

## Quantifying avian sexual dichromatism: a comparison of methods

Jessica K. Armenta\*, Peter O. Dunn and Linda A. Whittingham

University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA

\*Author for correspondence at present address: Lone Star College – CyFair, Cypress, TX, USA (email: jkarmenta@gmail.com)

Accepted 19 May 2008

### SUMMARY

Recent advances in portable spectrophotometers have allowed researchers to collect quantitative, objective data on colour. There are few comparisons of the different methods used to summarize and analyse spectrophotometer data, however. Using colour data on over 900 species of birds, we compared three methods of calculating sexual dichromatism using spectrophotometer data. We also compared sexual dichromatism calculated from spectrophotometer data, in both the ultraviolet (UV) and bird-visible range, with human estimates of sexual dichromatism. We found that all three methods, principal component analysis, segment classification and colour discriminability, yielded essentially comparable estimates of dichromatism for our extensive sample of birds. Certain methods may be better suited to a particular study depending on the questions addressed and the specific colours examined. We found that human visual estimates of dichromatism were similar to spectrophotometer estimates of dichromatism in the bird-visible range; however, human visual estimates did not predict the extent of UV dichromatism. Therefore, the conclusions of previous studies that relied on human vision to assess sexual dichromatism should be reliable. It is not possible, however, to predict *a priori* whether a species exhibits UV dichromatism without spectrophotometer measurements.

Supplementary material available online at <http://jeb.biologists.org/cgi/content/full/211/15/2423/DC1>

Key words: colouration, colour discriminability, dichromatism, principal component analysis, segment classification, ultraviolet.

### INTRODUCTION

The beautiful and varied colours of bird plumage have long captured the imagination of researchers. Many hypotheses about the evolution of bird colouration have been investigated, with a large proportion of studies focusing on sexual selection (Savalli, 1995; Badyaev and Hill, 2003; Owens, 2006). Studies of sexual selection usually focus on male–female differences in colouration, or sexual dichromatism. Most of these studies have relied on human observers to evaluate bird colouration, or have relied on colour charts (e.g. the Munsell colour system) that are calibrated to human vision (Bennett et al., 1994; Stevens and Cuthill, 2005).

The use of human vision to evaluate bird colouration has raised concerns because of differences between avian and human visual systems (Bennett et al., 1994). Avian colour cones contain coloured oil droplets that are not found in human eyes (Bowmaker, 1980). The oil droplets serve as a filter, preventing certain wavelengths of light from reaching the photoreceptor in the cone, narrowing the spectral sensitivity of the cones and sharpening colour distinctions (Bowmaker, 1980; Vorobyev, 2003; Hart and Vorobyev, 2005). Most importantly, although human eyes contain three colour cones, avian eyes contain four colour cones, one of which captures reflectance in the ultraviolet (UV) and near-UV portion of the electromagnetic spectrum (Bowmaker et al., 1997; Hart et al., 1998; Hart, 2001; Cuthill, 2006). The presence of this cone allows birds to perceive UV wavelengths to which humans are blind (Bennett and Cuthill, 1994).

It has been proposed that UV vision could be a special or hidden channel of communication for birds (e.g. Guilford and Harvey, 1998). Håstad and colleagues suggested that UV reflectance could be a ‘protected’ channel of communication because signals in these wavelengths will be less conspicuous to predators than to conspecifics (Håstad et al., 2005). Hausmann and colleagues

suggested that UV reflectance could be particularly important for sexual selection because the occurrence of UV reflectance in a body region is significantly associated with that body region being used for courtship displays (Hausmann et al., 2003). UV reflectance of plumage appears to be widespread (Eaton and Lanyon, 2003) and can be found in any colour of feather (Burkhardt, 1989). UV reflectance is important for mate choice and mating success in several species of birds, such as the zebra finch [*Taeniopygia guttata* (Bennett et al., 1996)], bluethroat [*Luscinia svecica* (Johnsen et al., 1998)] and blue tit [*Cyanistes caeruleus* (Andersson et al., 1998)]. Several cases of ‘hidden’ UV dichromatism have also been discovered, in which the sexes are monochromatic in the human-visible portion of the spectrum, but dichromatic in the UV-region, such that the sexes appear identical to humans but different to the birds themselves. Species exhibiting ‘hidden’ sexual dichromatism include the blue tit (Hunt et al., 1998), Picui dove [*Columbina picui* (Mahler and Kempenaers, 2002)], yellow-breasted chat [*Icteria virens* (Mays et al., 2004)] and European starling [*Sturnus vulgaris* (Cuthill et al., 1999)]. Eaton used a model that included parameters specific to avian vision (such as opsin sensitivity and number of colour cones) to examine whether species that are sexually monochromatic to humans would be capable of distinguishing between the sexes (Eaton, 2005). He found that in most species the sexes were dimorphic to other members of the same species and suggested that hidden sexual dichromatism may be much more widespread across birds than previously thought.

To measure colour objectively (i.e. independent of the human visual system), it is now recommended that researchers use reflectance spectrophotometers, which have lately become portable and relatively inexpensive (Andersson and Prager, 2006; Montgomerie, 2006). The use of spectrophotometers addresses concerns of human-biased evaluations of colour, but raises several

new questions. For example, how should we quantify and summarize the large amount of spectral data collected by spectrophotometers? Researchers should probably first ask whether the data are to be analysed in terms of the reflectance of the signal or the perception of the receiver. Signals can be objectively measured using reflectance data, but how those signals are interpreted will depend on the visual system and neurobiology of the signal receiver (Endler, 1990; Bennett et al., 1994; Grill and Rush, 2000). In this case, the most appropriate type of data will depend on the objectives of the study. Even when analysing reflectance data, however, there are several possible ways of summarizing the data, including principal component analysis (PCA) (e.g. Cuthill et al., 1999; Mays et al., 2004) and segment classification (e.g. Zuk and Decruyenaere, 1994; Endler and Thery, 1996). To date, few studies have directly compared methods of avian plumage colour analysis (Grill and Rush, 2000), and debate about different methods continues. Our large data set represents most families of birds and provides one of the first direct analyses of different methods of assessing sexual dichromatism across birds.

