When to measure plumage reflectance: a lesson from Collared Flycatchers *Ficedula albicollis*

GERGELY HEGYI,[1]* (iD) MIKLÓS LACZI,[1,2] NÓRA BOROSS,[1,3] MÓNIKA JABLONSZKY,[1] DÓRA KÖTÉL,[1]
KATALIN KRENHARDT,[1] GÁBOR MARKÓ,[1,4] GERGELY NAGY,[1] BALÁZS ROSIVALL,[1] ESZTER SZÁSZ,[1]
LÁSZLÓ ZSOLT GARAMSZEGI[5,6] & JÁNOS TÖRÖK[1,7]

[1]*Behavioural Ecology Group, Department of Systematic Zoology and Ecology, Eötvös Loránd University, Pázmány
Péter sétány 1/C, Budapest H-1117, Hungary*
[2]*The Barn Owl Foundation, Temesvári út 8, Orosztony H-8744, Hungary*
[3]*Balaton Limnological Institute, Centre for Ecological Research, Hungarian Academy of Sciences, Klebelsberg Kuno
utca 3, Tihany H-8237, Hungary*
[4]*Department of Plant Pathology, Szent István University, Ménesi út 44, Budapest H-1118, Hungary*
[5]*Department of Evolutionary Ecology, Estación Biológica de Doñana-CSIC, c/Americo Vespucio s/n, Seville 41092,
Spain*
[6]*MTA-ELTE Theoretical Biology and Evolutionary Ecology Research Group, Department of Plant Systematics,
Ecology and Theoretical Biology, Eötvös Loránd University, Pázmány Péter sétány 1/C, Budapest H-1117, Hungary*
[7]*Ecology Research Group of the Hungarian Academy of Sciences, Pázmány Péter sétány 1/C, Budapest H-1117,
Hungary*

Sexually selected colour traits of bird plumage are widely studied. Although the plumage is replaced only at one or two yearly moults, plumage colour has long been shown to change between moults. Nevertheless, most studies measure colour weeks to months after the courtship period, typically at nestling rearing, and it is unclear whether these measurements yield relevant data concerning the primary process of sexual selection. Here we analyse repeated spectrometric data taken from male Collared Flycatchers during social courtship and nestling rearing. We show that some spectral traits are not correlated between the two measurements and that within-individual correlation declines significantly with the likely exposure of the plumage area to damage and soiling. There is an overall decline in spectral trait exaggeration during breeding, but trait decline is not closely related to measurement latency, especially not in the damage-exposed areas. Finally, sexual selection estimates differ depending on whether they are derived from spectra measured during courtship or during nestling rearing. These results suggest that, contrary to current practice, measurements of plumage reflectance should be made during the primary period of sexual signalling. Spectral trait decline during breeding could also be studied as a possible signal for mates and neighbours.

Keywords: brightness, chroma, sexual selection, signal deterioration, spectrometry.

A high proportion of sexual selection research focuses on plumage colour. The only time when large plumage tracts are systematically replaced is moult. Plumage colour is therefore typically measured only once (most often in one stage of breeding) and this expression state is attributed to the individual for the whole season (breeding

or non-breeding). However, seasonal within-individual change in colour without moult has been found repeatedly (McGraw & Hill 2004, Hasegawa *et al.* 2008, Surmacki *et al.* 2011, Evans *et al.* 2012). Importantly, these findings refer to plumage colour regarded as pigment- or structure-based, and not 'cosmetic' colour (Montgomerie 2006). These findings should have stimulated methodological change. Researchers should measure colour at the time when it is

*Corresponding author.
Email: gehegyi@caesar.elte.hu

principally used in sexual selection (i.e. monogamous courtship period) or social selection (winter territory establishment period, winter food scarcity and competition period). We conducted a literature survey of measurement time in plumage reflectance studies and found no sign of such methodological change (see Supporting Information S1 for details). The greatest problem arises if reflectance changes non-predictably or non-repeatably with time (e.g. Surmacki *et al.* 2011). If so, reflectance data taken long after courtship (e.g. during breeding) will be of limited relevance to assessment of signal meaning and sexual selection. No study has yet directly tested the repeatability of individual plumage reflectance between courtship and breeding, although robust, directional, within-individual colour shifts seem to occur in this period (Örnborg *et al.* 2002, McGraw & Hill 2004, Delhey *et al.* 2006, Hasegawa *et al.* 2008).

