Visual Neuroscience (2004), **21**, 431–436. Printed in the USA.

DOI: 10.1017/S0952523804043469

Evaluation of an updated HRR color vision test

JAMES E. BAILEY,[1] MAUREEN NEITZ,[2] DIANE M. TAIT,[2] AND JAY NEITZ[2]

[1]Southern California College of Optometry
[2]Medical College of Wisconsin

(RECEIVED November 17, 2003; ACCEPTED May 8, 2004)

Abstract

The HRR pseudoisochromatic plate (pip) test was originally designed as a screening and diagnostic test for color vision deficiencies. The original HRR test is now long out of print. We evaluate here the new 4th edition of the HRR test, produced in 2002 by Richmond Products. The 2002 edition was compared to the original 1955 edition for a group of subjects with normal color vision and a group who had been previously diagnosed as having color vision deficiencies. The color deficient subjects spanned the range of severity among people with red-green deficiencies except for one individual who had a mild congenital tritan deficiency. The new test compared favorably with the original and in at least two areas, outperformed it. Among subjects with deutan defects the classification of severity correlated better with the anomaloscope results than the original; all the subjects who were classified as dichromats on the anomaloscope were rated as "severe" on the new HRR, while those diagnosed as anomalous trichromats were rated as mild or medium on the new test. Among those with moderate and severe defects the new test was highly accurate in correctly categorizing subjects as protan or deutan. In addition, a mild tritan subject made a tritan error on the new test whereas he was misdiagnosed as normal on the original.

Keywords: Pseudoisochromatic test, Color blindness, Color vision, Saturation discrimination, Dichromatic confusion lines

Introduction

Sixty years ago LeGrand Hardy, Gertrude Rand, and M. Catherine Rittler (Fig. 1) commenced development of a color vision test that later was made commercially available under the name "HRR Pseudoisochromatic Plates" (Hardy et al., 1954). It was designed as a single test for red-green (protan and deutan) and blue-yellow (tritan and tetartan) types of deficient color vision that combined both screening and diagnostic capabilities. American Optical (AO) produced the first edition of the test in 1955, followed in 1957 by a second edition with the order of the first-edition plates slightly altered. Many have considered the AO HRR to be a good test (Walls, 1959; Paulson, 1971; Vos et al., 1972), particularly for screening of red-green color defects and, to some degree, further classification of these according to type (qualitative diagnosis; protan, deutan, tritan, and tetartan) and extent of defect (quantitative diagnosis; "mild," "medium," "strong"). The AO HRR has been out of print and unavailable since about 1970. In 1991 Richmond International (RI) produced a third edition that aimed to reproduce the 1957 AO edition, but it has been unavailable since early 2002.

Walls (1959) and Schmidt (1952) conducted two of the now classic evaluations of the AO HRR test compared to the Nagel anomaloscope. Both found the sensitivity of the test as a screener for red-green color defects to be very high (0.96 and 0.98, respectively). Accuracy in dichotomizing protan and deutan types of color defect was slightly less satisfactory (0.84, 0.81), and only modest (0.63, 0.50) for classifying as "strong" the extent of the most severe dichromatic color defects: protanopia and deuteranopia. Six of 12 dichromats studied by Schmidt (1952), for example, were classified as having a "medium" extent of defect. The six who were misdiagnosed were all deuteranopes according to the anomaloscope. In agreement with the older studies, Birch (1997) found that the HRR did not distinguish dichromats and anomalous trichromats; 54% of dichromats were classified as moderate rather than severe. Birch (1997) also provides an excellent review of the evaluations of the AO HRR in diagnosing red-green deficiencies that have been conducted over the last half century. Little is known about how well the AO HRR identifies persons with blue-yellow color defects. Paulson (1971), however, reported that 4 tritan color defectives passed all of the blue-yellow test plates of the AO HRR. To our knowledge there is not a single published report of results of testing the theoretically possible tetartan-type color defect. Recent spectrophotometric measurements by Neitz and Neitz (2001) of the protan, deutan, and tritan colors in the AO and RI versions of the HRR test showed varying degrees of imprecision in achieving congruence of the printed colors with theoretical ideal confusion colors—the dichromatic confusion lines—accepted during the time of the development of the original test. Discrepancies between the modern accepted dichromatic confusions and the HRR colors can account for some of the imperfections in the test noted above which could be remedied by printing a new test using state-of-the-art methods.