We calculated indices of sexual dichromatism using three different methods. We used two traditional methods, segment classification and PCA (Endler, 1990), which are independent of the visual system of the receiver. We also used a more recent method, a receptor noise-limited model of colour discriminability (Vorobyev et al., 1998; Eaton, 2005), which focuses on the signal received by other members of the same species of birds. Colour discriminability calculates the quantum catch of the photoreceptors in the eye using data on ambient light, the transmission spectrum of the ocular media, the transmission spectrum of the oil droplets contained within the colour cones, the spectral sensitivity of individual opsins and the effects of photoreceptor noise. We examined the correlations between the sexual dichromatism scores calculated using these three methods of colour analysis in order to determine whether the different methods render fundamentally similar estimates of dichromatism across species. Many studies characterize spectra using the wavelength of maximum reflectance, but this method is problematic when comparing spectra with multiple peaks (e.g. in the UV and human-visible portions of the spectrum) and different shapes. Our data set contains spectra with a wide variety of shapes and colours, so we did not include this method in our comparisons.

We also compared these estimates with human estimates of dichromatism to assess the accuracy of previous studies based on human visual perception (i.e. without a spectrophotometer). Specifically, we wanted to determine: (1) the correlation between dichromatism estimates from the spectrophotometer (in the human visible range) and human visual estimates (Dunn et al., 2001), (2) the correlation between dichromatism estimates in the avian visual range (320–700 nm) from the spectrophotometer and human visual estimates, and (3) the correlation between UV sexual dichromatism (320–400 nm) and estimates of overall avian sexual dichromatism from the spectrophotometer (320–700 nm) or human visual estimates.

## MATERIALS AND METHODS

### Data collection

We measured spectral reflectance of plumage colours from museum specimens of 1003 species of birds. The 1003 species represented 91 of 144 families and 20 of 23 orders in Sibley and Monroe (Sibley and Monroe, 1991). For 987 species, we sampled three male and three female adult specimens in breeding plumage, but for 16 species we could only obtain two specimens for one sex. Thus, we were able to calculate segment classification scores for all 1003 species,

but could only calculate PCA and colour discriminability scores for the 987 species with complete measurements of three individuals of each sex. We considered these sample sizes to be adequate for avoiding type I errors, as over 90% of the variation in our colour variables (see below) lay between rather than within species (Harmon and Losos, 2005).

Five repeated reflectance measurements were taken from six body regions of each specimen: crown, back, throat, belly, wing coverts and tail. We sampled recently collected specimens (90% <50 years old) to avoid problems with colour fading (Armenta et al., 2008).

All reflectance measurements of plumage colour were made with an Ocean Optics USB2000 spectrophotometer and a PX-2 xenon light source (Ocean Optics, Dunedin, FL, USA) and calibrated against a WS-1 white standard, which reflects >98% of light from 250 to 1500 nm wavelengths. We took both white and dark reference measurements at the beginning of every session of measurements. A black rubber test tube stopper mounted on the end of the probe held the probe at a 90 degree angle to the feathers and kept it a fixed distance from the feathers while blocking ambient light. Reflectance measurements were made from 320 to 700 nm, as this spectrum encompasses the bird visible spectrum (Burkhardt, 1989).

### Segment classification

For segment classification, each reflectance measurement was transformed into variables of hue, chroma and brightness following methods in Endler (Endler, 1990) and implemented using the program Spectre (<http://www.uwm.edu/~pdunn/Spectre/Spectre.html>). These three variables are similar to the human perceptions of hue, saturation and lightness ['brightness' (Montgomerie, 2006)]. Spectre calculates brightness as the amount of light reflected by the sample (reflected radiance) in the human-visible spectrum (400–700 nm) relative to the amount of light reflected by the white standard. The program calculates chroma and hue using the formulas of Endler (Endler, 1990). This method divides the human-visible spectrum (400–700 nm) into four equal regions that are approximately the violet/blue (B), green (G), yellow/orange (Y) and red (R) wavelengths and uses the relative difference in reflectance between the red and green segments and the relative difference in reflectance between the yellow/orange and violet/blue segments as axes to construct a colour space. Chroma was calculated as the Euclidean distance to the spectrum from the origin of the colour space based on the relative reflectance of each segment {see Endler's equation 16,  $\text{chroma} = \sqrt{[(R-G)^2 + (Y-B)^2]}$  (Endler, 1990)}, and the value for chroma increases as differences between segment reflectance become greater. Hue is the clockwise angle as measured in the colour space between the spectrum and a spectrum with reflectance only in the R segment. Hue increases as you move through the colour wheel, progressing from red to violet (or in the case of birds, progressing from red to UV). The hue calculation incorporates the relative difference between the reflectance of segments and the chroma of the reflectance curve {see Endler's equation 17,  $\text{hue} = \arcsin[(Y-B)/\text{spectral distance}]$  (Endler, 1990)}. Within each spectrum, the values for brightness, chroma and hue for the five repeated samples were averaged for each body region of each bird.

Spectre calculated the segment classification index of sexual dichromatism for each species based on the difference (Euclidean distance) within each body region between males and females in average hue, chroma and brightness (Endler, 1990). For each body region, an average female hue, chroma and brightness were calculated as well as an average male hue, chroma and brightness. Hue, chroma and brightness were treated as three independent axes so that one point in space represented the average male colour and

one point represented the average female colour for each body region. A distance value was calculated between the male and female points. The distance value for each of the six body regions was summed to produce a dichromatism score for the species. A score of zero indicated a completely monochromatic species, while higher scores indicated increasing dichromatism.

#### PCA estimates of dichromatism

In order to calculate an index of dichromatism using PCA, we averaged the reflectance data within each of 19 bins spanning 20 nm portions of the avian visual spectrum (320–700 nm). The value of each bin was the mean reflectance across those wavelengths contained within the bin. All bin calculations were performed by Spectre. We performed a PCA using the 19 mean reflectance values in JMP v.5 (SAS Institute Inc., Cary, NC, USA). We also performed a PCA using 40 nm bins. Results using the 20 and 40 nm bins were qualitatively similar; therefore, subsequent analyses were carried out using the 20 nm bins.

We used PC1 and PC2 to calculate the index of dichromatism as these two principal components explained more than 97% of the variation in each data set. We calculated the PCA index for each species based on the difference (Euclidean distance) between males and females using PC1 as the *x*-axis and PC2 as the *y*-axis (Endler, 1990). Similar to the segment classification scores, the distance value for each of the six body regions was summed to produce a dichromatism score for the species, with a score of zero indicating a completely monochromatic species and higher scores indicating increasing dichromatism.

