Biological Journal of the Linnean Society, 2008, **94**, 231–240. With 4 figures

A simple method for measuring colour in wild animals: validation and use on chest patch colour in geladas (*Theropithecus gelada*)

THORE J. BERGMAN[1,2]* and JACINTA C. BEEHNER[1,3]

[1]*Department of Psychology, University of Michigan, Ann Arbor, MI 48109, USA*
[2]*Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA*
[3]*Department of Anthropology, University of Michigan, Ann Arbor, MI 48109, USA*

Received 23 April 2007; accepted for publication 31 July 2007

Adaptive hypotheses about colour variation are widespread in behavioural ecology, and several methods of objective colour assessment have been proposed and validated for use in a wide variety of taxa. However, to date, the most objective and reliable methods of assessing colour are not readily applied to wild animals. In the present study, we present a simple method for assessing colour in unrestrained, wild subjects using digital photography. The method we describe uses a digital camera, a colour standard, and colour analysis software, and can be used to measure any part of the visible colour spectrum. We demonstrate that the method: (1) is accurate and precise across different light conditions; (2) satisfies previous criteria regarding linearity and red, green, and blue equality; and (3) can be independently validated visually. In contrast with previous digital methods, this method can be used under natural light conditions and can be readily applied to subjects in their natural habitat. To illustrate this, we use the method to measure chest colour in wild geladas (*Theropithecus gelada*). Unique among primates, geladas have a red patch of skin on their chest and neck, which, for males, is thought to be a sexually selected signal. Offering some support to this hypothesis, we found differences in chest 'redness' for males across different age groups, with males in their reproductive prime exhibiting the reddest chests. *Biological Journal of the Linnean Society*, 2008, **94**, 231–240.

ADDITIONAL KEYWORDS: colour measurement – colour vision – primate – signal.

INTRODUCTION

Colour variation can be found in a multitude of animal taxa, and adaptive hypotheses about animal coloration date as far back as Darwin (1859, 1871). Many studies in both behavioural ecology and evolutionary biology provide evidence that colour serves adaptive functions ranging from crypsis (Endler, 1984; Litvaitis, 1991; Westmoreland & Kiltie, 1996), mimicry (Bower, 1958; Malcolm, 1990), and warning coloration (Bower, 1958; Endler, 1986), to sexual signalling (Evans & Morris, 1996; Gerald, 2001; Bourne, Breden & Allen, 2003; Cooper & Hosey, 2003; Waitt *et al*., 2003), and social communication (Losey, 2003). However, objective determination of colour is hampered by the fact that human observers perceive colours differently under different light conditions (Endler, 1990). Methods employing either arbitrary categorical rankings (e.g. dark versus light) or visual colour matching to colour standards (The Munsell Book of Color, 1976) are necessarily subjective (Endler, 1990) because measurements may vary under different light conditions, across different observers, across time (as observers grow fatigued), or across a study (if colour standards weather or fade). Moreover, these methods are not amenable to statistical analyses. Therefore, to adequately test hypotheses about colour variation, researchers should measure colour using methods that do not rely on a human observer's perception of colour (Endler, 1990; Kilner, 1997).

*Corresponding author. E-mail: thore@umich.edu

The most objective and reliable method for measuring colour is spectrophotometry (Zuk & Decruyenaere, 1994), which measures the distribution of wavelengths reflected via a digital device known as a spectrophotometer. Spectrophotometry can also measure colour beyond that which is visible to humans (e.g. ultraviolet wavelengths), and therefore is useful for animals that might perceive colours beyond the spectrum visible to humans. However, in addition to being expensive, spectrophotometry does not readily apply to wild animals because the lens must be placed within a few inches of a flattened sample of a fixed size, and samples must be illuminated by a constant light source.

More recently, due to its ease and affordability, digital photography has become the method of choice for measuring animal colour (Kilner, 1997; Goda & Fujii, 1998; Villafuerte & Negro, 1998; Gerald *et al*., 2001; Losey, 2003; Stevens *et al*., 2007). Digital photography (in combination with various computer software programs) is able to objectively quantify colour and is more sensitive than human vision at detecting differences in colour (Villafuerte & Negro, 1998). Despite its ease and accessibility, quantifying colour using digital photography also has its problems. For example, a recent review highlights several common methodological concerns that relate to problems inherent in most digital cameras (Stevens *et al*., 2007); these mainly include the nonlinearity of a camera's response to light intensity and biases in a camera's processing of an image towards particular wavelengths. Such problems, if not corrected, can result in inaccurate colour measurements. Additionally, most digital photography methods require either: (1) a standard in the same photo as the subject or (2) controlled lighting conditions, neither of which is typically possible with unrestrained subjects in their natural habitat. Therefore, the most objective and reliable methods of assessing colour are only applicable to preserved, captive or sedated specimens/animals. To date, no objective method of colour assessment has been validated for use in wild animals.

In the present study, we present a simple method for assessing colour in unrestrained subjects using digital photography. The method we describe uses a digital camera, a colour standard, and colour analysis software. We demonstrate that the method: (1) is accurate and precise; (2) satisfies the criteria regarding linearity and equalization laid out by Stevens *et al*. (2007); and (3) can be independently validated visually. The primary advantage of the described method over previous ones is that it can be used under natural light conditions and can be applied to subjects in their natural habitat. To illustrate this, we apply the method to wild geladas (*Theropithecus gelada*). Unique among primates, geladas have a red patch of skin on the chest and neck. Although chest patch colour in females changes with reproductive condition, changes in male chest patch colour appear to be status-based and may be a sexually selected signal (Dunbar, 1984). However, variation in male chest colour has never been previously quantified. As a first step towards addressing this hypothesis, we examine broad colour variation in gelada males across different age groups. If, as previously suggested, male chest patch colour is a signal of male quality, then we expect that reproductively active males (i.e. prime adults) should have the reddest chests.

