers. In the third decade 40% of normal observers can read the 'hidden digit' plates designed to be invisible to normal observers, while few older observers can. Lakowski\textsuperscript{6} has also shown the poor reliability of qualitative pseudoisochromatic (PIC) plates in classifying the confusion axis. The most reliable plates are the 'transformation' and 'vanishing' types (plates 2 to 21 in the 38-plate edition). Many colour vision workers use an edited version of the Ishihara test, avoiding those plates with well known weaknesses.

The Okhuma test presents a foreground circle within which there are one or two PIC breaks. Plate 1 is a demonstration plate using the same colours as plate 1 in the Ishihara test. The test plates 2 to 5 use four pairs of hues arranged to present two breaks, one more easily seen by normal observers and the other more easily seen by those with a red-green colour weakness. These are equivalent to the transformation type of Ishihara plate. Plates 6 and 7 use similar hues in a more complicated background pattern to present a single break that is not seen by those with a red-green colour weakness. These are equivalent to the vanishing type of figure in the Ishihara plates. Plate 8 introduces a new desaturated colour and is equivalent to a hidden digit style, in which only those with abnormal colour vision should see a break. Plates 9 to 14 of the Okhuma test are intended for grading only, and consist of foreground circles each with a single PIC break, arranged in an ascending order of difficulty. The first in each pair is intended to detect protans and the second to detect deutans.

Methods

We performed a prospective survey of 414 consecutive patients over four years of age attending a general ophthalmology practice, of whom 400 were able to complete the required Ishihara and Okhuma test plates. The population was unselected, but a large number of armed services personnel were referred during the study, increasing the number of young males in the series. All were tested under constant illumination consisting of daylight fluorescent lamps (Phillips TLD 90 Super de Luxe) at 200 lux, combined with an incandescent desk lamp which contributed 100 lux. This combination was thought to represent the most appropriate illumination available in the usual consulting room.
environment. The illuminant specified for these
tests is CIE source C, which is seldom used in clini-
cal practice, but an open window facing daylight
(as distinct from direct sunlight), or artificial
daylight globes are often available, and either gives
reliable results provided the level of illumination
is adequate.

Each subject underwent specific colour vision
assessment before routine orthoptic and ophthal-
mic assessment. They were asked about a history of
colour vision deficiencies, and then underwent
examination with an edited Ishihara 38 plate series,
consisting of the demonstration plate (no. 1), and the
‘vanishing’ and ‘transformation’ plates (nos 2 to
21 inclusive). Each subject was then shown an
equivalent edited series from the Ohkuma plates (1
to 7 inclusive), plus the Ohkuma ‘hidden digit’ plate
(no. 8). If an error was detected in the edited Ishi-
harah set the subject was then shown both the
Ishihara grading plates (nos 22 to 25), and Ohkuma
grading plates (nos 9 to 14) before undergoing
anomaloscope with the modified Pickford-Nicolson
anomaloscope. The mean of three measurements
at each end of the matching range was determined.
All colour vision testing was performed binocularly.

Two of 20 observers with congenital colour vision
deficiency were excluded from analysis due to inade-
quate anomaloscope data, and unfortunately both
live in remote communities and would not return for
retesting. The responses of the same group of 18
confirmed colour deficient observers were com-
pared for the Ishihara and Ohkuma edited
series. The sensitivity of the Ohkuma series was
compared to the Ishihara results, and the specific-
ity of each was determined by anomaloscope.

Results

The results obtained by those with normal colour
vision are summarised in Table 1. Of the 14
subjects unable to complete the test, most were at
the extremes of the age distribution, and were
restricted by immaturity on the one hand or disa-
bling medical conditions on the other. Three of the
elderly patients could not see well enough to
complete the test.

Twenty-four of the 400 completing the tests were
proven abnormal by anomaloscope after failing one
or more Ishihara plates. Of these 24 observers, four
(three males and one female), had an anarchic colour
vision defect associated with definite bilateral
macular disease, giving an incidence of 1% (four
in 400) binocular acquired colour vision defects.
This is less than expected in a population referred
for ophthalmic assessment, but may be explained
by the fact that 240 of the 400 in this series were
under 40 years of age.

All 20 patients with suspected congenital red-green
colour vision defects failed both the
Ohkuma and Ishihara tests, giving an incidence of
10% in males and 0.55% in females. This is close
to the expected proportion.