Address correspondence and reprint requests to: James E. Bailey, Southern California College of Optometry, 2575 Yorba Linda Blvd., Fullerton, CA 92831, E-mail: cvddoc@aol.com

Fig. 1. The developers of the HRR pseudoisochromatic plate test, LeGrand Hardy (left), Gertrude Rand (middle), and M. Catherine Rittler (right).

Given the attributes and limitations of the HRR test it would seem desirable not only to reproduce those parts of the original AO test that worked well (screening plates) but also to rebuild what needs improvement (a subset of the diagnostic plates), taking into account new knowledge about color vision that has emerged over the last half-century. For example, much experimental data supports a change in the deuteranopic co-punctal (convergence) point from longstanding values of $x = 1.08$, $y = -0.08$, generally accepted at the time of the development of the AO HRR, to $x = 1.40$, $y = -0.40$ (Smith & Pokorny, 1975). Earlier spectrophotometric analysis of the AO HRR (Neitz & Neitz, 2001) revealed that colors of the deutan plates corresponded closely to the old deuteranope confusion point ($x = 1.08$, $y = -0.08$); however, they deviate from the line derived using the now accepted deuteranope convergence point ($x = 1.40$, $y = -0.40$). Theoretically, protan and deutan color defects would be better dichotomized (qualitative diagnosis) if the deutan colors were adjusted. In addition, by adjusting the colors on the new deutan confusion line the problem of failure to classify deuteranopes as "strong" (quantitative diagnosis) would also be remedied. On the other hand, the screening plates of the AO HRR test should be duplicated in order to preserve the excellent qualities of that part of the original test. The strategy used to build the HRR 2002, involved carefully selecting a different ink for each hue in the test to ensure these objectives were obtained. Here we report on our work to evaluate the Richmond Products (RP) 4th edition of the original HRR test, including spectrophotometric measurements of the colors of the new test and results of administering the 2002 RP HRR test to a group of 23 observers with red-green color defects, one observer with a blue-yellow color defect and 29 subjects with normal color vision.

Materials and methods

Color vision classification

Color matching was performed using a Nagel type I anomaloscope following the testing procedure suggested by Linksz (1964) including the use of the preadaptation screen when applicable. Red-green color deficient subjects were classified according to their performance on this test. Those subjects who accepted the entire range of mixtures of the primary test lights as matching the standard comparison light were diagnosed as dichromats. The intensity of the 589 nm comparison light required to match the brightness of the mixture light was used to classify dichromats as protanopes or deuteranopes. Protanopes set the 589 nm comparison light to intensity near zero (usually < 5) when matching the red primary and they set it to a value brighter than normal when matching the green primary. Deuteranopes set the brightness of the 589 comparison to a value near the normal brightness match for all red-green mixtures. Color deficient individuals who accepted less than the full range of mixtures of the primaries were diagnosed as anomalous trichromats. Those who required a higher proportion of the red primary than normal to complete the match were diagnosed as protanomalous trichromats and those who required a relatively higher proportion of the green primary compared to normal were diagnosed as deuteranomalous trichromats.

Subjects

Twenty-three color deficient Caucasian males ranging in age from 20 to 56 with a mean age of 29 were tested. One participant had been classified earlier as having a mild congenital tritan color vision deficiency due to a mutation that produced an amino acid change in the encoded S photopigment which would be expected to affect S cone function (Gunther et al., 2003).

Twenty-nine individuals (18 female 11 male) with normal color vision were also tested. They all passed Ishihara's pseudoisochromatic plates, The Neitz Test of Color Vision, The D15, Lanthony's desaturated 15 hue test, and they made a normal Rayleigh color match on the Nagel anomaloscope. Experiments involving human subjects were conducted in accordance with the principles embodied in the declaration of Helsinki, and were approved by Institutional Review Boards.

Spectrophotometric measurements

The spectral reflectances from a representative sample of colored dots from the Richmond HRR were measured spectrophotometrically using a Colortron II 32 band color sensor manufactured by X-Rite Incorporated (Grandville, MI). The spectrophotometric measurements of the colored dots were converted to chromaticity coordinates that corresponded to the plates being viewed under illuminant C. The calibration of the Colortron II spectrophotometer was validated earlier by measuring monochromatic lights (Neitz & Neitz, 2001).