MATERIAL AND METHODS

Equipment

As a colour standard, we used a GretagMacbeth ColorChecker colour rendition chart (product no. 50105, manufactured by Munsell Colour, division of GretagMacbeth). The GretagMacbeth ColorChecker chart (hereafter referred to as the ColorChecker chart) is commonly used in photographic and video fields, either for colour calibration or for assessing the colour rendering accuracy of an imaging device. The ColorChecker chart itself consists of a checkerboard array of 24 coloured squares formulated to represent common natural colours (including skin colours), in addition to primary colours, and a six-step grey scale (Fig. 1).

Because we use a conventional digital colour camera, this method is applicable for assessing colours in the human-visible spectrum of approximately 400–700 nm. All photographs were taken with a Nikon COOLPIX 8700 digital camera, with an effective pixel count of just under 8.0 megapixels. This is a mid-priced camera with high quality optics (Párraga, Troscianko & Tolhurst, 2002) and full control over metering and exposure. All photos in this study were taken at the 'fine' quality setting, which has a compression ratio of 1 : 4 [creating a 2–3 MB Joint Photographic Experts Group (JPEG) file per photo]. We initially saved photos in Tagged Image File Format (TIFF) (*sensu* Stevens *et al*., 2007) but, for several reasons (see below), we settled on low-compression JPEG files instead. We used manual settings for integration time (shutter speed) and lens aperture, and the white balance was set to 'daylight'. All photos were taken outdoors and used natural lighting only. Digital images consist of a matrix of microscopic photocells where colour is recorded as brightness values (in the range 0–255) of red, green, and blue (RGB). We purposely underexposed all photos (by 1–2 f-stops) to guard against 'clipping'

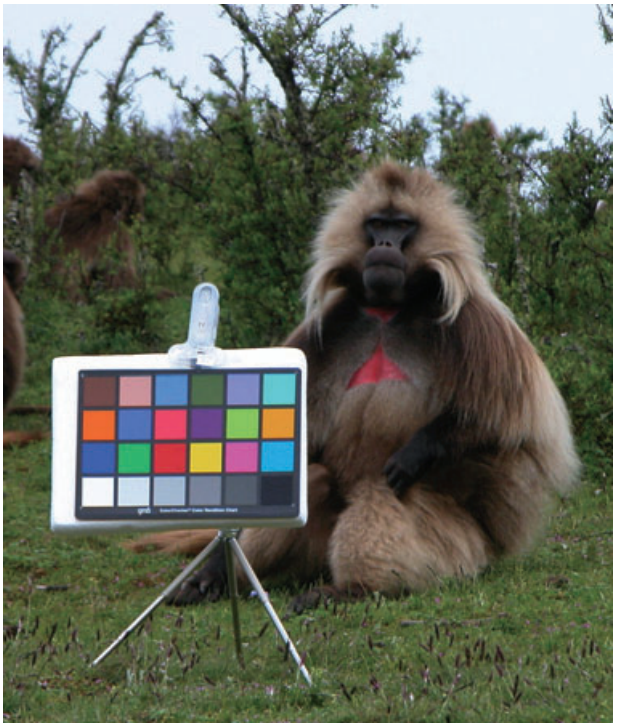

Figure 1. Setup for the adjacent method of colour assessment: an uncorrected photograph of the GretagMacBeth ColorChecker chart next to an adult male gelada (our test subject). Note that, in this photograph, the angle of the ColorChecker chart and the gelada's chest are in line with one another. Achieving such congruence is not always easy in wild subjects, which is why we also tested another method of colour assessment: the sequential method. The 'light skin' square is the second square from the left in the top row; the 'moderate red' square is the third square from the left in the second row from the top.

(Stevens *et al*., 2007), which occurs when the light levels for any of the RGB channels reach the upper limits of the camera (255). If any of the RGB values for any square was 255, then clipping probably occurred and the photo was discarded.

Photos were later analysed in Adobe Photoshop CS2 (9.0.1) using the inCamera 4.0.1 filter plug-in, a specific plug-in designed to be used with the GretagMacbeth ColorChecker chart. When using the plug-in, our light source value was set to D55 (corresponding to the CIE colour temperature of 5500). When saving photos in Adobe Photoshop, photos were saved as maximum quality JPEG files.

Method description

We tested two variations of a method for assessing colour in wild animals. The first version (the 'adjacent method') requires that the ColorChecker chart be in the same photograph as the animal whose colour is being measured. In other words, the chart must be positioned such that the coloured part of the animal can be photographed next to the chart (a feat that is easier for some species than others; Fig. 1). Because obtaining a standard in the same photo as a wild animal is not always possible, we also tested a second version of this method (the 'sequential method'), where separate photographs of the ColorChecker chart and the subject are taken in the same location and in rapid succession (within 1–2 min and under presumably identical light conditions).