The 376 normal observers were classified either
on the basis of making no errors in the Ishihara test,
or else making one or more errors and then going
on to a normal anomaloscope result. As the sen-
sitivity of the single error cut-off in the Ishihara test
is over 99% the expected number of false negative
results from the 400 subjects is from none to three.
Every subject classified as normal by the Ishihara
test was also judged normal by the Ohkuma test,

### Table 1. Survey results from 400 clinic patients

<table>
<thead>
<tr>
<th>Decade of Life</th>
<th>Males</th>
<th>Females</th>
<th>Break seen in Ohkuma plate no. 8</th>
<th>Total errors by normal observers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
<td>Not normal</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>1st</td>
<td>16</td>
<td>2</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>2nd</td>
<td>28</td>
<td>3</td>
<td>29</td>
<td>51</td>
</tr>
<tr>
<td>3rd</td>
<td>25</td>
<td>3</td>
<td>18</td>
<td>38</td>
</tr>
<tr>
<td>4th</td>
<td>33</td>
<td>4</td>
<td>13</td>
<td>42</td>
</tr>
<tr>
<td>5th</td>
<td>27</td>
<td>4</td>
<td>28</td>
<td>47</td>
</tr>
<tr>
<td>6th</td>
<td>22</td>
<td>3</td>
<td>27*</td>
<td>34</td>
</tr>
<tr>
<td>7th</td>
<td>20</td>
<td>1</td>
<td>34*</td>
<td>26</td>
</tr>
<tr>
<td>8th</td>
<td>13</td>
<td>1</td>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>9th</td>
<td>8</td>
<td>1</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>Total</td>
<td>192</td>
<td>22</td>
<td>184*</td>
<td>269</td>
</tr>
</tbody>
</table>

*One female in each of sixth and seventh decades was abnormal, so total females tested was 186.
Table 2. Ishihara results from 18 abnormal observers

<table>
<thead>
<tr>
<th>Ishihara plate number</th>
<th>Normals read as</th>
<th>Deuteranomals (n = 11) and deutanope* (n = 1)</th>
<th>Protanomals (n = 5) and protanope* (n = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Ratio misread</td>
<td>Per cent</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>9/12</td>
<td>75%</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>7/12</td>
<td>58%</td>
</tr>
<tr>
<td>4</td>
<td>29</td>
<td>10/12</td>
<td>83%</td>
</tr>
<tr>
<td>5</td>
<td>57</td>
<td>11/12</td>
<td>92%</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>11/12</td>
<td>92%</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>9/12</td>
<td>75%</td>
</tr>
<tr>
<td>8</td>
<td>15</td>
<td>9/12</td>
<td>75%</td>
</tr>
<tr>
<td>9</td>
<td>74</td>
<td>12/12</td>
<td>100%</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>8/12</td>
<td>67%</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>10/12</td>
<td>83%</td>
</tr>
<tr>
<td>12</td>
<td>97</td>
<td>12/12</td>
<td>100%</td>
</tr>
<tr>
<td>13</td>
<td>45</td>
<td>9/12</td>
<td>75%</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>11/12</td>
<td>92%</td>
</tr>
<tr>
<td>15</td>
<td>7</td>
<td>11/12</td>
<td>92%</td>
</tr>
<tr>
<td>16</td>
<td>16</td>
<td>9/12</td>
<td>75%</td>
</tr>
<tr>
<td>17</td>
<td>73</td>
<td>12/12</td>
<td>100%</td>
</tr>
</tbody>
</table>

*The single protanope and deutanope each misread all plates.

which suggests that they may be ‘testing the same thing’. If they aren’t ‘testing the same thing’, then the concordance of normal results suggests the false negative rate is closer to nil than to three.

The Ishihara plates most commonly misread by observers with normal colour vision were nos 17, 9 and 13 (Table 1). Lakowski reported the highest error rate for plate no. 9. One normal observer made five errors from the edited Ishihara set, so the cut-off level required to achieve 100% specificity is five errors in this population.

No normal observer made more than one error from the edited Ohkuma test (plates 1 to 7) and the most commonly mistaken plates were nos 4 and 5. If a cut-off level of one error is used with the seven Ohkuma plates the resulting sensitivity is 100% of those detected by the Ishihara set, and the specificity is 100% in this series. Plate no. 8 in the Ohkuma set is a ‘hidden digit’ type, but the break in the foreground circle was seen by two-thirds of the normal observers, to whom it should have remained hidden. Furthermore these false-positive responses were most frequent in normal observers under 60 years of age (233 of 296; 78%) and less in those over 60 years (38 of 104; 36%). The inclusion of plate 8 in the Ohkuma test contributes no information in subjects under 60 years of age.