Here we aim to assess the change of plumage reflectance between courtship and nestling rearing, the within-individual correlation of reflectance traits across this period (a measure analogous to repeatability) and some practical consequences. We use repeated spectrometric data from individual Collared Flycatcher *Ficedula albicollis* males and adopt a meta-analytical approach to synthesize our results across reflectance traits. We first assess the correlation of spectral traits between the two measurements, the directional deterioration of spectral trait expression and the correlation between trait change and time between the two measurements. We also estimate the change of the visual stimulus due to the directional reflectance deterioration. We then calculate sexual selection estimates for spectral trait values measured at courtship and those measured at nestling rearing. Finally, we assess the exposure of the given plumage area to damage (structural deterioration and soiling) as a possible predictor of spectral trait dynamics and sexual selection estimates. We predict low and variable within-individual correlation and a directional decline in spectral trait expression during the breeding season. We further predict differences between sexual selection estimates based on early and late spectral data. Specifically, we predict that signal deterioration blurs systematic patterns of sexual selection by the time of nestling rearing. Finally, if damage influences plumage reflectance, we predict significant effects of damage exposure on spectral trait change and within-individual correlation, but less strong effects on sexual selection estimates.

METHODS

We conducted this study in our long-term study plots in the Pilis-Visegrádi Mountains, Hungary (47°42′N, 19°01′E), between 2014 and 2017. We monitored male courtship daily and captured males at their first appearance on a territory (defined as firm ownership of a nestbox). We regularly checked the progress of breeding in the nestboxes, determined laying date of the first egg, and caught the parents at 8–10 days of nestling age. Male age (yearling or adult) was determined based on the colour of the remiges and wing patch size (Török *et al.* 2003). Sexual selection on male attributes was estimated as time to pairing, i.e. the latency between date of capture at courtship and the first egg date of the brood. This measure has previously given interpretable results for song attributes, behavioural traits and white patch sizes (Garamszegi *et al.* 2006, 2008, Hegyi *et al.* 2010). Time to pairing is relevant to fitness in our population where recruit production declines steeply with breeding date in the second half of the breeding season, when most of the broods are initiated (Herényi *et al.* 2014).

Spectra of males when first caught during courtship and when caught later during nestling rearing were recorded using the same equipment in the same way. Measurements were made in five plumage areas: forehead patch (white), crown (black), wing patch (white), wing coverts (black in adults, brown in yearlings) and breast (white). Spectral data were collected using a USB2000 spectrometer equipped with a DH-2000 deuterium-halogen light source and a QR400-7-SR-BX fibre optic sensor (Ocean Optics Europe). We fixed a matt black plastic tube around the end of the sensor to exclude ambient light and to standardize the measurement distance from the plumage (3 mm). The sensor was orientated perpendicular to the plumage surface. Two consecutive readings were taken of each plumage area while removing the sensor from the plumage between them. Calibrating with a WS-1-SS white standard (Ocean Optics Europe) and a black standard (no incoming light to the sensor), percentage intensity at each measured wavelength value was calculated and stored in OOIBASE32 software (Ocean Optics Europe) using the following formula: $((R_{sample} - R_{dark})/$

($R_{\text{white}} - R_{\text{dark}}$)) × 100. Dark and white signals were re-measured regularly to prevent signal drift. Previous exploratory analyses indicated that the main axes of spectral variation are the same in all plumage areas irrespective of pigmentation (Laczi *et al.* 2011), so we used two descriptors of reflectance variation throughout: overall (320–700 nm) mean reflectance (brightness, hereafter BRI) and relative ultraviolet (320–400 nm) reflectance (UV chroma, hereafter UVC). Previous analyses of mating patterns suggested that all measured plumage areas operate as one integrated ornamental trait in terms of both BRI and UVC (Laczi *et al.* 2011). Here we are interested in differences in repeatability and reflectance change between different plumage areas, but interpret their sexual selection estimates in an integrated way. We achieve this by means of a meta-analytical approach.