#### Colour discriminability

Using Spectre, we calculated the colour discriminability for each body region using the method described by Eaton (Eaton, 2005). The output of each single cone receptor type in a representative avian eye was calculated using the Vorobyev–Osorio colour discrimination model [(Vorobyev et al., 1998); see their equation 1] which takes into account the spectral sensitivity of each cone type, the reflectance of the sample, the background against which the sample is viewed, and the irradiance spectrum of the ambient light. Following Eaton (Eaton, 2005), we set the irradiance to one in all calculations to simulate viewing in a standardized light environment. For the background reflectance spectrum we used deciduous forest canopy because it is similar to several other natural backgrounds in colour discriminability calculations (Eaton, 2005). The reflectance spectrum of deciduous forest canopy was obtained from the online ASTER spectral library maintained by NASA (<http://speclib.jpl.nasa.gov/>) and used as the background in all calculations. We used the spectral sensitivity data for a representative UV-tuned avian eye and a representative violet-tuned avian eye from Endler and Mielke (Endler and Mielke, 2005) because the specific eye characteristics utilized by the model have been measured for only a limited number of species. The sample reflectance is the average of the male or female reflectance curves for a particular body region. The discriminability of each body region was calculated by comparing the difference in the cone stimulation produced by viewing the male colour and the cone stimulation produced by viewing the female colour, taking into account the signal-to-noise ratio for each cone [(Vorobyev et al., 1998); see their equation 8]. For any given male–female comparison, a discriminability of 1.0 is understood to be the smallest difference between two colours that the bird can detect (Siddiqi et al., 2004). Increasing discriminability values represent pairs of colours that are increasingly easy to tell apart. The colour discriminability score for each species was

calculated by summing the discriminability of each individual body region. The discriminability scores calculated for a UV-tuned avian eye and a violet-tuned avian eye were highly correlated ( $r^2=0.99$ ,  $N=978$ ,  $P<0.0001$ ), and we employed the UV-tuned discriminability scores in all further analyses.

#### Human estimates of dichromatism

Human visual estimates of sexual dichromatism were obtained from the paper by Dunn and colleagues (Dunn et al., 2001). In their study, sexual differences in plumage were scored over five regions of the body (head, nape–back–rump, throat–belly, tail and wings) using three scores: 0, no difference between the sexes; 1, difference in shade or intensity; 2, difference in colour or pattern. These estimates were summed for the five body regions of each species, so overall estimates of dimorphism ranged from 0 (monomorphic) to 10 (maximum dichromatism).

#### Statistics

In order to compare the four dichromatism indices, which differ in their range of possible values, we standardized the scores for each index to a mean of zero and a standard deviation of one. The standardized scores were used in all subsequent statistical analyses.

We used the residuals from regressions to identify the species that diverged the most between different indices of dichromatism and between spectrophotometer and human estimates of dichromatism. In order to compare the dichromatism scores calculated with the four different indices, we regressed the colour discriminability scores, the segment classification scores for the avian visual range (320–700 nm) and the human visual estimates of dichromatism against the PCA scores for the avian visual range (320–700 nm) using geometric mean regression (Ricker, 1973). Geometric mean regression was used for all regressions because all scores were standardized to a mean of zero and a standard deviation of one. The three indices of dichromatism do not give the same weight to differences in the spectral properties of brightness and spectral shape. The PCA scores are calculated using the first two principal components, and earlier investigations (Endler, 1990; Cuthill et al., 1999; Grill and Rush, 2000) have interpreted the first principal component as most closely corresponding to brightness, whereas the second principal component corresponds more closely to the shape of the reflectance curve (which has a greater influence on hue and chroma). Therefore, differences in brightness contribute to half of the PCA score. On the other hand, colour discriminability scores take into account differences in the shape of the spectral curves but do not directly account for differences in brightness (Vorobyev et al., 1998; Endler and Mielke, 2005). Segment classification scores are intermediate between the other two methods, because they weight differences in brightness, hue and chroma each as one-third of the score (Endler, 1990).

## RESULTS

#### Comparisons of segment classification, colour discriminability and PCA

The different methods of calculating dichromatism produced similar estimates; however, segment classification and colour discriminability were more in agreement with each other than either was with PCA. The estimates of sexual dichromatism for the avian visual range (320–700 nm) obtained using spectrophotometer measurements were all significantly correlated with each other (PCA and segment classification:  $r^2=0.70$ ,  $N=985$ ,  $P<0.0001$ ; PCA and colour discriminability:  $r^2=0.70$ ,  $N=977$ ,  $P<0.0001$ ; segment classification and colour discriminability:  $r^2=0.87$ ,  $N=979$ ,

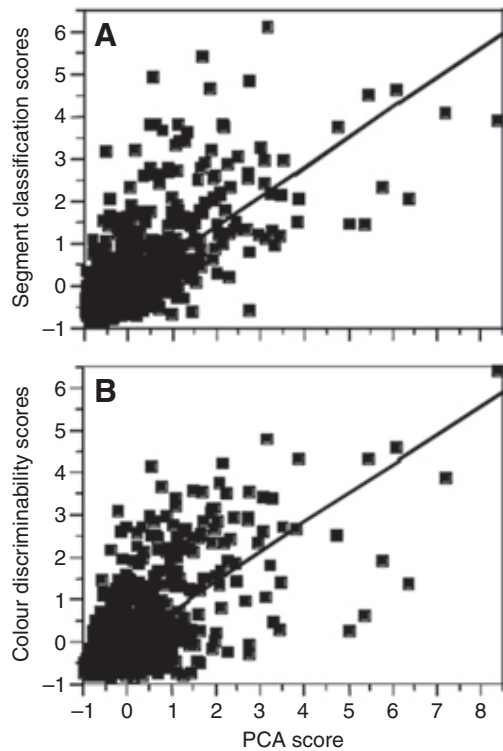


Fig. 1. Comparison of different indices of dichromatism. Geometric mean regression of colour discriminability and segment classification scores against principal component analysis (PCA) scores. Colour discriminability, segment classification and PCA dichromatism scores were calculated using spectrophotometer data. Scores for each index were standardized to a mean of zero and a standard deviation of one. (A) The solid line indicates the regression line of the segment classification scores ( $b=0.71\pm 0.02$ ) against the PCA scores. (B) The solid line indicates the regression line of the colour discriminability ( $b=0.70\pm 0.02$ ) scores against the PCA scores.