Tables 2 and 3 summarise the results obtained by the 18 colour-deficient observers sorted into groups according to anomaloscope results. Neither set of PIC plates was able to predict the anomaloscopic grading accurately from the use of the grading plates (P > 0.1), or by using the total error score.

The proportion of males with a congenital colour vision defect in this survey was 10%, which is close to the predicted number of 8%. If the unconfirmed female subject was a true positive then the proportion of affected females would have been 0.5%, which is also close to the expected number.

Both tests performed well as screening tools with 100% specificity with reference to anomaloscope when the appropriate cut-off was applied (one error for the Ohkuma test; five for the Ishihara).

Table 3. Ohkuma results from 18 abnormal observers

<table>
<thead>
<tr>
<th>Ohkuma plate number</th>
<th>Deuteranomals (n = 11) and deutanope (n = 1)</th>
<th>Protanomals (n = 5) and protanope (n = 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ratio misread</td>
<td>Per cent</td>
</tr>
<tr>
<td>2</td>
<td>12/12</td>
<td>100%</td>
</tr>
<tr>
<td>3</td>
<td>11/12</td>
<td>92%</td>
</tr>
<tr>
<td>4</td>
<td>10/12</td>
<td>83%</td>
</tr>
<tr>
<td>5</td>
<td>10/12</td>
<td>83%</td>
</tr>
<tr>
<td>6</td>
<td>9/12</td>
<td>75%</td>
</tr>
<tr>
<td>7</td>
<td>9/12</td>
<td>75%</td>
</tr>
<tr>
<td>8</td>
<td>12/12</td>
<td>100%</td>
</tr>
<tr>
<td>9</td>
<td>0/12</td>
<td>0%</td>
</tr>
<tr>
<td>10</td>
<td>1/12</td>
<td>8%</td>
</tr>
<tr>
<td>11</td>
<td>1/12</td>
<td>8%</td>
</tr>
<tr>
<td>12</td>
<td>3/12</td>
<td>25%</td>
</tr>
<tr>
<td>13</td>
<td>5/12</td>
<td>42%</td>
</tr>
<tr>
<td>14</td>
<td>9/12</td>
<td>75%</td>
</tr>
</tbody>
</table>

Discussion
We found the results from the two tests similar, but there are four advantages to the Ohkuma test for routine screening. It is quicker and yields fewer ambiguous responses than the Ishihara. For example, if 73 is read as 23 or 78 it can be difficult to score the plate, especially if the decision is critical to the test result. It can be applied to those who are unable to speak but are able to point. It cannot be memorised, since each plate can be rotated at will by the examiner. In this study the cut-off point for the Ohkuma was determined more readily than the cut-off for the Ishihara test. We find the ease of use and greater confidence in the test score are significant advantages.
Since this study we no longer use the Ohkuma grading plates (9 to 14) in sequence, as recommended by the test authors. Instead we offer a 'forced choice' test, by reversing alternate grading plates so that they are viewed in pairs, (9 + 10, 11 + 12, and 13 + 14). The subject is asked to indicate which of each pair has the most obvious break. By indicating the 'protan' or 'deutan' plate a score out of three is obtained. This result appears to correlate better with the confusion axis determined by anomaloscopy, but unfortunately no significant results are available for this modified method as it was implemented after the study.

Most ophthalmologists use fluorescent or incandescent lighting with the PIC plates of their choice for routine colour vision screening. We used a combination of both, but selected fluorescent tubes with good colour rendering properties (Ra 95, colour temperature 5000 °K). Those who fail should be retested under natural daylight if source C is unavailable. The prohibitive cost and recent lack of availability of the Macbeth easel lamp have forced some colour workers to devise alternatives. A Schott BG 34 filter combined with a halogen source is one such currently available alternative. Even with a high error score the PIC tests do no more than answer the question, 'Is there a red-green confusion axis?' Neither the Ishihara, Ohkuma, Hardy-Rand Rittler, or Farnsworth F-2 are reliable as grading tools for red-green deficiencies. There is no substitute for anomaloscopy when a grading is required.

As our understanding of the molecular genetics of congenital red-green colour deficiencies improves, there is a tantalising possibility that more deterministic psychophysical tests may allow correlations to be made with the underlying genetic defects. Rapid advances in the areas of colour science of relevance to ophthalmology may well place greater demands on our ability to screen for colour vision deficiencies accurately and efficiently. We believe the Ohkuma PIC test used under adequate illumination is a reliable screening tool, but like all PIC tests it does not grade defects accurately. A modification to the recommended grading method may be advantageous.

References