A total of 139 captures of courting males were made during the 4 years, and 85 of these captures were also repeated during nestling rearing. For logistical reasons, plumage spectra from both courtship and breeding could be obtained at only 49 of the 85 double captures, of which eight data points were excluded as pseudoreplicates of birds measured twice in previous years, thereby leaving a final sample size of 41 males for the present analysis. Among these, eight polygynous males that were caught and measured while courting at their secondary nestboxes (date of primary courtship unknown) were omitted from the analyses of time to pairing but not the other analyses. Latency between the two measurements of the 41 birds ranged from 21 to 61 days and its effect was explicitly estimated in the present analysis. Importantly, there is little evidence that spectral trait expression at courtship affected the probability of measuring the bird for a second time at nestling rearing (results not shown), which implies that our sample of recaptured males is not systematically biased with respect to reflectance at first measurement.

In this paper, we analysed the spectral data by pooling yearling and adult males after the necessary standardizations (see below). We pooled the two age categories because despite age-dependence in the absolute value of some spectral measures, the focus of our study, the within-individual change of reflectance, was not significantly age-dependent for any of the spectral measures we use here (repeated measures general linear models, details not shown). Statistical analyses of the spectral measurements (five BRI and five UVC measures from each individual during courtship and nestling rearing) were carried out in three steps: estimation of correlations, meta-analysis and meta-regression. First, for each spectral trait ($n = 10$), we calculated effect sizes of (1) within-individual correlation, (2) overall within-individual change, (3) effect of latency between measurements on trait change and (4) relationships of early (courtship) and late (nestling rearing) spectral trait expression with time to pairing. Within-individual correlation was estimated using Pearson correlation of spectral variables after year- and age-standardization where necessary (Table S1). We chose this correlation approach because although negative correlations (when the initially most conspicuous individuals become the least conspicuous) occur and are meaningful in our case, they are not interpretable in standard repeatability calculations that work with variance components. If we assign the spectral traits with a negative within-individual correlation of zero, the resulting values are perfectly ($r > 0.99$, details not shown) correlated and have practically identical values, with standard repeatabilities calculated from variance components with the rptGaussian function of the rptR package in R (Nakagawa & Schielzeth 2011). In this paper, therefore, the within-individual correlation can safely be interpreted as a signed value analogous to repeatability. Measurement repeatabilities of spectral traits for the present data are positive and high (Pearson $r = 0.507$–0.868) and do not correlate with the across-season within-individual correlation we use here (details not shown). We further calculated paired *t*-tests for systematic within-individual change based on raw uncorrected data, and converted the results to Pearson r (Dunlap *et al.* 1996). We also calculated the correlation between measurement latency (log-transformed to ensure normality) and within-individual change for each spectral trait. Finally, we correlated standardized spectral variables (from the courtship or the nestling stage) with time to pairing (log-transformed number of days between first capture and first egg date).

As a second step, we meta-analysed Fisher *Z*-transformed effect sizes across the 10 spectral measurements for within-individual correlation, within-individual change, measurement latency effects, and correlations with time to pairing using the metafor procedure in R (Viechtbauer 2010). For the meta-analysis, we converted the

sign of all effect sizes except that of within-individual correlation to take into account the different directions of signal exaggeration; that is, a greater signal exaggeration implies higher UVC in all areas, higher BRI in white areas, but lower BRI in the melanized areas. In the converted effect sizes, positive values meant positive within-individual change in trait expression, positive measurement latency effect on trait expression (increasing positive change in spectral trait exaggeration with greater time between measurements) and facilitation of pairing by high trait expression (shorter pairing latencies for greater values of spectral trait exaggeration). Using the converted effect sizes, we assessed whether effect sizes differed between spectral trait types (BRI or UVC). We also quantified the heterogeneity of effect sizes (i.e. the degree to which they can be regarded as coming from different 'populations' of effect sizes) and estimated mean effect size across all spectral traits. As a third step, we also assessed how the four effect size measures depended on the exposure of the given plumage area to potential damage (hereafter 'damage exposure'), used as a covariate in the meta-analysis. Damage exposure was scored as 1 for the forehead patch, as this area hardly touches anything except very large food items. It was scored as 2 for the upward-facing plumage areas (crown, wing patch, wing coverts), as these do make some contact with vegetation and the nestbox entrance. Finally, the downward-facing breast received a score of 3, as it is in frequent contact with many objects including the nest and the ground at feeding and drinking and may undergo the strongest abrasion and soiling. We used this low-resolution variable (three values) as a covariate and not as a factor, as hypothetical intermediate (non-integer) values would be meaningful and because we predicted a directional effect of this variable on our effect sizes.