$P<0.0001$ ). The colour discriminability ( $b=0.70\pm 0.02$ ) and segment classification (320–700 nm;  $b=0.71\pm 0.02$ ) scores had similar slopes, while the human visual ( $b=0.65\pm 0.02$ ) scores for dichromatism were slightly lower relative to PCA scores (Fig. 1). Therefore, the different indices yielded similar, but not identical estimates of dichromatism. The relative dichromatism scores of each species can be found in the supplementary material Table S1.

To determine where the estimates of dichromatism differed most, we examined the residuals of the bivariate regression of segment classification scores on PCA scores. Overall, PCA gave lower estimates of dichromatism than segment classification particularly in very colourful groups of birds (Fig. 1). The subfamilies represented above the 90th percentile of residuals ( $N=98$  species) were primarily birds of paradise (Paradisae;  $N=7$  of 15 species), cardinals (Cardinalidae;  $N=7$  of 11 species) and blackbirds (Icteridae;  $N=7$  of 24 species). These families contain species in which both sexes are often brightly coloured, thereby minimizing differences in brightness (see supplementary material Table S2A for lists of species). As differences in brightness are weighted more heavily in the PCA index than in the segment classification index, species with sexes of similar brightness will have lower PCA scores. We also examined species of birds below the 10th percentile and found that PCA generally gave a higher dichromatism estimate than segment classification methods particularly for groups of birds in which the colour of both sexes

is drab. This bottom 10% of residuals represents species in which PCA assigned a higher dichromatism score than segment classification. The subfamilies represented in the bottom 10% were primarily sandpipers (Scolopacidae;  $N=5$  of 33 species), gulls (Laridae;  $N=5$  of 8 species), flycatchers (Tyrannidae;  $N=6$  of 41 species) and ducks (Anatidae;  $N=12$  of 24 species). These families mostly contain birds in which the plumage of both sexes is brown, grey or white, minimizing variation in the shapes of the reflectance curves. Differences in the shapes of male and female reflectance curves (primarily hue and chroma) are weighted more heavily in the segment classification index than in the PCA index. Therefore, smaller differences between the sexes in the shape of curves will lead to lower scores for the segment classification than for the PCA method.

We also examined the residuals of the bivariate regression of colour discriminability scores on PCA scores. The species in the top 10% of residuals ( $N=97$  species) had a higher dichromatism score from colour discriminability than from PCA. These species were primarily manakins (Pipridae;  $N=6$  of 8 species), tanagers (Thraupidae;  $N=8$  of 20 species), cardinals (Cardinalidae;  $N=8$  of 11 species) and blackbirds (Icteridae;  $N=10$  of 24 species). The species in these families often have sexes that are both brightly coloured but the sexes may exhibit different colours, meaning that they differ primarily in the shape of the reflectance curves (see supplementary material Table S2B). Species in which the sexes have reflectance curves with different shapes but similar total brightness would be given a larger score by colour discriminability. The bottom 10th percentile of residuals represents species in which PCA assigned a higher dichromatism score than colour discriminability. The subfamilies represented in the bottom 10th percentile were primarily parulid warblers (Parulidae;  $N=5$  of 47 species) and ducks (Anatidae;  $N=14$  of 24 species). Again, these families mostly contain birds whose plumage is brown, grey or white and the sexes are similar except for small plumage regions. Therefore, PCA, which gives a heavier weight to brightness than to differences in spectral shape, renders a higher dichromatism score than colour discriminability because variation in the shapes of the reflectance curves is minimal.

We also compared the species that differed the most between the segment classification and colour discriminability dichromatism scores. The species in the top 10% of residuals ( $N=97$  species) represent those species in which colour discriminability assigned a higher dichromatism score than segment classification. The subfamilies best represented in the top 10% were primarily parulid warblers (Parulidae;  $N=7$  of 47 species), tanagers (Thraupidae;  $N=8$  of 20 species) and blackbirds (Icteridae;  $N=8$  of 24 species). These members of the parulid warbler and blackbird families tend to be the more colourful members of their respective families and the species that show greater differences between the sexes (see supplementary material Table S2C). Colourful species receive a relatively higher score from colour discriminability (which considers spectral shape but does not consider brightness) than from segment classification. The bottom 10% of residuals represents species in which segment classification assigned a higher dichromatism score than colour discriminability. The subfamilies best represented in the bottom 10% were monarch flycatchers (Monarchidae;  $N=6$  of 14 species), butcherbirds (Artamidae;  $N=9$  of 13 species) and crows (Corvidae;  $N=9$  of 17 species). In these species, the sexes differ little in spectral shape and usually differ more in brightness. Segment classification, which considers brightness along with spectral shape, will give higher scores to these species than will colour discriminability, which only considers spectral shape.

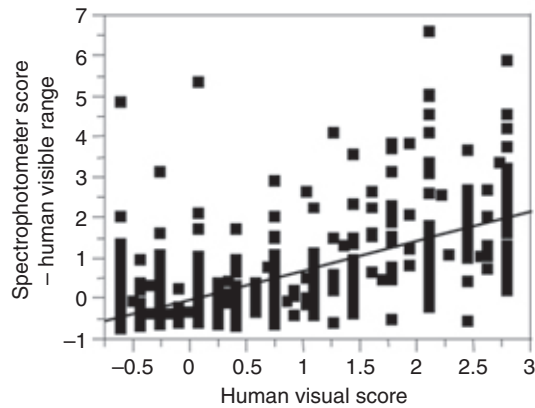


Fig. 2. Assessing dichromatism in the human visual range. Geometric mean regression of segment classification estimates of sexual dichromatism in the human visual (400–700 nm) range against human visual estimates ( $r^2=0.73$ ,  $N=978$ ,  $P<0.0001$ ). Segment classification scores were calculated using spectrophotometer data. Scores for each index were standardized to a mean of zero and a standard deviation of one.