After transferring all photos to a computer, we used the inCamera plug-in for Adobe Photoshop to open the photo of the ColorChecker chart. First, we manually aligned a grid (provided by the plug-in) to the 24 squares of the ColorChecker chart in the photo, thus, sampling each colour in the chart. Second, we used the 'check capture' feature of the inCamera plug-in to ensure that 'noise' (i.e. variability within each colour square) was minimal (SD < 3.0). In cases where noise is high, the inCamera manual suggests blurring the image with the standard; however, we did not need to resort to this. Third, inCamera was used to create a colour profile that adjusts the colour in the photograph to the known colour levels in each square of the ColorChecker chart. Fourth, we assigned and converted the photo to be measured (which, for the adjacent method, is the same photo) to this newly-created colour profile using the settings recommended by the inCamera plug-in. Finally, following this colour correction, we made our colour measurements by: (1) selecting the appropriate area of the photograph using the rectangular 'marquee' tool (a minimum of 200 pixels) and (2) recording the RGB levels using the histogram palette (averaged over selected pixels). In our subjects, we were primarily interested in differences in the colour red. Because the actual value in each channel is only informative relative to the values in other channels, we analysed the ratio of red to green (hereafter, the R/G ratio).

Method validation

We validated both versions of the method using two ColorChecker charts: one chart served as the 'test chart' (simulating the subject) and the other chart served as the 'control chart' (or the standard). First, we selected two of the 24 ColorChecker squares, the 'light skin' square and the 'moderate red' square, for colour validation (Fig. 1). We chose these two squares because they span the range of colour observed for male gelada chest patches. The light skin square was specifically designed by GretagMacBeth to mimic the reflectance of skin and is similar in redness to the chest patch of an immature male gelada. The moderate red square, although not as red as primary red, is

redder than any of the gelada chest patches that we measure.

To validate the adjacent method, one chart (the control chart) was placed 3–7 m in front of the other chart (the test chart). We selected this distance because it approximates the distance between the chart and our gelada subjects during our initial attempts to use this method. The charts were held approximately 0.6 m above the ground (to approximate the height of the chest patch of a seated gelada) by a mini-tripod outfitted with a large clip on an adjustable arm that allowed us to position the chart in three dimensions (Fig. 1). Both ColorChecker charts were positioned to face the same direction, but the horizontal and vertical angles were separately adjusted by 0–15° in a random fashion. As such, the two charts faced in directions that could differ by up to 30° in each dimension. In the field condition, it would be impossible to perfectly align the angle of the chart with a gelada chest patch in the same photograph. Therefore, we wanted to incorporate the error that different angles might introduce into this method. A single photograph was taken from a distance of 5 m to the closest chart. The control chart was used to create the colour profile and then colour measurements were taken (after profile conversion) from both the control chart (as a control) and the test chart.

To validate the sequential method, we also used two ColorChecker charts. First, the test chart was placed 5–10 m from the camera, facing to within 30° of the camera (both horizontally and vertically) and a photo was taken. The photographer remained in the same place and another person removed the test chart and, after waiting 1 min to simulate the time it takes our subject to move away, placed the control chart in the same spot. The photographer used the first photo to direct the person as to how they should position and align the control chart so that it resembled the test chart (this is the same procedure that we use with our gelada subjects for the sequential method). The second photo was taken 1–2 min after the first using the same exposure and metering settings. The colour profile was then created using the first picture and applied to the second picture before measuring colour in the test squares.

Linearization and equalization

We checked the reliability of our method for measuring several aspects of the test chart across different light conditions. First, using the six-step grey scale on the test chart as our set of grey reflectance standards, we tested our corrected photos for linearity (i.e. a linear relationship between each of the RGB values and the nominal reflection values across the six grey squares) and RGB equality (by definition, greys should have equal values in all three colour channels) as recommended by Stevens *et al*. (2007). Second, using the light skin and moderate red squares on the test chart, we examined the R/G ratio. However, actual RGB values for squares of the ColorChecker chart vary considerably depending on the ambient light and the colour space being used. Therefore, to create a control value for the light skin and moderate red squares (to which we could compare our test values), we calculated a mean R/G ratio for the light skin and medium red squares from the control charts under all light conditions. This represents the value to which the inCamera plug-in is attempting to adjust the ColorChecker squares for the Adobe RGB 1998 colour space set to D55 lighting. We then compared the test values with these control values under the different light conditions.

Different light conditions/times of day

To validate both versions of the method, we took ten photos (or ten pairs of photos for the sequential method) under each of four light conditions: (1) full sun (i.e. chart in direct sunlight); (2) full shade (i.e. sun is out but chart is entirely in the shade of a tree); (3) backlit (i.e. sun hitting the chart from behind); and (4) cloud (i.e. sun blocked by heavy clouds). Because sunlight varies throughout the day, we also took ten full sun pictures in the morning (within 1 h of sunrise), midday (within 1 h of noon), and evening (within 1 h of sunset). All other photos were taken at midday. We never took photos when light conditions were changing (e.g. partial clouds). Photos taken in each light condition were separated by at least 10 min and taken on at least two different days. The ColorChecker charts were moved across a horizontal arc relative to the photographer to ensure that, within each light condition, the charts had a range of orientations to the sun.