To test whether the reflectance change we detect is relevant to colour signalling, we estimated the visual discriminability of mean reflectance spectra of five plumage areas of the same 41 males using the vismodel and coldist functions of the R package pavo (Maia *et al.* 2013). The scores (chromatic contrast dS and achromatic contrast dL) are in 'just noticeable difference' (JND) units and therefore values greater than one indicate discriminability. Details of the visual system of the study species are unknown, so the models had to be built based on parameters of other species and therefore the estimates must be treated with caution. We used cone sensitivity (single and double cones) and ocular media transmittance data from the Eurasian Blue Tit *Cyanistes caeruleus* (a built-in setting of the R package pavo), relative cone density data from the Blue Tit (based on Hart *et al.* 2000), a green foliage background and a Weber fraction of 0.05. We present estimates using two different ambient light settings (a built-in setting of pavo), but the results are practically the same.

RESULTS

Within-individual correlation, overall trait change and measurement latency patterns

All meta-analytical results are presented in Table 1. When analysed across all traits, within-individual correlation was marginally non-significantly positive (mean = 0.161; 95% confidence interval (CI) lower = −0.013; 95% CI upper = 0.336) and did not differ significantly between BRI and UVC, although it showed highly significant heterogeneity. Within-individual change between courtship and nestling rearing was highly significant and in the direction of signal deterioration (mean = −0.197; CI lower = −0.330; CI upper = −0.064), with marginally non-significant heterogeneity, and it did not differ between BRI and UVC. Table 2 shows raw means and standard deviations of BRI and UVC values from our males at courtship and nestling rearing. The relationship between spectral trait change and latency between measurements was overall non-significant (mean = 0.084; CI lower = −0.061; CI upper = 0.229), with a significant heterogeneity and no difference between BRI and UVC.

Time to pairing

When looking at spectral data obtained at courtship, the across-trait estimate of effects on time to pairing was marginally non-significant in the predicted direction (Fig. 1, see all details in Table 1). In contrast, for spectral data measured at nestling rearing, the (apparent) time to pairing effects gave a near-zero combined value (Fig. 1). Neither effect showed significant heterogeneity or significant difference between BRI and UVC (Table 1).

Table 1. Meta-analytical results across spectral traits when comparing reflectance data at courtship vs. nestling rearing.

	Difference from zero (Z)	Heterogeneity (Q)[a]	BRI vs. UVC (Z)	Exposure effect (Z)	Residual heter. (Q)[b]
Within-individual correlation	1.813[†]	27.085**	−0.760	−3.788***	9.694
Directional change of trait exaggeration	−2.901**	15.712	−0.271	1.601	NA
Effect of measurement latency on change	1.137	18.193*	−0.037	2.501*	10.210
Time to pairing with courtship spectra	1.824[†]	12.107	1.017	−0.792	NA
Time to pairing with nestling rearing spectra	0.516	9.921	−0.734	−1.482	NA

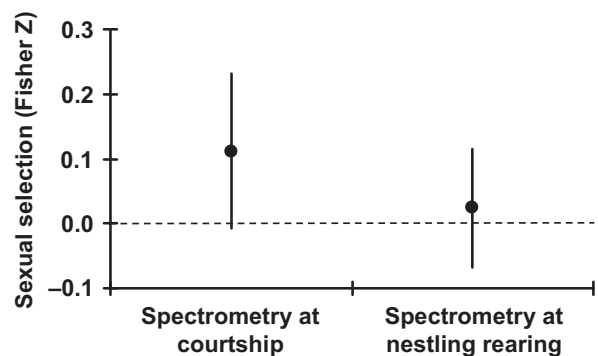[a]df = 9. [b]df = 8. [†]$P < 0.07$. *$P < 0.05$. **$P < 0.01$. ***$P < 0.001$.

Table 2. Mean (SD) plumage reflectance traits for each measured plumage area of the same 41 male Collared Flycatchers during courtship and nestling rearing, and effect sizes for within-individual spectral trait change (Pearson r). The values only serve illustrative purposes as they result from the pooling of yearling and older males, which differ in some reflectance traits.

