Ringing & Migration (2005) **22**, 193-197

Sexing Starlings *Sturnus vulgaris* using iris colour

E.L. SMITH[1]*, I.C.CUTHILL[1], R. GRIFFITHS[2], V.J. GREENWOOD[1], A.R. GOLDSMITH[1] and J.E. EVANS[1]

[1]Centre for Behavioural Biology, School of Biological Sciences, University of Bristol, Woodland Road, Bristol BS8 1UG, UK [2]Molecular Laboratory, DEEB, Graham Kerr Building, Glasgow University, University Avenue, Glasgow G12 8QQ, UK

We took blood samples from 100 post-fledging juvenile Starlings *Sturnus vulgaris* for DNA sexing in late August, and scored the traits that are commonly thought to predict sex in adults: iris colour, length and shape of the throat feathers, degree of speckling of plumage, body mass and tarsus length. Using logistic regression, the iris colour alone predicted the sex, as determined from the DNA, with 97% accuracy. By using iris colour and the length of the throat feathers together, 98% of birds were sexed accurately and greater separation of groups was achieved. These features were still reliable in the following February, when the birds had come into breeding plumage. At this stage, the colour of the base of the bill was 100% accurate in predicting sex.

It is commonly thought that it is difficult to sex juvenile Starlings *Sturnus vulgaris* from their appearance. Although previous reports suggest it is possible to determine sex in juveniles by iris colour (Krätzig 1936, mentioned but not fully referenced in Svensson 1992), field guides typically do not describe any differences between male and female juveniles (eg Cramp & Perrins 1994, Snow & Perrins 1998). Likewise, despite reports that adult Starlings are either visibly sexually dimorphic or differ reliably in mass between sexes (Harrison 1928, Hick 1934, Parks 1962, Klijn 1975), such differences are considered to be unreliable (Svensson 1992). Given that Starlings may require specific conservation effort (Feare 1994, Baillie *et al* 2001), a reliable non-invasive sexing method for Starlings would be a valuable aid in monitoring the sex structure of their populations. Here, we demonstrate that it is possible to ascertain the correct sex of most post-fledging juvenile and first-winter Starlings by using morphological characters alone.

METHODS

We caught and ringed 100 juvenile wild Starlings under English Nature licence during July 2000 near Somerton, Somerset (51° 5'N 2° 44'W). We maintained these birds in captivity for 12 months either in outdoor aviaries, or indoors under full spectrum lighting where the lighting was set to mirror natural changes in photoperiod throughout the year.

Blood samples were taken from each bird for sexing at the end of August 2000 by VJG, about two months post-fledging. Using venepuncture, 50 μl of blood was collected from the alar vein using heparinised capillary tubes. Each sample was added to 0.5 ml of BLB buffer (50mM EDTA, 0.5% SDS, 50mM Tris). DNA sex identification was carried out by RG following the protocol used by Griffiths *et al* (1998), with modified thermal conditions: an initial 94°C for 90 sec followed by 30 cycles of 49°C for 30 sec, 72°C for 30 sec, 94°C for 15 sec with a final 49°C for 60 sec and 72°C for 300 secs. Any samples which did not give a clear DNA band were subsequently re-run using an iterative procedure until sex could be firmly established. A second re-run using a new gel showed very high repeatability between runs (Kendall's $W = 0.98$, $df = 99$, $P < 0.001$). In the single case where the two runs did not match, the sample was run a third time. In two out of three cases the gel showed the bird was female, so that was the sex we accepted for that animal.

At the time of blood sampling, ELS rated each bird on several plumage and morphological characters (Table 1), and ICC measured the tarsi length (maximum tarsus, Redfern & Clark 2001). At this stage, birds still had juvenile plumage, although most had a few first-winter feathers starting to grow through on the chest. These ratings were repeated by ELS in February 2001, when the birds were in breeding condition. Such measurements

---

*Correspondence author
Email: emma.smith@bristol.ac.uk

Table 1. Description of variables used to assess morphology and appearance of juvenile Starlings in August (A), and the following February (F). Only variables found to be significant predictors of sex were remeasured in February. Bill colouration was assessed only in February, as colour differences are only present in the breeding season.