#### Comparisons of human and spectrophotometer scores

Finally, we compared the human visual estimates of dichromatism (see Dunn et al., 2001) and other methods. In the human visual range, segment classification estimates of sexual dichromatism from the spectrophotometer were correlated with human visual estimates (Fig. 2;  $r^2=0.73$ ,  $N=978$ ,  $P<0.0001$ ). Human visual estimates of dichromatism were also correlated with those from the spectrophotometer over the avian visual range (320–700 nm;  $r^2=0.73$ ,  $N=978$ ,  $P<0.0001$ ). The subfamilies that differed the most between spectrophotometer (segment classification) and human estimates of dichromatism (top 10% of residuals;  $N=97$  species) were primarily fairy-wrens (Maluridae;  $N=10$  of 12 species), birds of paradise (Paradisae;  $N=6$  of 15 species) butcherbirds (Artamidae;  $N=7$  of 13 species) and monarch flycatchers (Monarchidae;  $N=7$  of 14 species). Many of the species in these subfamilies have iridescent plumage, and differences between the sexes in this iridescence may not be adequately quantified by human observers. Among the bottom 10% of residuals, which represents species that humans thought were more dichromatic than the spectrophotometer, were parulid warblers (Parulidae;  $N=10$  of 47 species), American sparrows (Emberizidae;  $N=9$  of 52 species), ducks (Anatidae;  $N=9$  of 24 species) and blackbirds (Icteridae;  $N=6$  of 24 species). In some of these species, the spectrophotometer may not have measured small patches of sexually dichromatic plumage (we only sampled six body regions), and this omission could have led to lower estimates of dichromatism than those by human observers who examined the entire body.

Human visual estimates of dichromatism could not accurately predict dichromatism in the UV range (as estimated by the PCA method), although there was a generally positive relationship (Fig. 3;  $r^2=0.25$ ,  $N=960$ ,  $P<0.0001$ ). Similarly, the relationship between spectrophotometer-based estimates of dichromatism in the UV range and dichromatism in both the human ( $r^2=0.50$ ,  $N=1003$ ,  $P<0.0001$ ) and avian visual range ( $r^2=0.55$ ,  $N=1003$ ,  $P<0.0001$ ) were significant, but of little predictive value. The subfamilies that showed the largest differences between UV dichromatism and human estimates of dichromatism (top 10% of residuals;  $N=100$  species) were primarily birds of paradise (Paradisae;  $N=10$  of 15 species), blackbirds (Icteridae;  $N=5$  of 24 species), parulid warblers (Parulidae;  $N=5$  of 47 species) and Australasian robins

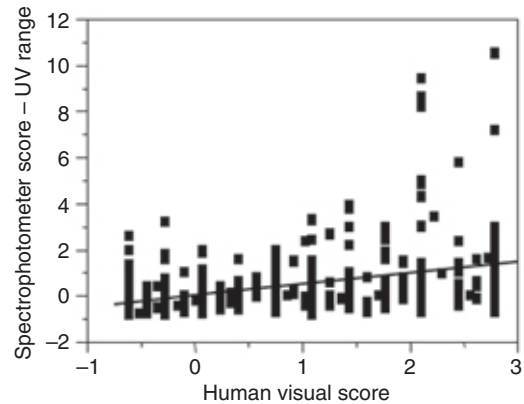


Fig. 3. Relationship between UV and human-visible dichromatism. Geometric mean regression of PCA estimates of dichromatism in the UV range (320–400 nm) against human visual estimates ( $r^2=0.25$ ,  $N=960$ ,  $P<0.0001$ ). PCA scores were calculated using spectrophotometer data and human-visible scores are from the study of Dunn and colleagues (Dunn et al., 2001). Scores for each index were standardized to a mean of zero and a standard deviation of one.

(Eopsaltriidae;  $N=5$  of 23 species). Several of these subfamilies also occurred in the top 10% of residuals between spectrophotometer and human-estimated dimorphism. As above, many of the species in these subfamilies have iridescent plumage that often reflects in the UV range and that may not be adequately quantified by human observers. Many of the species in these subfamilies also have yellow, orange and red plumage patches, which are presumably carotenoid based and also often reflect in the UV range. Among the bottom 10% of residuals, representing species that were more dichromatic in the human-visible range than they were in the UV range, were American sparrows (Emberizidae;  $N=9$  of 52 species), blackbirds (Icteridae;  $N=8$  of 24 species) and gulls (Laridae;  $N=7$  of 8 species). It is interesting that blackbirds appear in both the top and bottom 10% of the residuals, but the species found in the bottom 10% mostly lack carotenoid plumage patches and tend to be more uniformly black, a colour that does not usually reflect light in the UV range unless it is iridescent. Likewise, the American sparrows tend to be brown, another colour that also seldom has much UV reflectance.

#### DISCUSSION

Our data set represented almost 10% of all species and 63% of families of birds. In this large-scale analysis of sexual dichromatism we found that most methods (using spectrophotometer data) gave similar estimates of dichromatism ( $r^2$  values were all  $>0.70$ ), even though some focused on the signaller (PCA and segment classification) and others (colour discriminability) focused on the receiver of the colour signal (other members of the same species). Nevertheless, each of the three methods we used to quantify sexual dichromatism may have advantages for particular types of studies.

PCA has been one of the more widely used methods to estimate dichromatism (e.g. Cuthill et al., 1999; Grill and Rush, 2000; Mays et al., 2004). PCA was the simplest and quickest method of achieving sexual dichromatism scores, requiring only standard statistical software. PCA utilized data collected at all wavelengths of the spectrophotometer reading, because all data were summarized into 'bins', which represented the average reflectance across a range of wavelengths (Endler, 1990). Thus, PCA scores were the least transformed of the three methods and the closest to the raw spectrophotometer data.

The fact that PCA was not dependent upon avian visual system characteristics such as opsin sensitivity could be either an advantage or a disadvantage, depending on the question being addressed. As it did not account for the visual system of the viewer, PCA would not be suitable for studying the co-evolution of signals and eye design (Vorobyev et al., 1998). Colour signals could be under multiple selection pressures, however, not just those originating from conspecifics with the same visual system. In these cases, it could be most appropriate to use a vision-neutral method of analysis, such as PCA, to evaluate sexual dichromatism. Also, detailed data on many aspects of avian visual systems, such as the spectral sensitivity of cone cells, are unavailable for the majority of species (Eaton, 2005). Where a vision-neutral method was appropriate, PCA would be a valid method for describing differences within groups of reflectance readings (Endler, 1990; Vorobyev et al., 1998). If a study is concerned primarily with the receiver's visual response, however, treating all wavelengths equally ignores the fact that avian opsins do not respond to all wavelengths equally (Hart et al., 2000; Endler and Mielke, 2005). Therefore, differences that were detected statistically by PCA might or might not represent biologically meaningful differences to a receiver.