Plumage area	Trait	Courtship Mean	sd	Nestling rearing Mean	sd	Pearson r
Forehead patch	BRI	44.55	6.37	38.92	7.01	−0.387***
Crown	BRI	2.23	0.55	2.52	0.51	0.265**
Wing patch	BRI	32.66	7.31	30.79	6.43	−0.135*
Wing covert	BRI	5.34	1.92	4.79	1.61	−0.152*
Breast	BRI	39.77	7.04	36.79	6.17	−0.220*
Forehead patch	UVC	0.72	0.07	0.62	0.08	−0.532***
Crown	UVC	0.98	0.10	0.98	0.05	0.039
Wing patch	UVC	0.80	0.04	0.77	0.06	−0.285*
Wing covert	UVC	0.92	0.12	0.91	0.10	−0.089
Breast	UVC	0.81	0.08	0.79	0.07	−0.143

*$P < 0.05$. **$P < 0.01$. ***$P < 0.001$.

Effect of damage exposure

The degree of exposure to damage was highly negatively related to within-individual correlation between courtship and nestling rearing (see all details in Table 1). That is, areas potentially exposed to more damage exhibited lower spectral trait repeatability (Fig. 2a). Moreover, there was no significant residual heterogeneity in repeatability after the inclusion of exposure in the model. Exposure had no significant effect on within-individual change in spectral traits. However, the correlation of change with measurement latency was
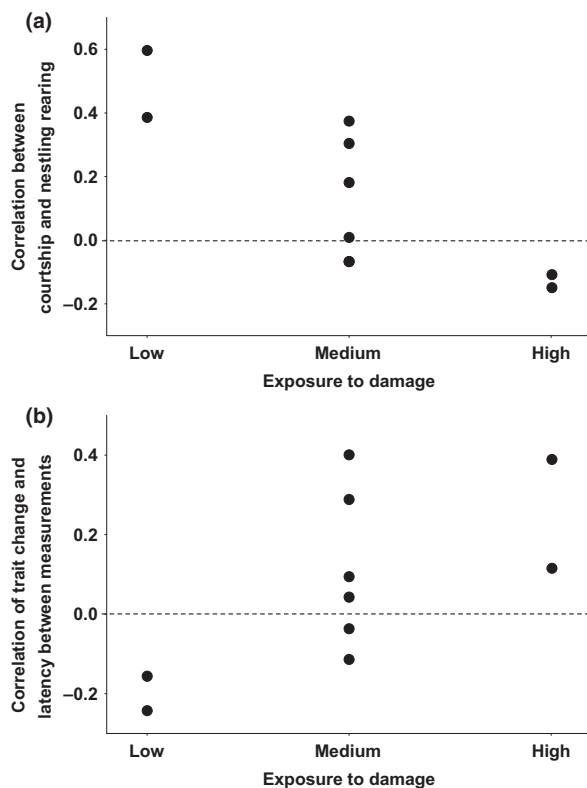

Figure 1. Sexual selection estimates on plumage reflectance using spectral trait values from courtship vs. nestling rearing (mean with 95% CI). Sexual selection was estimated using time between territory settlement and pairing. The dashed line marks zero correlation.

significantly positively related to damage exposure. Areas exposed to more damage frequently showed near-zero or positive relationships, whereas the expected moderate negative relationships dominated in less exposed areas (Fig. 2b). Again, there was no significant residual heterogeneity after the inclusion of exposure in the model. Finally, exposure had no significant effects on the effect sizes of correlations between spectral traits and time between territory settlement and pairing at either the courtship or the nestling rearing stage.

Visual modelling

Discriminability values for each plumage area are listed in Table 3. The mean spectra of the same males during courtship and nestling rearing were estimated to be perceptibly different for three of

Figure 2. The effect of potential exposure of the plumage area to damage on (a) within-individual correlation in spectral traits and (b) measurement latency effects. The dashed lines mark zero correlation.

the five measured plumage areas (forehead, crown and wing coverts).