JPEG versus TIFF photos

Stevens *et al*. (2007) recommend using RAW or TIFF over JPEG because JPEG files are compressed and information is lost in the compression process, with this problem being particularly pronounced at high levels of compression (90%). However, RAW and TIFF files have several disadvantages: (1) they are extremely large (> 20 MB with the Nikon COOLPIX 8700) and thus take up a lot of storage space and (2) it takes a long time (up to 20 s) for the camera to write the image to the media storage device, which limits the ability to take photos in rapid succession. Furthermore, for our purposes, we measured a relatively large area (several cm^2) and measured mean RGB values across several hundred pixels. Therefore, the loss of detail associated with compression should not negatively impact our results. To test this, we

used the adjacent method on ten low-compression JPEG photos and ten uncompressed TIFF photos taken under identical light conditions (full sun) and compared the results. Each JPEG and TIFF pair was taken within 30 s of each other.

Visual validation

As an additional validation, two observers compared pairs of adult males whose chests could be observed simultaneously (i.e. the males were sitting adjacent to one another with chests exposed to the same light, $N = 37$). Each observer independently scored which of the two chest patches appeared 'redder'. Cases where observers agreed that the chests were too similar to differentiate ($N = 6$) were excluded from analysis. In the other 31 cases, both observers agreed on the visual comparison. For these cases, we were able to take measurement photos of each male (always within 1 h of the visual comparison) using the sequential method outlined above. After digitally assigning 'redness' to the photographs of male chest patches, we determined how many times our visual assessment agreed with the digital measurement.

Gelada colour variation across age

We then applied this method of colour assessment to wild geladas living in the Simien Mountains National Park, Ethiopia. Research was conducted in the Sankaber area, where two bands have been under intensive behavioural study since October 2005 onward and are fully habituated to human observers on foot. For the field application, we used only the sequential method because it is easier in practice, and it yielded slightly better results in the validation (see Results).

To examine chest colour variation in male geladas, we assessed colour across age categories. Because we have less than 2 years of daily observations on this population, all ages of males were necessarily estimated. Prior to this study, we placed males in age categories based on a combination of (1) published descriptions of gelada age characteristics based on dental eruption schedules and morphological traits (Dunbar & Dunbar, 1975) and (2) our observations of physical size and developmental markers compared with baboons (*Papio* spp.) of known age. We took chest photos of ten different males in each of the following five age classes: old juvenile (3.5–4.5 years), adolescent (4.5–6.5 years), young adult (6.5–8.0 years), prime adult (8.0–13.0 years), and old adult (> 13.0 years).

For each male, we selected a portion of the chest patch within the 'sample area' (Fig. 2), avoiding shadows, scratches, blemishes, and scaly skin. The sample area was selected because it is the flattest part of the chest patch (and thus easiest to align with the ColorChecker chart) and is least likely to be shaded by other parts of the body. Where possible (81% of measurements), both right and left sides of the chest patch were measured and mean RGB values were recorded for each male.

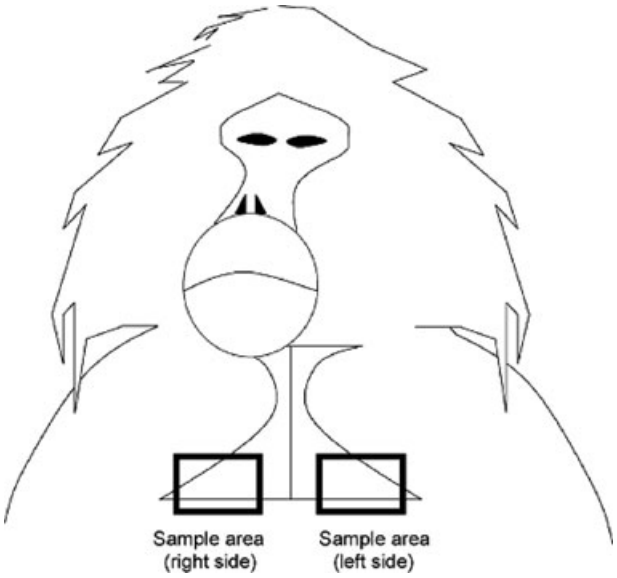

Figure 2. Sketch of a male gelada and his chest patch. Rectangular boxes indicate the 'sample area' from which we selected a smaller section to be measured (at least 200 pixels in size, but usually closer to 600 pixels, see Material and Methods).

RESULTS

Method validation

Linearization and equalization

To test for linearity, we used linear regression to assess the relationship between the measured RGB values for each square in the six-step grey scale of the test chart and the known nominal reflectance value for each square (20%, 35%, 50%, 65%, 80%, and 95% reflectance values, respectively). We examined the relationship for these six squares for all photographs taken under all light conditions (10 photos × 6 light conditions = 60 photos) for each colour channel (3 colour channels × 60 photos = 180 regressions).

Both methods showed a high degree of linearity, but the sequential method yielded slightly better results. For the adjacent method, the R^2 range was 0.974–1.000, with a mean ± SD of 0.994 ± 0.004 ($N = 180$). For the sequential method, the R^2 range was 0.989–1.000, with a mean ± SD of 0.995 ± 0.002 ($N = 180$).