Variable	Description	Time period
Iris colour	Iris colour changes from a washed out yellow-grey as a juvenile, to richer tawny/chestnut hues as an adult. Ignore this, and simply rate how light/dark the iris appears relative to the pupil (Fig 1). 1 = Much lighter, highly distinct ring, 2 = Lighter, clear ring, 3 = Dim ring, visible with careful observation 4 = So dark that it is indistinguishable from the pupil	A, F
Bill colour	In breeding plumage, bill is yellow but the base is either pink or blue. In juveniles and non-breeding first winter birds the entire bill is black. 1 = salmon pink 2 = grey-blue	F
Feather tip shape	Throat and chest feather tips range in shape from perfectly rounded, to angular 'V' shaped tips (see Fig 2). Shape of tips of the feathers of all visible feathers on the throat and belly were classified as: 1 = all rounded, 2 = mostly rounded, 3 = 50% round, 50% V-shaped 4 = mostly V-shaped, 5 = all V-shaped	A,F
Feather length	The throat and chest feathers either appear short and wide, or elongate and thin (see Fig 2). Appearance of feathers on the throat and entire belly were classified as: 1 = all short and wide, 2 = mostly short and wide, 3 = 50% short and wide, 50% long and thin, 4 = mostly long and thin, 5 = all long and thin	A,F
Speckling	Density of pale white/buff feather tips* on chest varies across birds. 2 x 2 cm paper template placed in centre of chest, without disrupting underlying feathers. Counted number of tips visible within square.	A
Mass	Grams	A, F
Tarsus length	Measured using callipers (mm)	A, F

*These pale tips usually become abraded during the breeding season, particularly in males (Feare 1984).

were always made blind to the genetic sex of each bird.

We chose plumage characters based on features that had previously been used to predict the birds' sex as adults, but yet were considered unreliable for juveniles (Feare 1984, Svensson 1992). We excluded measures that require elaborate equipment or complex analysis (eg spectrometric measures of feather reflectance, Cuthill *et al* 1999). We devised scales with which to rank the appearance, based on obvious categories of variation in each character. Characters rated included mass, tarsus size, iris colour, beak colour, number of speckles on the chest and the shape and length of the throat and chest feathers (see Table 1 and Figs 1 & 2 for details).

Our eye colour ratings assessed, on a four point scale, how light the iris was relative to the pupil: a score of 1 was a pale, whitish, iris, and a score of 4 a very dark one (Table 1 and Fig 1).

We used binary logistic regression in the statistical analysis programme SPSS to see which of our measured variables, or combination thereof, best predicted the DNA sex. We carried out individual binary logistic regressions on the data for each variable recorded in August. We then investigated whether any combination of these variables gave more accurate predictions than any single variable alone, using binary logistic regression using the forward stepwise method based on changes in Likelihood Ratio.

To investigate the reliability of these characters over time, the birds were scored again the following February, blind to the genetic sex and previous scores of each animal. We also recorded the colour of the base of the bill, which is sexually dimorphic during the breeding season (Feare 1984). We used a forward stepwise method based on changes in Likelihood Ratio to assess the factors that would indicate sex in first-winter birds in February. We had 96 animals in the study at this stage.

In August 2004, we carried out an inter-observer reliability test, in which ELS, ICC and a completely naïve rater, JEE, all scored 63 juvenile Starlings independently for iris colour and feather length. ELS scored the animals as described above, and ICC and JEE scored the animals by comparing them directly to the photographs and category descriptions provided in this paper.

RESULTS

Sexing of juveniles

According to the DNA sexing results, we had 44 females and 56 males. All variables except the number of speckles on the chest were significantly better than random at predicting sex (Table 2).

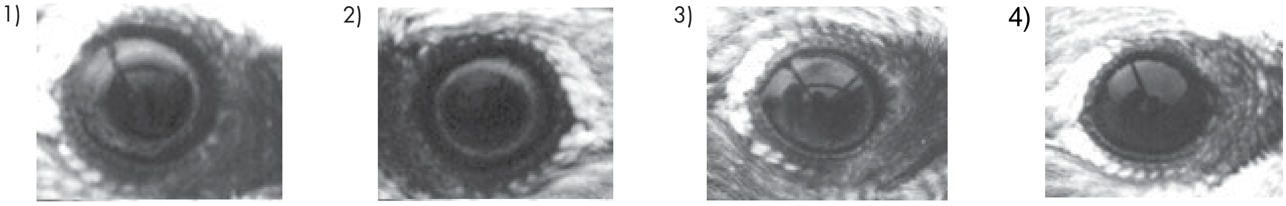

Figure 1. System for ranking perceived lightness of iris relative to pupil. The differences between categories, although visible here, are more striking when seen in colour. The pupil is always a dark chocolate brown/black colour. The surrounding iris is yellowish-grey in hue when birds are in juvenile plumage, changes to a pale tawny colour as the birds moult into their first winter plumage, and then becomes a rich chestnut colour as they come into breeding plumage. The exception to this is the eye rank 4 category, in which the iris is always such a deep brown that it is hard to distinguish from the pupil itself. Categories 1 and 2 are typically female, and categories 3 and 4 are typically male. These photographs are all from juveniles in August.