HRR administration

The HRR tests were administered according to guidelines that accompanied the tests. All subjects had taken the AO HRR previously as part of a large battery of color vision tests. After the previous testing session subjects were told an overall diagnosis that was deduced from performance on all the tests, but they were not given information about performance on individual tests. Thus, although the subjects had previous experience in taking the test they had not received feedback about the correct responses on individual plates. The experimenters were aware of the previous diagnosis as to presence or absence of a color vision defect, but to avoid experimenter bias the tester followed a set script for all subjects and the only other interaction between subject and experimenter during the test was to record the subject's responses. As indicated in the RP HRR 2002 instructions, testing was done using a Richmond Products "True Daylight Illuminator with Easel" (color temperature 6500, model number 1339R) which approximated illuminant C. The level of illumination at the surface of the test plates averaged about 550 Lux. Subjects were given both the AO HRR and the RP HRR 2002 in a single sitting and the order of administration was randomized from subject to subject. In practice it is not uncommon for users to administer the test under standard fluorescent lighting. Thus, for comparison, in a separate sitting, both tests were administered under conventional fluorescent lighting (Sylvania Cool White, color temperature 4800); for this administration of the test, the illuminance of the test plates was approximately the same, 550 Lux, as when the daylight illuminator was used.

Results and discussion

Spectrophotometric measurements

As shown in Fig. 2, the chromaticities as viewed under illuminant C measured for the RP HRR 2002 compare favorably with the ideal protan, deutan and tritan dichromatic confusions. Thus, the test closely meets one of the theoretical design objectives for the reengineered HRR test.