To test for RGB equalization, we examined whether R = G = B in each of the six grey squares on the test chart in all photos. We measured the deviation from

Table 1. Statistical comparisons between test and control colour squares for the adjacent method across different light conditions

Adjacent method					95% confidence interval	
R/G ratio	*t*	d.f.	*P**	Mean difference	Lower	Upper
Light skin square (control value = 1.25)						
Full sun	0.28	9	0.79	0.00	–0.01	0.02
Full shade	–1.07	9	0.31	–0.01	–0.03	0.01
Backlit	–1.08	9	0.31	–0.01	–0.03	0.01
Cloud	–3.09	9	**0.01**	–0.02	–0.03	0.00
Moderate red square (control value = 2.08)						
Full sun	1.44	9	0.18	0.06	–0.04	0.16
Full shade	0.40	9	0.70	0.02	–0.08	0.12
Backlit	–2.00	9	0.08	–0.04	–0.08	0.00
Cloud	–3.06	9	**0.01**	–0.08	–0.14	–0.02

*Bold indicates significant results at $P < 0.05$.

equality by calculating the absolute value of the difference between the RGB values (i.e. R–G, R–B, G–B) for a total of 1080 comparisons (60 photos × 6 squares × 3 RGB differences).

The sequential method demonstrated a higher degree of RGB equalization than the adjacent method. Out of a maximum possible difference of 255, the absolute value of the difference for the adjacent method was in the range 0–37 (0.00%–14.51%), with a mean ± SD difference of 4.46 ± 4.65 (1.75%, $N = 1080$). The percentage of RGB values that were within 5% of each other was 94.72%. For the sequential method, the absolute value of the difference was in the range 0–13 (0.00%–5.10%), with a mean ± SD of 2.83 ± 2.46 (1.11%, $N = 1080$). The percentage of RGB values that were within 5% of each other was 99.61%.

Different light conditions

Based on measurements from the control chart across all light conditions (10 measurements × 6 light conditions = 60 values per control mean), we obtained a 'control' R/G ratio of 1.25 for the light skin square, and 2.08 for the moderate red square.

To measure the accuracy of the adjacent method, we compared test values to the control values calculated above. For both the light skin square and the moderate red square, only photos taken under cloud conditions were significantly different from the control values (Table 1). To measure the precision (repeatability) of the adjacent method, we compared values across different light conditions to determine whether they were significantly different from each other. Although R/G ratios for the light skin square across light conditions were not significantly different from each other, R/G ratios for the moderate red square across light conditions were different [analysis of variance (ANOVA), light skin: $F_{3,36} = 0.95$, $P = 0.43$, moderate red: $F_{3,36} = 3.20$, $P = 0.04$; Fig. 3A].

The sequential method was more accurate than the adjacent method. None of the R/G ratios were significantly different from control values except for R/G ratios for the moderate red square taken under backlit conditions (Table 2). In terms of precision, none of the R/G ratios for either colour square were significantly different across light conditions (ANOVA, light skin: $F_{3,36} = 0.09$, $P = 0.97$, moderate red: $F_{3,36} = 1.05$, $P = 0.38$; Fig. 3B).

Different times of day

We also tested accuracy and precision across different times of the day: morning, midday, and evening (full sun only). For the light skin square, R/G ratios across different times of day did not differ from the control value for either the adjacent or the sequential method (Table 3). For the moderate red patch, morning R/G ratios were significantly different from the control value for both methods (Table 3). To measure precision, we compared R/G ratios across the different times of the day and found no significant differences for either the adjacent method (ANOVA, light skin: $F_{2,27} = 0.52$, $P = 0.60$, moderate red: $F_{2,27} = 0.59$, $P = 0.56$) or the sequential method (ANOVA, light skin: $F_{2,27} = 0.12$, $P = 0.19$, moderate red: $F_{2,27} = 2.53$, $P = 0.10$).

JPEG versus TIFF photos

We compared the reliability of ten pairs of photos taken under identical light conditions (full sun only): one set taken as low-compression JPEG files and the other set taken as uncompressed TIFF files. In no case did the JPEG files produce different results from the TIFF files.

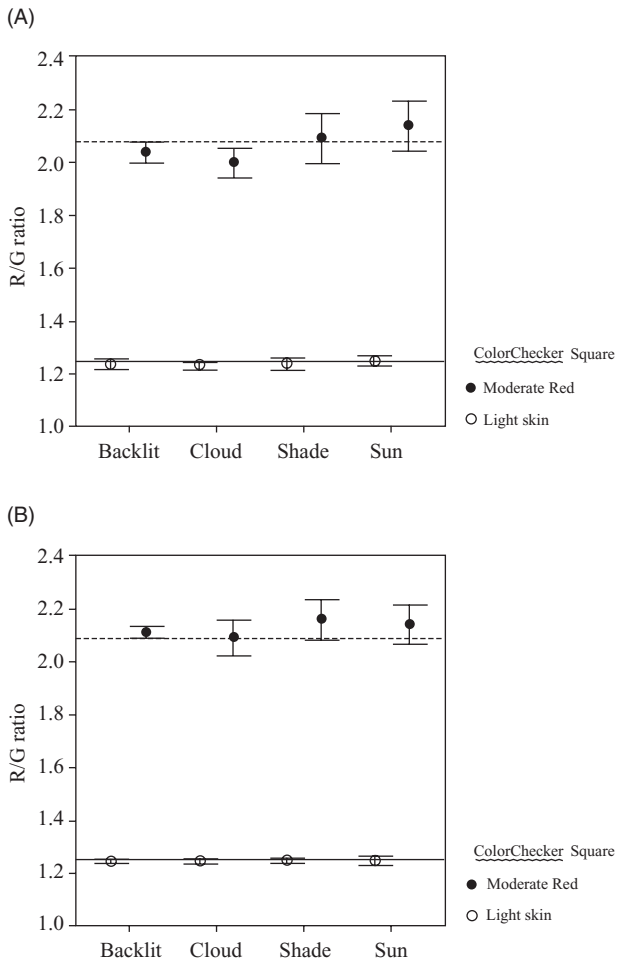

Figure 3. Adjacent (A) and sequential (B) methods of colour assessment. Mean ± SEM for the ratio of red to green (R/G) for two coloured squares of the GretagMacBeth ColorChecker chart (light skin and moderate red) measured under four different light conditions using each method. Lines represent the control R/G ratio values for the light skin (solid line, 1.25) and moderate red squares (dotted line, 2.08). Ten photographs were taken for each square per light condition.