Although PCA can easily summarize large amounts of spectrophotometer data, it has some statistical problems (Endler and Mielke, 2005). First, reflectance readings from adjacent bins of wavelengths are highly correlated, and thus PCA scores are not independent observations when using inferential tests (Endler and Mielke, 2005). Additionally, the principal components generated in an analysis are dependent upon the data entered into the PCA, so PCA scores from one study cannot be compared directly with those of another study without a re-analysis of the raw data (Endler, 1990; Cuthill et al., 1999).

PCA is most sensitive to brightness of the reflectance spectrum, and least sensitive to spectral shape (Endler, 1990; Grill and Rush, 2000). We found that PCA was more sensitive to differences between spectra with a similar, fairly uniform shape. Therefore, PCA might be best for a study focused on groups of birds with primarily brown, white or grey plumage. Grill and Rush (Grill and Rush, 2000) examined principal components calculated from standardized Munsell colour chips and suggested that PCA would be better for studies of variation within one colour class, such as when all spectral readings are various shades of yellow. We found that PCA was least able to quantify sex differences in very colourful species. In these colourful species, the brightness of the sexes could be very similar, but the multiple colours presented a variety of spectral shapes, and PCA was less able to quantify these types of differences between spectra. For example, in the summer tanager (*Piranga rubra*), both the male and female are brightly coloured, but the female is yellow while the male is red. Comparing these different colours with similar brightness, PCA yielded a relatively lower dichromatism score than either segment classification or colour discriminability.

Like PCA, segment classification simply summarized reflectance data from the point of view of the signaller and did not use avian visual characteristics (i.e. the receiver). However, the boundaries of the segments used in segment classification calculations were similar to the limits of spectral sensitivity for each of the four opsins found in an avian eye, especially for a UV-tuned eye (Hart et al., 2000; Endler and Mielke, 2005). Within each segment, however, all wavelengths were treated equally, and the same problem arose in segment classification as in PCA, namely that the avian eye does not treat all wavelengths as equal. In contrast, how sensitive opsins are to each wavelength of light is factored into calculations of colour

discriminability. Segment classification and colour discriminability, however, gave similar estimates of dichromatism.

Segment classification has an advantage over PCA and colour discriminability in that it generates estimates of hue, chroma and brightness that can be related to human attributes of colour (hue, saturation and lightness). It should be noted, however, that these variables would not translate directly to avian perception of colour because they only construct a two-dimensional colour space, and the tetrachromatic vision of birds would necessitate a three-dimensional colour space. Segment classification, like PCA, would be appropriate when a vision-neutral method is most suited to the study, or when visual characteristics of the study species are not readily available. Segment classification takes into account the shape of the spectral reflectance curve more than PCA does, however, and we found that segment classification was better than PCA at detecting differences between very colourful, but equally bright, spectra.

Segment classification produced dichromatism scores that were intermediate between PCA and colour discriminability in several ways. Segment classification required more computational steps than PCA, but still far fewer than those required to calculate colour discriminability scores. Segment classification puts a greater emphasis on brightness differences than colour discriminability, but weights brightness differences less heavily than PCA. Conversely, segment classification weights differences in spectral shape less heavily than colour discriminability, but puts a greater emphasis on these differences than PCA would.

Calculating colour discriminability dichromatism scores requires detailed information about ambient light, background colours and the avian eye, including properties of its ocular media, opsins and oil droplets (Vorobyev et al., 1998). Often, researchers are forced to use representative avian eye characteristics (Eaton, 2005; Håstad et al., 2005) or eliminate terms from the model (Eaton, 2005). Introducing more and very detailed parameters into the model may increase noise in the results more than it increases the accuracy of the signal, especially if the parameters are not specific to the species in question. In any case, our results from both PCA and segment classification produced similar estimates of dichromatism to those from colour discriminability (all  $r^2 > 0.70$ ) for this large data set.

If all of the necessary data are available, however, colour discriminability scores should be able to predict which colour differences are biologically relevant to birds and therefore can reveal new insights into the evolution of avian signals. Eaton calculated sexual dichromatism scores for 139 species of passerines classified as sexually monochromatic by humans and concluded that more than 90% of these species were likely to be dichromatic when viewed by other birds of the same species (Eaton, 2005). Using the same criteria, that a species must have at least one body region with a colour discriminability score greater than 1 to be considered dichromatic, we examined the 574 species in our data set that were monochromatic to humans and found that 48% of these species were likely to be dichromatic when viewed by other birds of the same species. The fact that our sample has a lower percentage of dichromatic birds is probably due to the fact that our sample includes both passerines and non-passerines, and non-passerines exhibit lower levels of dichromatism than passerines. Håstad and colleagues (Håstad et al., 2005) calculated colour discriminability scores for songbirds using data on the optical properties of a UV-tuned eye, found in songbirds (Odeen and Håstad, 2003), and a violet-tuned eye, which is found in avian predators such as raptors and crows (Odeen and Håstad, 2003). These two sets of colour discriminability scores allowed Håstad and colleagues to conclude that signal

evolution was actively occurring because the songbird colour patterns were more conspicuous to the eyes of conspecifics, and more cryptic to the eyes of predators (Håstad et al., 2005). Endler and colleagues used a similar calculation of photon catch (though a different statistical approach known as LSED-MRPP) to examine the evolution of bowerbird signals in both the birds' plumage and their bower decorations (Endler et al., 2005). They found that bower ornaments, instead of having colours similar to the plumage colours of the species, were separate colours that increased the overall signal to conspecifics.