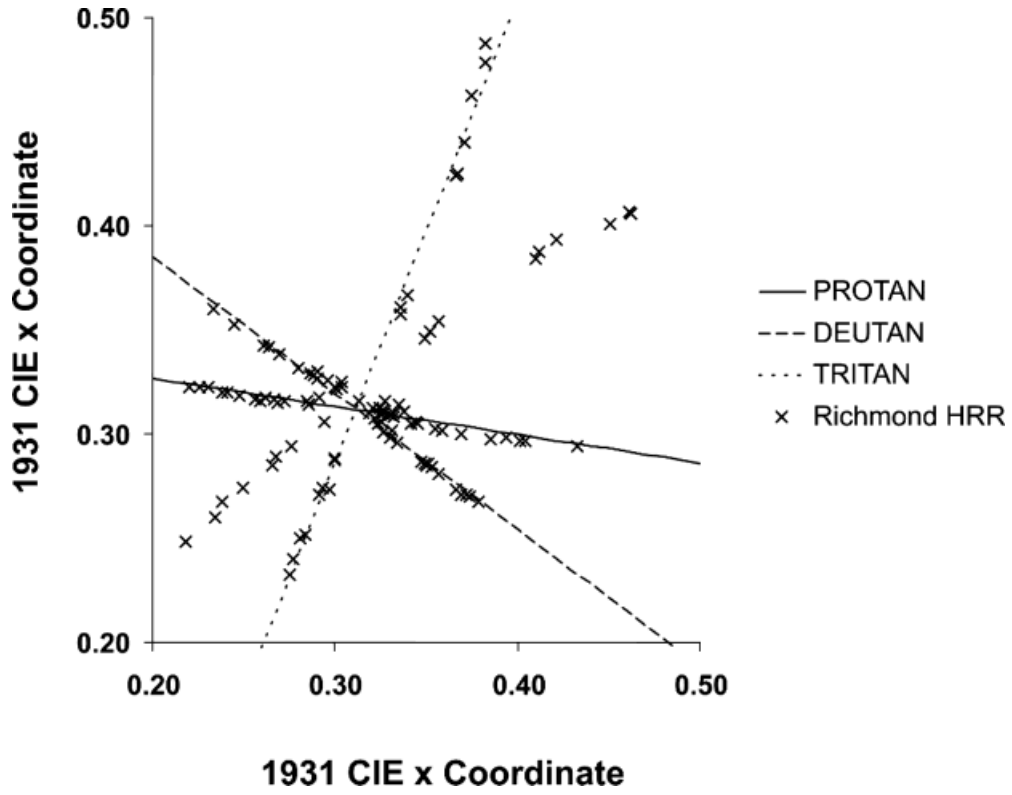

Fig. 2. Chromaticities of the gray and colored dots from the HRR 4th edition, produced by Richmond Products in 2002 under illuminant C compared with theoretical ideal protanopic, deuteranopic, and tritanopic isochromatic (confusion) lines plotted in 1931 CIE *x*, *y* coordinates. The earlier editions also included a series of plates to test for a fourth category "tetartanopia." Although this classification is not recognized by most modern vision scientists the plates are included in the HRR 2002 (x's with no isochromatic line) keeping the new test consistent for those familiar with using the earlier tests.

Subjects with normal color vision

All (100%) of subjects with normal color vision tested as normal on the HRR 2002. The instructions for the test indicate that if an error is made on one of the screening plates but no errors are made on the diagnostic plates the screening series should be re-administered with the book in a different orientation. One subject, incidentally a carrier of a deutan color vision defect, made an error on plate 7, the most difficult red-green screening plate. She read the plate correctly on a second administration, as recommended in the testing instructions. All the other subjects with normal color vision made zero errors when tested in a single administration. This indicates that the improved sensitivity (see below) is not at the expense of an increased rate of false positives.

Subjects with color vision defects

All (100%) of subjects who were diagnosed as having a color vision deficiency by the anomaloscope (and the battery of other tests as well) were diagnosed as having a deficiency using the new RP HRR 2002. The sample included individuals that span the range of severity of protan and deutan color vision deficiencies and one participant who had been identified previously as having a mild congenital tritan color vision deficiency (Gunther et al., 2003).

Mild tritan deficiency

From an earlier investigation (Gunther et al., 2003) a subject was found to be heterozygous for a missense mutation that substituted proline for a highly conserved leucine at amino acid position 56 in the S cone photopigment. This mutation was absent in 584 S cone photopigment genes from 262 subjects who did not make tritan errors. He made errors with a tritan axis on Lanthony's desaturated 15 hue test and he was originally identified because he made a tritan error but no red-green errors on the Neitz test of Color Vision. On clinical examination the subject's fundus and acuity were normal. This subject did not make any errors on the original (1955) AO-HRR. He did make one error on the first tritan screening plate of the RP HRR 2002. This suggests that the new test may be more sensitive in detecting mild tritan deficiencies than the original. The subject also made an error on plate 7, a red-green plate, on the first administration of the test, however, upon the second administration the subject did not make any red-green errors, but consistently made the tritan error.

From a practical perspective we note that this subject made the tritan error when tested on the RP HRR 2002 both when using the daylight illuminator as well as when the test was administered under cool white fluorescent lights. He made no errors on the original (1955) AO-HRR under either illuminant. We thought that any differences in test performance under the two illumination conditions might be greater for tritan deficiencies because the largest differences between cool white and daylight florescent light are in the shorter wavelengths. Conclusions about sensitivity of the test in detecting tritan defects when used under different illumination conditions cannot be made from this case. However, because, in practice, many people either do, or would like to, test color vision without special lighting, this matter should be investigated further in a future study. Overall we did not find any significant difference in test results when subjects were tested under standard cool white florescent illumination compared to the recommended daylight illuminator. There were a few scattered differences in the number of errors red-green color vision deficient subjects made between the illumination conditions but in no case was the categorization of presence or absence of deficiency, or type, or degree different when the test was administered under cool white compared to daylight.