Linearity: For the JPEG photos, R^2 was in the range 0.990–0.997, with a mean ± SD of 0.994 ± 0.004 ($N = 30$). For the TIFF photos R^2 was in the range 0.984–0.998, with a mean ± SD of 0.993 ± 0.003 ($N = 30$).

RGB equality: For the JPEG photos, the absolute value of the difference was in the range 0–11 (0.0%–4.3%), with a mean ± SD of 2.72 ± 2.22 (1.07%, $N = 180$). For the TIFF photos, the absolute value of the difference was in the range 0–9 (0.0–3.5%), with a mean ± SD of 2.58 ± 2.14 (1.01%, $N = 180$). RGB values for all JPEG and TIFF photos were within 5% of each other.

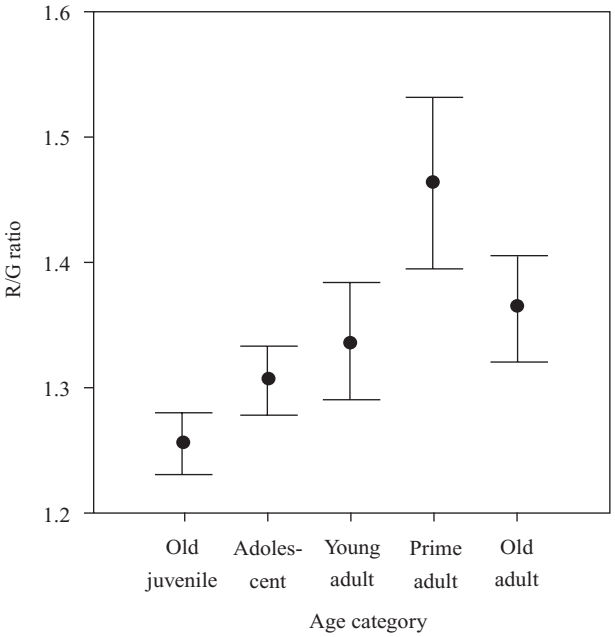

Figure 4. Mean ± SEM for the ratio of red to green (R/G) for gelada males from five different age categories. Ten photographs from ten different males were taken for each category.

Accuracy: When compared with control values for the light skin and moderate red squares, the two types of photos demonstrated comparable levels of accuracy. Neither the light skin square (JPEG: $t = -1.81$, d.f. = 9, $P = 0.104$, mean difference = 0.007; TIFF: $t = -0.85$, d.f. = 9, $P = 0.416$, mean difference = 0.004), nor the moderate red square (JPEG: $t = -0.54$, d.f. = 9, $P = 0.600$, mean difference = 0.005; TIFF: $t = -1.74$, d.f. = 9, $P = 0.117$, mean difference = 0.017) was significantly different from control values.

Visual validation

In 31 cases where we were able to both visually compare and collect measurement photographs on pairs of males, the male that observers visually scored as the 'redder' of the two males was also the male with the higher R/G ratio as measured digitally. This success rate is significantly above chance (Binomial, $P < 0.001$).

Gelada colour variation across age

We found significant differences in R/G ratios across the five age categories (ANOVA, $F_{4,45} = 11.62$, $P < 0.001$), with prime age males exhibiting significantly redder chests than all other males (Tukey's post-hoc test: $P < 0.05$), and old juveniles exhibiting significantly paler chests than old males (Tukey's post-hoc test: $P = 0.012$; Fig. 4).

Table 2. Statistical comparisons between test and control colour squares for the sequential method across different light conditions

Sequential method					95% confidence interval	
R/G ratio	*t*	d.f.	*P**	Mean difference	Lower	Upper
Light skin square (control value = 1.25)						
Full sun	0.11	9	0.91	0.00	–0.02	0.02
Full shade	–0.42	9	0.68	0.00	–0.01	0.01
Backlit	–1.24	9	0.25	0.00	–0.01	0.00
Cloud	–0.48	9	0.64	0.00	–0.01	0.01
Moderate red square (control value = 2.08)						
Full sun	1.82	9	0.10	0.06	–0.01	0.14
Full shade	2.13	9	0.06	0.08	0.00	0.16
Backlit	3.22	9	**0.01**	0.03	0.01	0.05
Cloud	0.26	9	0.80	0.01	–0.06	0.08

*Bold indicates significant results at $P < 0.05$.