DISCUSSION

In this study we found that the overall within-individual correlation estimate of spectral trait values between the courtship and the nestling rearing phases of Collared Flycatcher males was positive but not significant, and there was high heterogeneity in within-individual correlation among spectral traits. There are spectral traits for which measurements at nestling rearing give values that have essentially nothing to do with the earlier reflectance of the bird at courtship (see raw correlation values in Fig. 2a). This is consistent with a previous study of Eastern Bluebirds *Sialia sialis* that focused on reflectance at two consecutive breeding bouts within a season and found no significant repeatability between them (Surmacki *et al.* 2011). Moreover, as expected (e.g. Willoughby

Table 3. Visual discriminability values of mean reflectance spectra of five plumage areas of the same 41 Collared Flycatcher males, estimated in two different light environments.

Area	D65 daylight		Forest shade	
	dS	dL	dS	dL
Forehead	1.694	1.128	1.762	1.118
Crown	0.257	1.148	0.246	1.152
Wing patch	0.400	0.538	0.429	0.532
Wing coverts	0.440	1.117	0.412	1.112
Breast	0.220	0.718	0.259	0.715

dL, achromatic contrast; dS, chromatic contrast.

et al. 2002), the potential exposure of the given plumage area to damage very strongly predicted seasonal within-individual correlation, with little or no correlation in highly vulnerable areas.

When taking into account the probable direction of sexual selection on the given spectral trait, we found an almost uniform decline in spectral trait expression from courtship to nestling rearing. When adopting an avian visual model approach (Maia *et al.* 2013), mean courtship and nestling rearing spectra of the same males were estimated to be visibly different for three of the five plumage areas. Previous within-individual and cross-sectional studies detected various directions of colour change after moult, from decline (Avilés *et al.* 2008) to non-linear change (Delhey *et al.* 2010) and marked increase (Enbody *et al.* 2017). In our population, length of time between measurements did not have the expected negative effect on the spectral trait change. Despite the overall decline, its magnitude did not systematically increase with latency between measurements, and this was especially true for damage-exposed plumage areas. Finally, the effects of spectral measures during courtship on a sexual selection estimate (time between territory settlement and pairing) were marginal but in the predicted direction (quicker pairing of individuals with more exaggerated spectral trait values), whereas spectral data from the nestling period gave estimates scattered around zero and therefore no indication of sexual selection.

These results suggest that conducting spectral measurements at nestling rearing is a highly uncertain way to estimate values of spectral traits during the primary sexual selection event (e.g. courtship and pairing). This uncertainty is particularly characteristic of plumage areas exposed to damage and

soiling, and probably arises from non-linear or abrupt changes of reflectance characteristics. Our results further indicate how this measurement uncertainty may impinge on the conclusions concerning the function of plumage reflectance. A literature survey of studies of plumage reflectance (Supporting Information S1) indicates that a large majority measured reflectance well after the courtship period. Until several further studies demonstrate that the phenomenon we have described here is not general (which is unlikely, based on the literature), utmost caution is needed with reflectance measurements collected during nestling rearing. In particular, it does not seem entirely justified to use such measurements to estimate certain aspects of sexual selection on plumage reflectance traits (e.g. social pairing, assortative mating or relationship with breeding date or maternal effects).

In addition to a methodological warning, our results have two important implications. First, the mechanism of spectral signal change during breeding requires further studies. Here we generally saw a deterioration of signal strength which involved a decline of BRI and UVC for the white patches, and an increase of BRI and no change of UVC for the dark areas (Table 2). These changes probably originate from macrostructural feather degradation and changes in dirt and wax cover, but this needs to be confirmed. In general, deterioration of plumage spectral traits may involve mechanical abrasion and soiling (Pérez-Rodríguez *et al.* 2011, Surmacki *et al.* 2011, Enbody *et al.* 2017), bacterial keratinolysis (Shawkey *et al.* 2007, Kilgas *et al.* 2012), changes in the thickness and composition of the preen gland wax (Pérez-Rodríguez *et al.* 2011, Surmacki 2011) and, for carotenoid-based traits, also pigment degradation (McGraw & Hill 2004, Surmacki 2008, Higginson *et al.* 2016).