Red-green deficiency

In this evaluation, all subjects who were diagnosed as having red-green color vision deficiencies on the anomaloscope also made red-green errors on the HRR 2002. Thus, for this group the new test was 100% accurate in detecting color vision deficiencies diagnosed by anomaloscope testing. In addition to being able to detect them the HRR was designed to classify red-green color vision deficiencies according to type (qualitative diagnosis; protan, deutan, tritan, and tetartan) and the extent of defect (quantitative diagnosis; "mild," "medium," "strong"). To evaluate these aspects of the test, in addition to the comparison with the anomaloscope, each subject's performance on the RP HRR 2002 and the 1955 AO-HRR were compared. The results are given in Fig. 3 which illustrates the plate by plate performance of all subjects on both tests. The red-green plates decrease in difficulty progressively from the front to the back of the book. At every difficulty level, including the screening plates, there is one symbol designed to have colors which lie on a protan confusion line and one symbol designed to be on a deutan confusion line. The colored dots in the most difficult plates are very desaturated which makes the chromaticity differences between protan and deutan plates small. Thus, the difficult plates which allow the detection of the mild color vision deficiencies are less able to differentiate protan and deutan defects than the more saturated plates which are missed by people with more severe deficiencies. To evaluate each of the two HRR tests in their ability to provide a correct qualitative diagnosis (protan vs. deutan) and display the results in Fig. 3, each symbol was assigned a qualitative diagnosis value. This value was progressively larger for the plates as they decreased in difficulty, as follows, starting with the first red-green plate (i.e., plate # 7 HRR 2002, plate 1AO HRR): 0.13, 0.26, 0.38, 0.38, 0.51, 0.64, 0.64, 0.77, 0.90, 1.02, 1.15, 1.28, 1.4, 1.53. The test is constructed so that in all but two cases there is a deutan and a protan symbol on each plate. For the two exceptions there is only one symbol per page and the deutan and protan symbols are on consecutive pages. In the two cases where a qualitative diagnosis value is duplicated above (i.e., 0.38 and 0.64) it is because the deutan and protan symbols appear on consecutive plates. For each deutan error a subject made, the subject's error total was increased by the corresponding plate's qualitative diagnosis value. For each protan error a subject made, his qualitative diagnosis score, displayed on the *Y* axis of Fig. 3, was decreased by the corresponding plate's qualitative value. The plot shows the number that corresponds to a subject's cumulative score. These numbers were designed so that a person making deutan errors on all plates and and no protan errors would receive a cumulative score of $+10$, and a person making protan errors on all plates and no deutan errors would receive a cumulative score of -10. Plates 5 and 6 are not used in the 1955 HRR graph because they test tritan color vision. The screening plates start with number 5 in the Richmond test and the two tritan plates are 5 and 6, thus there is no data shown for those two plates in Fig. 3. The results for each subject are shown as a solid line which connects data points which are triangles for subjects who tested as protan on the anomaloscope and circles for subjects who tested as deutan on the anomaloscope. Each subject's line ends at