Table 3. Statistical comparisons between test and control colour squares for both methods across different times of day

					95% confidence interval	
R/G ratio	*t*	d.f.	*P**	Mean difference	Lower	Upper
Time of day: adjacent method						
Light skin square (control value = 1.25)						
Morning	1.66	9	0.13	0.01	0.00	0.02
Midday	0.28	9	0.79	0.00	–0.01	0.02
Evening	1.29	9	0.23	0.00	0.00	0.01
Moderate red square (control value = 2.08)						
Morning	2.37	9	**0.04**	0.13	0.01	0.26
Midday	1.44	9	0.18	0.06	–0.04	0.16
Evening	1.17	9	0.27	0.06	–0.06	0.18
Time of day: sequential method						
Light skin square (control value = 1.25)						
Morning	–0.56	9	0.59	0.00	–0.01	0.00
Midday	0.11	9	0.91	0.00	–0.02	0.02
Evening	0.44	9	0.67	0.00	–0.01	0.01
Moderate red square (control value = 2.08)						
Morning	2.29	9	**0.05**	0.06	0.00	0.12
Midday	1.82	9	0.10	0.06	–0.01	0.14
Evening	–0.72	9	0.49	–0.02	–0.08	0.04

*Bold indicates significant results at $P < 0.05$.

DISCUSSION

Both the adjacent and the sequential method of colour assessment demonstrated accuracy and precision, even across widely varying light conditions. With few exceptions, colour measurements taken on the test chart did not differ from control values, and measurements taken across different light conditions and different times of day did not differ from each other. Additionally, measurements taken on the light skin square (designed to simulate skin reflectance) proved more reliable than measurements taken on the moderate red patch, demonstrating the utility of this method for measuring colour from natural surfaces such as skin.

The sequential method proved slightly more reliable than the adjacent method, exhibiting a higher degree of linearity and RGB equality, and proving

more accurate across different light conditions. The sequential method has the added advantage of being easier to use in the field. However, in situations where the ColorChecker chart can easily be positioned near study animals and when the angle between the subject and the chart can be precisely matched (particularly in situations where subjects are more sedentary and/or less likely to move within a minute), the adjacent method would still be effective. Furthermore, under conditions where natural light is rapidly changing (i.e. thin cloud cover), the adjacent method would be preferable to the sequential method.

For colour assessment on relatively large areas, we found that taking and storing our photographs as low-compression JPEG images produced results that were indistinguishable from using TIFF images. On our wild subjects, JPEG photographs were easier to acquire because we could take several photographs in rapid succession and use the one that best displayed the chest patch. The long period of time required to write the TIFF images to the storage media precluded this ability. Therefore, we used low-compression JPEG images throughout this study. However, in cases where smaller, more detailed comparisons are necessary (e.g. patterns on the wing of a moth), it is preferable to use uncompressed images (TIFF or RAW) as recommended by Stevens *et al*. (2007).

The sequential method of colour assessment closely matched visual assessments of male chest redness. Because the visual system is highly conserved among Old World primates (Bowmaker *et al*., 1991; Deeb *et al*., 1994; Waitt & Buchanan-Smith, 2006), it is likely that our visual assessments are similar to the assessments geladas themselves would make. The very close correspondence of our visual assessments with digital measurements indicates that this method is able to detect and accurately score biologically meaningful variation in 'redness'.

The utility of this method, like all methods using digital cameras, is limited to colour variation within the spectral range of the camera and is not useful for variation in the ultraviolet range. Furthermore, with the exception of tri-chromatic primates, the RGB values generated by this method are better suited for quantifying variation in the production of colour than for measuring how that colour might be perceived. Because colour is a function of the perceptual system of the intended receiver, Bennett, Cuthill & Norris (1994) stress the importance of quantifying colour as it is perceived by the animal receiving the light. Thus, mapping RGB values to differing spectral sensitivities for other species is recommended whenever possible (Stevens *et al*., 2007). Furthermore, even in species whose spectral range does not match that of humans, it is possible that the variation of interest falls within the human range and could, as an initial step, be measured using this method. Finally, the RGB values generated by this method can be converted to colour quantification systems based on human perception that define colour in terms of hue (the shade of the colour), chroma (saturation, or the intensity of the colour), and brightness (how light the colour is) for comparison with measurements based on such systems (e.g. visual matches to Munsell colour charts).

In our gelada subjects, we were able to use this method to detect significant differences in R/G ratios across age categories. Males most likely to have reproductive access to females (prime adults) exhibited the reddest chests, whereas the least mature males (old juveniles) had the palest chests. The pattern of chest colour in relation to age (Fig. 4) suggests that chest redness may be a secondary sexual trait associated with sexual maturation, reproduction, and/or aggression associated with defending reproductive access to females. If so, a male's chest colour might also serve as a sexually selected signal to other geladas (Andersson, 1994) and this hypothesis warrants further exploration.

Because we are interested in gelada chest colour, we focused our validation on the red squares of the ColorChecker chart. However, the high degree of linearity and equality in the grey squares suggests that the method does a good job of standardizing all colours across different light conditions. Thus, the method should be useful for researchers interested in any part of the visible colour spectrum. In addition to being reliable, this method is inexpensive, non-invasive, and has the potential to be used on wild or captive subjects.

ACKNOWLEDGEMENTS

We are grateful to the Ethiopian Wildlife Conservation Department, the Amhara National Regional State Parks Development and Protection Authority, and the wardens and staff of the Simien Mountains National Park for granting us the permission to conduct this research. We are extremely grateful to H. Gelaye for his help with data collection in the field and D. Kitchen for her comments on an earlier draft of the manuscript. Funding was provided by the Wildlife Conservation Society (SSF Grant #67250), the National Geographic Society (Grant #8100-06), and the University of Michigan. This research was approved by the University Committee on Use and Care of Animals (UCUCA) at the University of Michigan and adhered to the laws and guidelines of Ethiopia.