As a second implication, the direction and magnitude of within-individual reflectance change may represent a signal and may convey information on aspects of individual state and quality for individuals that see each other throughout the season (mates, neighbours). It is known that the sizes of ornamental plumage patches change due to abrasion (Tökölyi *et al.* 2008) and experimental studies have shown that patch size change is a meaningful trait that alters the reproductive investment of social mates (Nakagawa *et al.* 2007, Hasegawa & Arai 2015). Analogously, several studies have demonstrated change in partner behaviour or investment in response to manipulated plumage

reflectance (e.g. Safran *et al.* 2005, Kingma *et al.* 2009, partly reviewed by Horváthová *et al.* 2012). It is now important to find out how widespread rapid natural reflectance shifts are during breeding (Surmacki *et al.* 2011) and what information they convey to conspecifics (Delhey *et al.* 2006, Safran *et al.* 2010).

REFERENCES

Avilés, J.M., Solis, E., Valencia, J., de la Cruz, C. & Sorci, G. 2008. Female and male plumage brightness correlate with nesting failure in Azure-winged Magpies *Cyanopica cyanus*. *J. Avian Biol.* **39**: 257–261.

Delhey, K., Peters, A., Johnsen, A. & Kempenaers, B. 2006. Seasonal changes in Blue Tit crown color: do they signal individual quality? *Behav. Ecol.* **17**: 790–798.

Delhey, K., Burger, C., Fiedler, W. & Peters, A. 2010. Seasonal changes in color: a comparison of structural, melanin- and carotenoid-based plumage colors. *PLoS ONE* **5**: e11582.

Dunlap, W.P., Cortina, J.M., Vaslow, J.B. & Burke, M.J. 1996. Meta-analysis of experiments with matched groups or repeated measures designs. *Psychol. Methods* **1**: 170–177.

Enbody, E.D., Lantz, S.M. & Karubian, J. 2017. Production of plumage ornaments among males and females of two closely related tropical passerine bird species. *Ecol. Evol.* **7**: 4027–4034.

Evans, S.R., Summers, A.G.R. & Sheldon, B.C. 2012. Seasonality of carotenoid-based plumage coloration: modelling wavelength-specific change through spectral reconstruction. *J. Avian Biol.* **43**: 234–243.

Garamszegi, L.Z., Rosivall, B., Hegyi, G., Szöllősi, E., Török, J. & Eens, M. 2006. Determinants of male territorial behaviour in a Hungarian Collared Flycatcher population: plumage traits of residents and challengers. *Behav. Ecol. Sociobiol.* **60**: 663–671.

Garamszegi, L.Z., Eens, M. & Török, J. 2008. Birds reveal their personality when singing. *PLoS ONE* **3**: e2647.

Hart, N.S., Partridge, J.C., Cuthill, I.C. & 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* L.) and the Blackbird (*Turdus merula* L.). *J. Comp. Physiol. A.* **186**: 375–387.

Hasegawa, M. & Arai, E. 2015. Experimentally reduced male ornamentation increased paternal care in the Barn Swallow. *J. Ornithol.* **156**: 795–804.

Hasegawa, M., Arai, E., Watanabe, M. & Nakamura, M. 2008. Methods for correcting plumage color fading in the Barn Swallow. *Ornithol. Sci.* **7**: 117–122.

Hegyi, G., Szöllősi, E., Jenni-Eiermann, S., Török, J., Eens, M. & Garamszegi, L.Z. 2010. Nutritional correlates and mate acquisition role of multiple sexual traits in male Collared Flycatchers. *Naturwissenschaften* **97**: 567–576.

Herényi, M., Garamszegi, L.Z., Hargitai, R., Hegyi, G., Rosivall, B., Szöllősi, E. & Török, J. 2014. Laying date and polygyny as determinants of annual reproductive success in male Collared Flycatchers (*Ficedula albicollis*): a long-term study. *Naturwissenschaften* **101**: 305–312.

Higginson, D.M., Belloni, V., Davis, S.N., Morrison, E.S., Andrews, J.E. & Badyaev, A.V. 2016. Evolution of long-term coloration trends with biochemically unstable ingredients. *Proc. R. Soc. B* **283**: 20160403.

Horváthová, T., Nakagawa, S. & Uller, T. 2012. Strategic female reproductive investment in response to male attractiveness in birds. *Proc. Biol. Sci.* **279**: 163–170.

Kilgas, P., Saag, P., Mägi, M., Edenberg, M., Tilgar, V. & Mänd, R. 2012. Variation in assemblages of feather bacteria in relation to plumage color in female Great Tits. *Condor* **114**: 606–611.