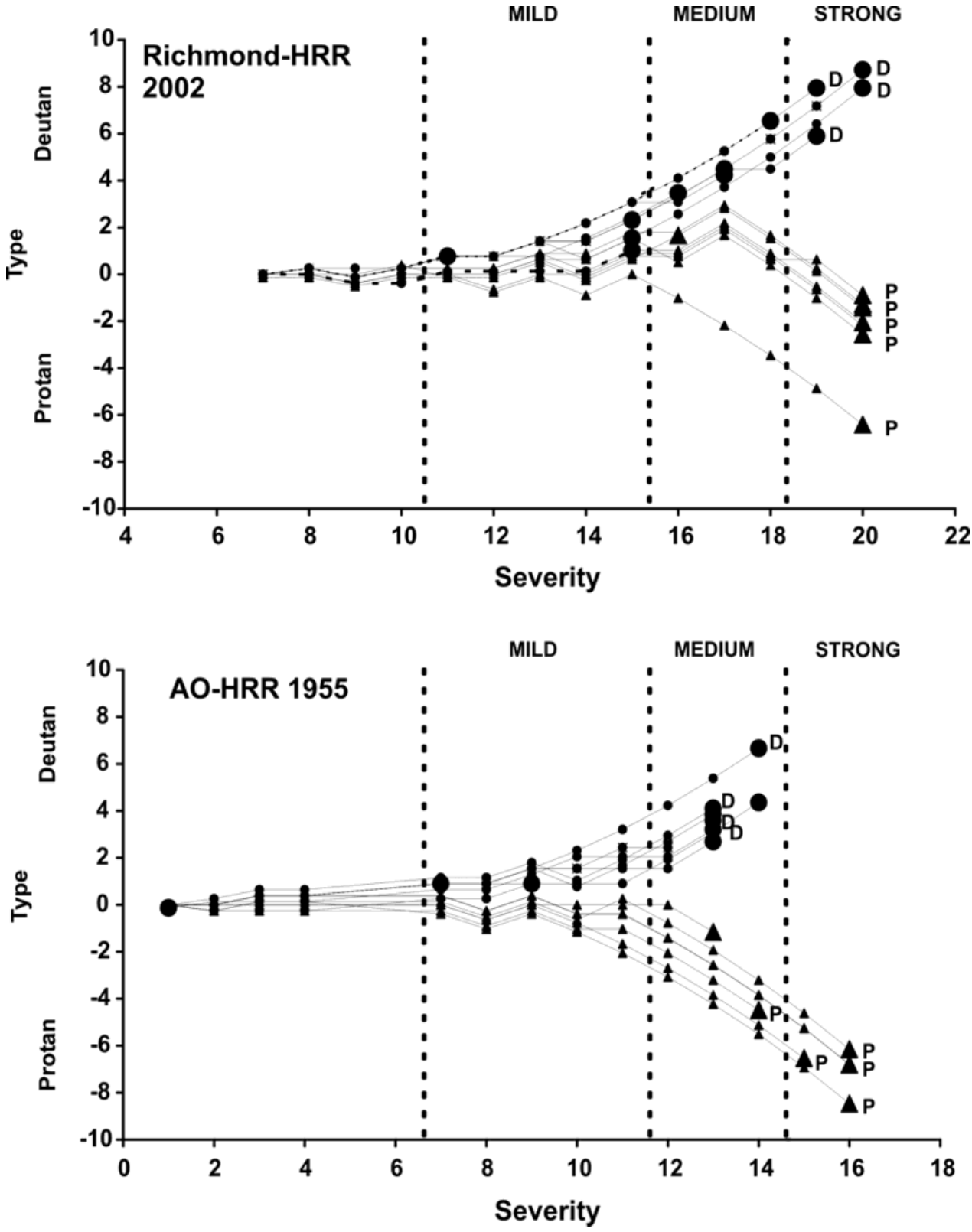

Fig. 3. Results for 23 subjects who were diagnosed as having red-green color vision deficiencies using the Nagel, type I anomaloscope (triangles, protan; circles, deutan; P, protanopes; D, deuteranopes). Graphs illustrate the plate-by-plate performance of all subjects on both tests; Richmond HRR 4th edition (top), AO-HRR (bottom). The results for each subject are shown as a solid line which connects the data points. Each subject's line ends at the last plate in which he made an error, indicated by a larger symbol. The position of the larger symbol provides a diagnosis by the test with regard to both type and severity. See text for further details. We note that one protanope (and no other red-green deficient subject) made a single error on one of the tritan screening plates, the results for which are not shown on this graph.

the last plate in which he made an error and the end of the each person's line is punctuated by a larger symbol. The position symbol on the graph marking the last plate missed provides a diagnosis by the test with regard to both type and severity. The line leading to the last point gives an indication of how each subject performed on each plate on the test up to his last error. Subjects who end up with low Y values are categorized as protan by the test being evaluated. Those with high values are categorized as deutan by the test. The *X*-axis indicates the severity of each subject as determined by each test. The data points that have the letter D next to them indicate subjects who were diagnosed as deuteranopes according to the anomaloscope. The data points that have the letter P next to them indicate subjects who were diagnosed as protanopes according to the anomaloscope. All the other subjects were diagnosed as anomalous trichromats by the anomaloscope.

