Error scores on the Farnsworth-Munsell 100-hue test were partitioned into those representing red-green and those representing blue-yellow losses. Data from two groups of normal observers were used. One group showed results characteristic of published norms; one group showed superior performance. Both observers showed a correlation between red-green and blue-yellow scores indicative of a strong performance factor in this test. The difference between blue-yellow and red-green scores eliminates their correlated variance and allows evaluation of the axis. Both groups showed an increase in difference scores, with age indicating development of a blue-yellow axis. This increase was significant for the observers characteristic of the norms. We suggest cutoff scores to allow a decision as to whether a given patient shows a blue-yellow or red-green axis.

The deterioration of color-vision discrimination with advancing age was first reported during World War II. The Farnsworth-Munsell 100-hue test has gained acceptance as a sensitive clinical test of color discrimination. For example, it can be used to follow the course of optic neuritis. Verriest's was the first to establish age norms based on an unselected population of 480 observers. The data collected by Verriest showed optimal discrimination (an average error score of 40) for the age decade of 20 to 29 years, followed by an increase in the average error score of approximately 15 errors per decade in subsequent decades. More recently Verriest, Van Laethem, and Uvijls, using 232 observers, published a new set of norms that differ little from the original set. Inter-eye comparisons were among the statistics they gave. Age norms are used by many ophthalmology centers to decide if a given score is pathologic. The new inter-eye norms allow a decision as to whether a score is pathologic in unioocular disease.

Errors on the Farnsworth-Munsell 100-hue test may accumulate for certain colors, yielding definite bipolar axes characteristic of specific types of color-vision defect. For acquired color-vision defects, the usual axes are blue-yellow or red-green. A question of interest to the clinician is whether an axis exists for a given patient. Based on the literature relating age to color vision, it may be predicted that discrimination of blues and yellows would deteriorate more than discrimination of greens and reds. Age norms for the development of an axis could therefore be helpful to the clinician in deciding if an axis is pathologic.

In this report, we examine two sets of data to evaluate the development of a blue-yellow axis with age. One data set was collected by Krill and Schneiderman; the second data set was collected by us more recently.

**Subjects and Methods**

In 1963, Krill and Schneiderman collected data from 50 observers with an age range of 10 to 50 years. All of the observers underwent an initial ocular examination to rule out eye disease, and all showed normal visual acuity with correction. In 1984, we tested 38 members of the International Society of Appraisers who were gathered in Chicago for their annual convention. Most were members of the sub-specialty gem and jewelry division. The subjects' age range was 26 to 68 years; their visual acuity was not checked. There were 20 male and 18 female subjects: one man failed the Ishihara plate test and proved to be a protanomalous trichromat, reducing our sample with normal color vision to 37 observers. One observer who customarily wore tinted lenses performed the test without correction. Other observers used their customary (clear) refractive correction.

We used pseudoisochromatic plate tests for initial screening (the Ishihara pseudoisochromatic plates,
16-plate edition) to rule out observers with congenital or severe acquired color-vision defects. We used the Farnsworth-Munsell 100-hue test to evaluate chromatic discrimination. In 1963, the tests were given using a MacBeth easel lamp, with a color temperature of 6,740 degrees Kelvin and an illumination of 170 lux. In 1984, the tests were given under 1,500-lux illumination from a desk lamp using two Verilux (F15T8VLX) 15-W fluorescent lamps especially designed for color work, with a color temperature of 6,200 degrees Kelvin and a color-rendering index of 94.8 The tests were given monocularly in 1963 and binocularly in 1984, according to standardized techniques, with no time limit.8 The 1963 study group performed the test twice. Scores from the initial test were used unless the retest showed a “decided change.” Error scores for the Farnsworth-Munsell 100-hue test were calculated from both anchors for each box. The total error score was the sum of the scores for the four boxes. Quadrant analysis10-11 was performed on the raw scores. The total errors were partitioned into blue-yellow and red-green partial scores. The blue-yellow partial scores included errors in caps 1 through 12, 34 through 54, and 76 through 84. The red-green partial scores included errors in caps 13 through 33 and 55 through 75. All of the statistics were calculated using the square root of the total or partial error scores, a transformation which yields a normal distribution of error scores.12-13 Observers were also grouped by decade and group means were calculated for each decade.

**Results**

The Farnsworth-Munsell 100-hue data for the two data sets are shown in Figure 1. The graph shows the square root of the total error score as a function of age. Also shown on the graph are the norms obtained by Verriest, Van Laethem, and Uvijls.7 The 1963 data set agrees well with the norms. The average scores overlap the means of the norms. An analysis of variance showed an age effect significant at a probability of .05. In comparison, the 1984 data show that only two of the 37 observers performed more poorly than the means obtained by Verriest, Van Laethem, and Uvijls.7 Six observers performed better than the 97.5 percentile. The remaining 29 observers showed above-average performance. Furthermore, there is no significant effect of age according to an analysis of variance.