Kingma, S.A., Komdeur, J., Vedder, O., von Engelhardt, N., Korsten, P. & Groothuis, T.G.G. 2009. Manipulation of male attractiveness induces rapid changes in avian maternal yolk androgen deposition. *Behav. Ecol.* **20**: 172–179.

Laczi, M., Török, J., Rosivall, B. & Hegyi, G. 2011. Integration of spectral reflectance across the plumage: implications for mating patterns. *PLoS ONE* **6**: e23201.

Maia, R., Eliason, C.M., Bitton, P.P., Doucet, S.M. & Shawkey, M.D. 2013. pavo: an R package for the analysis, visualization and organization of spectral data. *Methods Ecol. Evol.* **4**: 906–913.

McGraw, K.J. & Hill, G.E. 2004. Plumage color as a dynamic trait: carotenoid pigmentation of male House Finches (*Carpodacus mexicanus*) fades during the breeding season. *Can. J. Zool.* **82**: 734–738.

Montgomerie, R. 2006. Cosmetic and adventitious colors. In Hill, G.E. & McGraw, K.J. (eds) *Bird Coloration*. Vol. **1**: 399–427. Cambridge, MA: Harvard University Press.

Nakagawa, S. & Schielzeth, H. 2011. Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol. Rev.* **85**: 935–956.

Nakagawa, S., Ockendon, N., Gillespie, D.O.S., Hatchwell, B.J. & Burke, T. 2007. Does the badge of status influence parental care and investment in House Sparrows? An experimental test. *Oecologia* **153**: 749–760.

Örnborg, J., Andersson, S., Griffith, S.C. & Sheldon, B.C. 2002. Seasonal changes in a ultraviolet structural color signal in Blue Tits, *Parus caeruleus. Biol. J. Linn. Soc.* **76**: 237–245.

Pérez-Rodríguez, L., Mougeot, F. & Bortolotti, G.R. 2011. The effects of preen oils and soiling on the UV-visible reflectance of carotenoid-pigmented feathers. *Behav. Ecol. Sociobiol.* **65**: 1425–1435.

Safran, R.J., Neuman, C.R., McGraw, K.J. & Lovette, I.J. 2005. Dynamic paternity allocation as a function of male plumage color in Barn Swallows. *Science* **309**: 2210–2212.

Safran, R.J., McGraw, K.J., Wilkins, M.R., Hubbard, J.K. & Marling, J. 2010. Positive carotenoid balance correlates with greater reproductive performance in a wild bird. *PLoS ONE* **5**: e9420.

Shawkey, M.D., Pillai, S.R., Hill, G.E., Siefferman, L.M. & Roberts, S.R. 2007. Bacteria as an agent for change in structural plumage color: correlational and experimental evidence. *Am. Nat.* **169**: S112–S121.

Surmacki, A. 2008. Preen waxes do not protect carotenoid plumage from bleaching by sunlight. *Ibis* **150**: 335–341.

Surmacki, A. 2011. Natural soiling has a small effect on structurally-based plumage coloration. *Anim. Biol.* **61**: 441–455.

Surmacki, A., Liu, M., Mercadante, A. & Hill, G.E. 2011. Effect of feather abrasion on structural coloration in male Eastern Bluebirds *Sialia sialis. J. Avian Biol.* **42**: 514–521.

Tökölyi, J., Bókony, V. & Barta, Z. 2008. Seasonal color change by molt or by the abrasion of feather tips: a comparative study. *Biol. J. Linn. Soc.* **94**: 711–721.

Török, J., Hegyi, G. & Garamszegi, L.Z. 2003. Depigmented wing patch size is a condition-dependent indicator of viability in male collared flycatchers. *Behav. Ecol.* **14**: 382–388.

Viechtbauer, W. 2010. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **36**: 1–48.

Willoughby, E.J., Murphy, M. & Gorton, H.L. 2002. Moult, plumage abrasion, and color change in Lawrence's Goldfinch. *Wilson Bull.* **114**: 380–392.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Data S1. Methods, results and references of the literature survey.

Table 1. Effects of year and age on the focal spectral variables: general linear models with backward stepwise selection and reintroduction.