In our experience, the priorities for what people want to know from a color vision test are the presence or absence of a color deficiency and its severity. An inadequacy of the original HRR, which is evident from Fig. 3 and from earlier evaluations (see Birch, 1997 for a review), is that people who test as deuteranopes on the anaomaloscope are scored on the AO HRR as having only a medium defect. In the new RP HRR 2002, this was corrected by improving the colors of the deutan plates. This had the desired effect in that all deuteranopes tested as strong, which was a very high priority in making this a good test. However, the results here suggest that with this change there was also an increase in the number of times that people with protan deficiencies missed deutan plates. On both the old and new tests there were instances for the middle plates where protan subjects missed a deutan symbol on a particular plate but read the protan symbol correctly; however, this occurred more frequently with the new test. The plate for which this was a problem is #17; all protanopes, but one, missed the deutan symbol but read the protan symbol correctly. This affected the cumulative scores in Fig. 3 such that only on the last plate did all protanopes achieve a "protan" score and the cumulative scores were "deutan" for all but one protanope until the last two plates. This would have been remedied if plate #17 had been eliminated from the cumulative scoring. These results suggest that, in practice, plate 17 should not be used in separating protan from deutan subjects. However, cumulative scores on plates 18, 19, and 20 of the new test segregated the protan and deutan subjects perfectly for all who made errors on them. Finally, one subject diagnosed as protanomalous on the anomaloscope had a final score slightly on the deutan side. Inherent to the design of the test, the protan and deutan symbols are similar in chromaticity for the plates which are intended to be missed by individuals with less severe deficiencies. Thus, it is expected that the greatest ambiguity in qualitative diagnosis will be for the subjects with milder deficiencies. Both the old and new tests do a good job of correctly dichotomizing subjects as protan and deutan. However, the results presented here suggest that the old test may be slightly better at dichotomizing mild subjects. In designing a test like this, there are trade offs. The deutan series of symbols were modified in the new test to improve the detection and quantitative diagnosis of deutan disorders. These high priority improvements may have come with the price of some loss in ability to separate mild protans from deutans. In practice, this should pose very little problem because protanomalous individuals with a defect this mild are rare in the population ($<$1% of the total) and the test presumably would correctly identify them as having a defect and correctly identify them as having an anomalous trichromacy vs. dichromacy even though it might give an ambiguous or misleading result as to whether the defect is protan vs. deutan.

For the new test the severity corresponds closely with that determined by the anomaloscope–all the subjects determined to be deuteranopes on the anomaloscope were found to be strong with the HRR 2002. None of the deuteranomalous were categorized as strong with the old HRR. Finally, one person who was determined to be deuteranomalous on the anomaloscope made no errors on the old test but missed one screening plate on the new test and another deuteranomalous individual made only one error on the old test but made errors in the mild portion of the diagnostic series on the new test. This suggests that the new test is also more sensitive at detecting color vision defects than the old one.

Acknowledgments

This work was supported by NIH grants EY09303, EY09620, and EY01931, and Research to Prevent Blindness. MN is the recipient of an RPB Senior Scientific Investigator Award. JN and MN are the recipients of an Alcon Research Institute Award.

References

Birch, J. (1997). Clinical use of the American Optical Company (Hardy, Rand and Rittler) pseudoisochromatic plates for red-green colour deficiency. *Ophthalmic and Physiological Optics* **17**, 248–54.

Gunther, K., Neitz, J., & Neitz, M. (2003). A novel missense mutation in the S cone photopigment in a male who made tritan errors on the Neitz Test of Color Vision. *ARVO Abstracts CD-ROM*, abstract 1907.

Hardy, L.H., Rand, G. & Rittler, M.C. (1954). HRR polychromatic plates. *Journal of the Optical Society of America* **44**, 509–523.

Linksz, A. (1964). *An Essay on Color Vision and Clinical Color-Vision Tests*. New York: Grune & Stratton.

Neitz, M. & Neitz, J. (2001). A new test for mass screening of school age children for red-green color vision defects. *Color Research & Application* **26**, S239–S249.

Paulson, H.M. (1971). Color vision testing in the United States Navy. In *Visual Science. Proceedings of the 1968 International Symposium*, ed. Pierce, J.R. & Levene, J.R., pp. 164–176. Bloomington: Indiana University Press.

Schmidt, I. (1952). Comparative evaluation of the Hardy-Rand-Rittler polychromatic plates for testing color vision. USAF School of Aviation Medicine, Project Number 21-31-013. Randolph Field, Texas.

Smith, V.C., Pokorny J. (1975). Spectral sensitivity of the foveal cone photopigments between 400 and 500 nm; *Vision Research* **15**, 161–171.

Vos, J.J., Verkaik, W., Boogaard, J. (1972). The significance of the TMC and HRR color-vision-tests as to red-green defectiveness. *American Journal of Optometry and Archives of the American Academy of Optometry* **49**, 847–859.

Walls, G.L. (1959). How good is the HRR test for color blindness? *American Journal of Optometry and Archives of the American Academy of Optometry* **36**, 169–193.