When the intended receiver of a signal is known, and the receiver's visual system is well described, it is advantageous to compare spectral data using the discriminability model in order to include the influence of the receiver's optical properties on the perception of colour differences from the receiver's point of view. We found that colour discriminability was particularly adept at detecting differences between species in which both sexes had very colourful spectra with different shapes. For example, colour discriminability detected relatively larger differences than segment classification or PCA between the sexes in colourful species such as tanagers and blackbirds, species in which reflectance spectrum readings of males and females have similar brightness but different spectral shapes. Colour discriminability, however, is unable to detect differences between spectra when most of the variation is in brightness instead of spectral shape. The colour discriminability model does not include brightness as a term because brightness is processed separately from colour by the visual system (Vorobyev et al., 1998; Endler and Mielke, 2005). This omission does not necessarily mean that differences in brightness are not meaningful to birds, however, and, indeed, there are several studies of mate choice that indicate females respond to male brightness [e.g. blue tit (Hunt et al., 1999); black-capped chickadee, *Poecile atricapilla* (Doucet et al., 2005); golden-collared manakin, *Manacus vitellinus* (Stein and Uy, 2006)].

Many previous studies were carried out without the benefit of spectrophotometers. When we evaluated human visual estimates of dichromatism, we found that they were correlated with estimates of dichromatism obtained using the spectrophotometer when the analysis was restricted to the human visible range. Also, human vision gave similar, but not identical, estimates of dichromatism to spectrophotometer estimates of dichromatism for the entire avian visible range. Human estimates failed to predict UV dichromatism reliably, however, as did dichromatism scores obtained using the spectrophotometer over the avian visible range. Therefore, in the majority of cases, human estimates of dichromatism appear to predict reliably the amount of dichromatism that birds would see. Unfortunately, if we rely on human visual estimates alone, we cannot predict *a priori* for which species those estimates will be wrong. Previous studies that utilized human visual estimates of dichromatism should be mostly reliable because of the good correlation between human dichromatism estimates and bird-visible dichromatism estimates. Particular caution should be used, however, when the studies deal with either iridescent or carotenoid colours, as human vision appears to be particularly poor at assessing these groups of colours.

Despite major differences in how the methods calculated sexual dichromatism, all methods yielded comparable estimates of dichromatism. Which of the three methods is most appropriate will depend upon the question being addressed. Researchers should consider whether there are one or multiple intended receivers of the signal and how much is known about the visual system of the receiver. For insights into the co-evolution of signals and the visual system, researchers should rely upon colour discriminability. Also

important to consider is how much variation in spectral shape will be present in the data set. If the data set contains little variation in shape and dichromatism scores do not need to be comparable to other studies, then PCA may be useful. If there is a lot of variation in the shape of the reflectance spectra in the sample, then segment classification or colour discriminability will be better able to characterize differences between spectra. Segment classification will be the most useful method when there is a large amount of variety in both spectral brightness and shape, and receiver-neutral estimates of dichromatism are desired.

We thank the curators and collection managers of the following museums for access to specimens: American Museum of Natural History (New York), Australian National Wildlife Collection (Canberra), Field Museum of Natural History (Chicago), Louisiana State University Museum of Natural Science (Baton Rouge), Museum Victoria (Melbourne), and National Museum of Natural History (Washington, DC). We thank John Endler and Nathan Hart for providing avian spectral sensitivity data. John Berges and Ken Yasukawa provided helpful comments. This work was supported by NSF DEB-0215560 to P.O.D. and L.A.W. and by an NSF graduate research fellowship to J.K.A.

## REFERENCES

- Andersson, S. and Prager, M. (2006). Quantifying colours. In *Bird Coloration*. Vol. 1 (ed. G. E. Hill and K. J. McGraw), pp. 41-89. Cambridge, MA.: Harvard University Press.
- Andersson, S., Ornborg, J. and Andersson, M. (1998). Ultraviolet sexual dimorphism and assortative mating in blue tits. *Proc. R. Soc. Lond. B* **265**, 445-450.
- Armenta, J. K., Dunn, P. O. and Whittingham, L. A. (2008). Effects of specimen age on plumage color. *Auk*. (in press).
- Badyaev, A. V. and Hill, G. E. (2003). Avian sexual dichromatism in relation to history and current selection. *Annu. Rev. Ecol. Syst.* **34**, 27-49.
- Bennett, A. and Cuthill, I. C. (1994). Ultraviolet vision in birds: what is its function? *Vision Res.* **34**, 1471-1478.
- Bennett, A. T. D., Cuthill, I. C. and Norris, K. J. (1994). Sexual selection and the mismeasure of color. *Am. Nat.* **144**, 848-860.
- Bennett, A. T. D., Cuthill, I. C., Partridge, J. and Maier, E. J. (1996). Ultraviolet vision and mate choice in zebra finches. *Nature* **380**, 433-435.
- Bowmaker, J. K. (1980). Color vision in birds and the role of oil droplets. *Trends Neurosci.* **3**, 196-199.
- Bowmaker, J. K., Heath, L. A., Wilkie, S. E. and Hunt, D. M. (1997). Visual pigments and oil droplets from six classes of photoreceptor in the retinas of birds. *Vision Res.* **37**, 2183-2194.
- Burkhardt, D. (1989). UV vision: a bird's eye view of feathers. *J. Comp. Physiol. A* **164**, 787-796.
- Cuthill, I. C. (2006). Color Perception. In *Bird Coloration*. Vol. 1 (ed. G. E. Hill and K. J. McGraw), pp. 3-40. Cambridge, MA.: Harvard University Press.
- Cuthill, I., Bennett, A., Partridge, J. and Maier, E. (1999). Plumage reflectance and the objective assessment of avian sexual dichromatism. *Am. Nat.* **153**, 183-200.
- Doucet, S. M., Mennill, D. J., Montgomerie, R., Boag, P. T. and Ratcliffe, L. M. (2005). Achromatic plumage reflectance predicts reproductive success in male black-capped chickadees. *Behav. Ecol.* **16**, 218-222.
- Dunn, P. O., Whittingham, L. A. and Pitcher, T. E. (2001). Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution* **55**, 161-175.
- Eaton, M. D. (2005). Human vision fails to distinguish widespread sexual dichromatism among sexually "monochromatic" birds. *Proc. Natl. Acad. Sci. USA* **102**, 10942-10946.
- Eaton, M. D. and Lanyon, S. M. (2003). The ubiquity of avian ultraviolet plumage reflectance. *Proc. R. Soc. Lond. B, Biol. Sci.* **270**, 1721-1726.
- Endler, J. A. (1990). On the measure and classification of colour in studies of animal colour patterns. *Biol. J. Linn. Soc. Lond.* **41**, 315-352.
- Endler, J. A. and Mielke, P. W. (2005). Comparing entire colour patterns as birds see them. *Biol. J. Linn. Soc. Lond.* **86**, 405-431.
- Endler, J. A. and Thery, M. (1996). Interacting effects of lek placement, display behavior, ambient light, and color patterns in three neotropical forest-dwelling birds. *Am. Nat.* **148**, 421-452.
- Endler, J. A., Wescott, D. A., Madden, J. R. and Robson, T. (2005). Animal visual systems and the evolution of color patterns: sensory processing illuminates signal evolution. *Evolution* **59**, 1795-1818.
- Grill, C. P. and Rush, V. N. (2000). Analyzing spectral data: comparison and application of two techniques. *Biol. J. Linn. Soc. Lond.* **69**, 121-138.
- Guilford, T. and Harvey, P. H. (1998). The purple patch. *Nature*, **392**, 867-868.
- Harmon, L. J. and Losos, J. B. (2005). The effect of intraspecific sample size on type I and type II error rates in comparative studies. *Evolution* **59**, 2705-2710.
- Hart, N. S. (2001). The visual ecology of avian photoreceptors. *Prog. Retin. Eye Res.* **20**, 675-703.
- Hart, N. S. and Vorobyev, M. (2005). Modelling oil droplet absorption spectra and spectral sensitivities of bird cone photoreceptors. *J. Comp. Physiol. A* **191**, 381-392.
- Hart, N. S., Partridge, J. C. and Cuthill, I. C. (1998). Visual pigments, oil droplets and cone photoreceptor distribution in the European starling (*Sturnus vulgaris*). *J. Exp. Biol.* **201**, 1433-1446.
- Hart, N. S., Partridge, J. C., Cuthill, I. C. and Bennett, A. T. D. (2000). Visual pigments, oil droplets, ocular media and cone photoreceptor distribution in two