We partitioned the total error scores into those reflecting red-green discrimination errors and those reflecting blue-yellow discrimination errors. The partitioned scores are shown in Figure 2, where the square root of the blue-yellow scores are plotted against the square root of the red-green scores. For the 1963 data set, the Pearson product-moment correlation was 0.71. The 1984 data set showed a Pearson product-moment correlation of 0.46. Two observers lay outside the main scatter of points (Fig. 2, bottom); when these points were omitted, a correlation of 0.67 was obtained. These correlations indicate that almost 50% of the individual variation can be ascribed to variables that affect both blue-yellow and red-green discrimination.

To examine the age effect more closely, we calculated the difference between the square roots of the partitioned scores and show these results for both individual scores and group means in Figure 3. These difference scores do not contain the correlated variance shown in Figure 2. For the 1963 data set, the Pearson product-moment correlation was 0.49 for observers more than 20 years of age. An analysis of variance of the group means showed a significant effect of age (P < .03 for subjects more than 20 years of age, P < .05 for the entire group). For the 1984 data set, there was a Pearson product-moment correlation of 0.22, which is not statistically significant, given the small sample size.

The regression line for the 1963 data set is shown in Figure 3 as a solid line. The standard deviation for the entire data was 1.4. The regression line ± 2 S.D. based on the 1963 data set is also shown plotted on the 1984 data set. Although the 1984 data set showed minimal correlation with age, the scatter of data points fell well within the variance of the 1963 data set. Deviations greater than 2.8 were noted in two observers from the 1984 data set. A 31-year-old woman had a score of −5.3, indicating a red-green axis. A 41-year-old man showed a score of +4.2, indicating a blue-yellow axis. Ocular disease was not excluded nor was visual acuity checked in the 1984 data set.

Significance levels for test-retest and inter-eye comparisons have previously been published,7 but we do not know of significance levels for determining if an axis exists. Although our data sets do not contain sufficient numbers to derive norms for each decade, a preliminary estimate based on the regression line and the overall standard deviation may be made and is shown in the Table.

**Discussion**

The 1963 data set showed results characteristic of the published norms. This data set showed a significant age effect and the group means by the age
Fig. 1 (Smith, Pokorny, and Pass). Square root of the total error score as a function of age. Open triangles indicate individual data; closed triangles, group means; solid and broken lines, means ±2 S.D. obtained by Verriest, Van Laethem, and Uvijis. Top, Data from 1963 study group. Bottom, Data from 1984 study group.
Fig. 2 (Smith, Pokorny, and Pass). Square roots of the partitioned error scores. The square root of the blue-yellow (B-Y) scores is plotted against the square root of the red-green (R-G) scores. Top, Data from the 1963 study group. Bottom, Data from the 1984 study group. Two observers in the 1984 group (open triangles) were omitted from the correlation.
Fig. 3 (Smith, Pokorny, and Pass). The difference between blue-yellow (B-Y) and red-green (R-G) scores as a function of age. The square root of the B-Y score minus the square root of the R-G score is plotted for individual observers (open triangles) and for group means (closed triangles). Solid line indicates the least-squares regression calculated for observers older than 20 years of age in the 1963 data set; broken lines, ± 2 S.D. from the regression. Top, Data from the 1963 study group. Bottom, Data from the 1984 study group.
Calculation of the difference scores for the partitioned blue-yellow and red-green errors eliminated the correlated performance variable. The 1963 data showed a significant age effect, while the 1984 data showed a positive but nonsignificant correlation. The difference scores for the 1984 data set did fall well within the difference scores of the 1963 data set. The difference scores eliminate only the correlated variance of blue-yellow and red-green scores. Possibly some performance elements important for blue-yellow performance remain. For example, Lakowski has shown that adjacent caps in the “blue” quadrant (caps 35 through 75) show smaller color differences than those in other quadrants. Thus, this quadrant is “more difficult” for unselected observers, and we may expect error scores to accumulate in the blue quadrant and to be linked with overall performance. Our data indicate that performance elements remain important in the difference scores. In the 1963 data set, the observers in the youngest age decade (10 to 19 years of age) showed higher blue-yellow difference scores than the next age decade (20 to 29 years of age). These young observers also showed poor performance on the total error scores.

Our partition of the scores into blue-yellow and red-green quadrants is somewhat arbitrary, but it should be appropriate for the majority of acquired color-vision defects of minimal to moderate extent. In more severe color-vision defects, a “scotopic” axis may occur. This axis occurs between the blue-yellow and the red-green axes. A different approach to axis analysis was suggested by Kitahara. Klahra rotated the quadrants to maximize the axis and then calculated a relative difference score. As yet, no analysis of a large sample of color-normal observers has been made using this technique. However, relative scores may be expected to distribute between ±1 in persons with normal color vision (especially for observers making few errors). Thus, it would be very difficult to establish cutoff scores indicative of a pathologic axis.

Acknowledgment

Steven K. Shevell, Ph.D., provided statistical advice.

References