REFERENCES

Andersson M. 1994. *Sexual selection*. Princeton, NJ: Princeton University Press.

Bennett A, Cuthill I, Norris K. 1994. Sexual selection and the mismeasure of color. *American Naturalist* **144:** 848–860.

Bourne G, Breden F, Allen T. 2003. Females prefer carotenoid colored males as mates in the pentamorphic live bearing fish, *Poecilia parae*. *Naturwissenschaften* **90:** 402–405.

Bower J. 1958. Experimental studies of mimicry in some North American butterflies. 1. The monarch, *Danaus plexippus*, and viceroy, *Limenitis archippus*. *Evolution* **12:** 3–47.

Bowmaker J, Astell S, Hunt D, Mollon J. 1991. Photosensitive and photostable pigments in the retinae of Old World monkeys. *Journal of Experimental Biology* **156:** 1–19.

Cooper V, Hosey G. 2003. Sexual dichromatism and female preference in *Eulemur fulvus* subspecies. *International Journal of Primatology* **24:** 1177–1188.

Darwin C. 1859. *On the origin of species*. London: J. Murray.

Darwin C. 1871. *The descent of man and selection in relation to sex*. London: J. Murray.

Deeb S, Jorgensen A, Battisti L, Iwasaki L, Motulsky A. 1994. Sequence divergence of the red and green visual pigments in great apes and humans. *Proceedings of the National Academy of Sciences of the United States of America* **91:** 7262–7266.

Dunbar RIM. 1984. *Reproductive decisions: an economic analysis of gelada baboon social strategies*. Princeton, NJ: Princeton University Press.

Dunbar RIM, Dunbar P. 1975. Social dynamics of gelada baboons. In: Kuhn H, Luckett WP, Noback CR, Schultz AH, Starck D, Szalay FS, eds. *Contributions to primatology*. Basel: S. Karger.

Endler J. 1984. Progressive background matching in moths, and a quantitative measure of crypsis. *Biological Journal of the Linnean Society* **22:** 187–231.

Endler J. 1986. Defense against predation. In: Feder M, Lauder G, eds. *Predator-prey relationships: perspectives and approaches from the study of lower vertebrates*. Chicago, IL: University of Chicago Press, 109–134.

Endler J. 1990. On the measurement and classification of colour in studies of animal colour patterns. *Biological Journal of the Linnean Society* **41:** 315–352.

Evans M, Morris K. 1996. The importance of carotenoids in signaling during aggressive interactions between male firemouth cichlids (*Cichlasoma meeki*). *Behavioral Ecology* **7:** 1–6.

Gerald MS. 2001. Primate colour predicts social status and aggressive outcome. *Animal Behaviour* **61:** 559–566.

Gerald M, Bernstein J, Hinkson R, Fosbury R. 2001. Formal method for objective assessment of primate color. *American Journal of Primatology* **53:** 79–85.

Goda M, Fujii R. 1998. The blue coloration of the common surgeonfish, *Paracanthurus hepatus* – II. Color revelation and color changes. *Zoological Science* **15:** 323–333.

Kilner R. 1997. Mouth colour is a reliable signal of need in begging canary nestlings. *Proceedings of the Royal Society of London Series B, Biological Sciences* **264:** 963–968.

Litvaitis J. 1991. Habitat use by snowshoe hares, *Lepus americanus*, in relation to pelage color. *Canadian Field Naturalist* **105:** 275–277.

Losey G. 2003. Crypsis and communication functions of UV-visible coloration in two coral reef damselfish, *Dascyllus aruanus* and *D. reticulatus*. *Animal Behaviour* **66:** 299–307.

Malcolm S. 1990. Mimicry: status of a classical evolutionary paradigm. *Trends in Ecology and Evolution* **5:** 57–62.

Munsell AH. 1976. *Munsell Book of Color – Matte Finish Collection*. Baltimore, MD: Munsell Color.

Párraga C, Troscianko T, Tolhurst D. 2002. Spatiochromatic properties of natural images and human vision. *Current Biology* **12:** 483–487.

Stevens M, Párraga C, Cuthill I, Partridge J, Troscianko T. 2007. Using digital photography to study animal coloration. *Biological Journal of the Linnean Society* **90:** 211–237.

Villafuerte R, Negro J. 1998. Digital imaging for colour measurement in ecological research. *Ecology Letters* **1:** 151–154.

Waitt C, Buchanan-Smith H. 2006. Perceptual considerations in the use of colored photographic and video stimuli to study nonhuman primate color. *American Journal of Primatology* **68:** 1054–1067.

Waitt C, Little AC, Wolfensohn S, Honess P, Brown AP, Buchanan-Smith HM, Perrett DI. 2003. Evidence from rhesus macaques suggest that male coloration plays a role in female primate mate choice. *Proceedings of the Royal Society of London Series B, Biological Sciences* **270:** S144–S146.

Westmoreland D, Kiltie R. 1996. Egg crypsis and clutch survival in three species of blackbirds (Icteridae). *Biological Journal of the Linnean Society* **58:** 159–172.

Zuk M, Decruyenaere J. 1994. Measuring individual variation in colour: a comparison of two techniques. *Biological Journal of the Linnean Society* **53:** 165–173.