- species of passerine bird: the blue tit (*Parus caeruleus*) and the blackbird (*Turdus merula*). *J. Comp. Physiol. A* **186**, 375-387.
- Håstad, O., Victorsson, J. and Odeen, A.** (2005). Differences in color vision make passerines less conspicuous in the eyes of their predators. *Proc. Natl. Acad. Sci. USA* **102**, 6391-6394.
- Hausmann, F., Arnold, K. E., Marshall, N. J. and Owens, I. P. F.** (2003). Ultraviolet signals in birds are special. *Proc. R. Soc. Lond. B, Biol. Sci.* **270**, 61-67.
- Hunt, S., Bennett, A. T. D., Cuthill, I. C. and Griffiths, R.** (1998). Blue tits are ultraviolet tits. *Proc. R. Soc. Lond. B, Biol. Sci.* **265**, 451-455.
- Hunt, S., Cuthill, I. C., Bennett, A. T. D. and Griffiths, R.** (1999). Preferences for ultraviolet partners in the blue tit. *Anim. Behav.* **58**, 809-815.
- Johnsen, A., Andersson, S., Ornborg, J. and Lifjeld, J. T.** (1998). Ultraviolet plumage ornamentation affects social mate choice and sperm competition in bluethroats (*Aves: Luscinia s. svecica*): a field experiment. *Proc. R. Soc. Lond. B, Biol. Sci.* **265**, 1313-1318.
- Mahler, B. A. and Kempnaers, B.** (2002). Objective assessment of sexual plumage dichromatism in the Picui Dove. *Condor* **104**, 248-254.
- Mays, H. L., McGraw, K. J., Ritchison, G., Cooper, S., Rush, V. and Parker, R. S.** (2004). Sexual dichromatism in the yellow-breasted chat *Icteria virens*: spectrophotometric analysis and biochemical basis. *J. Avian Biol.* **35**, 125-134.
- Montgomerie, R.** (2006). Analyzing colours. In *Bird Coloration* Vol.1 (ed. G. E. Hill and K. J. McGraw), pp. 90-147. Cambridge, MA.: Harvard University Press.
- Odeen, A. and Håstad, O.** (2003). Complex distribution of avian color vision systems revealed by sequencing the SWS1 opsin from total DNA. *Mol. Biol. Evol.* **20**, 855-861.
- Owens, I. P. F.** (2006). Ecological explanations for interspecific variability in coloration. In *Bird Coloration*. Vol. 2 (ed. G. E. Hill and K. J. McGraw), pp. 380-416. Cambridge, MA.: Harvard University Press.
- Ricker, W. E.** (1973). Linear regressions in fishery research. *J. Fish. Res. Board Can.* **30**, 409-434.
- Savalli, U. M.** (1995). The evolution of bird coloration and elaboration: a review of hypotheses. *Curr. Ornith.* **12**, 141-190.
- Sibley, C. G. and Monroe, B. L.** (1991). *Distribution and Taxonomy of Birds of the World*. New Haven, CT: Yale University Press.
- Siddiqi, A., Cronin, T. W., Loew, E. R., Vorobyev, M. and Summers, K.** (2004). Interspecific and intraspecific views of colour signals in the strawberry poison frog *Dendrobates pumilio*. *J. Exp. Biol.* **207**, 2471-2485.
- Stein, A. C. and Uy, J. A. C.** (2006). Plumage brightness predicts male mating success in the lekking golden-collared manakin, *Manacus vitellinus*. *Behav. Ecol.* **17**, 41-47.
- Stevens, M. and Cuthill, I. C.** (2005). The unsuitability of html-based-colour charts for estimating animal colours – a comment on Berggren and Merilä (2004). *Front. Zool.* **2**, 1-9.
- Vorobyev, M.** (2003). Coloured oil droplets enhance colour discrimination. *Proc. R. Soc. Lond. B, Biol. Sci.* **270**, 1255-1261.
- Vorobyev, M., Osorio, D., Bennett, A. T. D., Marshall, N. J. and Cuthill, I. C.** (1998). Tetrachromacy, oil droplets and bird plumage colours. *J. Comp. Physiol. A* **183**, 621-633.
- Zuk, M. and Decruyenaere, J. G.** (1994). Measuring individual variation in colour – a comparison of two techniques. *Biol. J. Linnean Soc. Lond.* **53**, 165-173.