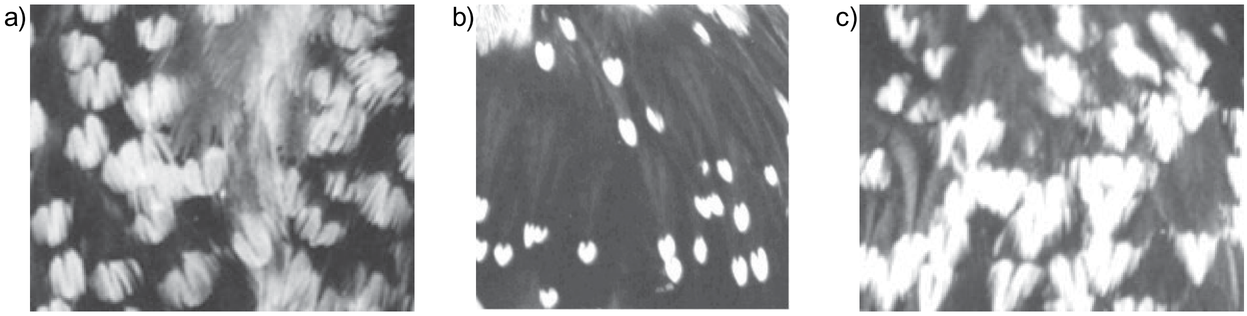

Figure 2. The shape of the pale feather tips on the throat and chest varies between being short, broad and rounded (Fig 2a, usually genetic female) to being elongate and thin, which makes the feather tips appear more angular and 'V' shaped (Fig 2b, usually genetic male). Typically, feathers of the type shown in Fig 2a are not very iridescent, whereas the long thin variety in 2b tend to be highly glossy and iridescent. Some birds have a mixture of the two feather types. Occasionally, one finds a bird with feathers of ambiguous intermediate characters. For example, Fig 2c shows a bird with broad, short feathers that are fairly iridescent, and which have V-shaped tips. These images are of the centre of the breast of juveniles in October.

Using iris colour alone, the model developed was extremely successful at classifying sex, even in juveniles (Table 2, 98% correctly classified). Eye scores of 1 and 2 were typically female, and scores of 3 and 4 were typically male (Fig 1). A model based on throat and chest feather length enabled us to correctly classify the sex of 94% of birds (Table 2 and Fig 2). We scored feathers on a five point scale ranging from being all short and wide (score=1) to all elongate and thin (score=5). A score of 3 indicated an ambiguous case, where there was either an even mixture of both types, or atypical feather shapes. Feather length scores of less than three were invariably female and more than four always male (Table 1), however for intermediate categories there were errors, particularly for the nine birds in category 3, which the model classified as female but were actually five females and four males. However, a model based on both iris colour and feather length gave significantly greater discriminatory power than using just the eyes or feathers alone (improvement in model fit, $\chi^2 = 8.217$, df = 1, $P = 0.004$). Although two birds were still misclassified (one of these was also misclassified in the 'eyes alone' model), the confidence with which borderline cases could be discriminated was improved by using both variables. Adding further variables did not significantly improve discrimination of the sexes, nor did any other combination of variables perform as well as the iris colour and feather length model. The precise probability of a bird being male, based on these characters, can be calculated using the formula:

$$P = \frac{e\text{\^{}}\,(-14.68 + 2.298^{*}\text{ eye score} + 2.58^{*}\text{ feather score})}{(1 + e\text{\^{}}\,(-14.68 + 2.298^{*}\text{ eye score} + 2.58^{*}\text{ feather score}))}$$

where 'e' represents the base number used in natural logarithm calculations (approximately 2.718) and '^' represents exponentiation (ie 'e to the power of'). If P is less than 0.5, then the predicted sex is female, if P is greater than 0.5 the predicted sex is male. The probability of a bird being female is calculated as 1-P.

It is of interest to consider the birds that were incorrectly classified by the models. In the 'feather length alone' model, six animals were misclassified, two of which were females (both with scores greater than three), and four of which were males (all score three).

Table 2. Analysis of each variable separately at each time period (A = August, F = February), showing the accuracy with which each can be used to predict sex. With the exception of bill colour, binary logistic regression, with a P value based on the change in log-likelihood as a result of adding the factor in question, was used throughout; this has an approximately χ^2 distribution. In the case of bill colour a Pearson chi-square contingency test was used, as the logistic regression could not compute statistics for a model that was a perfect fit to the data. For significant results, the cut-off point at which the model discriminated males from females is shown as a footnote.

Variable	Time period	% correct	Coefficient	χ^2	df	P
Bill colour[1]	F	100	-	97	1	<0.001
Iris colour[2]	A	98	4.12	118.1	1	<0.001
	F	98	5.10	117.9	1	<0.001
Feather length[3]	A	94	3.46	113.6	1	<0.001
	F	93	2.27	91.5	1	<0.001
Feather tip shape[4]	A	89	1.69	75.2	1	<0.001
	F	81	2.28	58.9	1	<0.001
Mass[5]	A	72	0.22	19.0	1	<0.001
	F	70	0.24	28.4	1	<0.001
Tarsus length[6]	A	65	0.92	11.5	1	0.001
	F	67	0.70	9.1	1	0.003
Speckling	A	56	0.24	–	1	0.231
	F	55	0.23	–	1	0.355

[1] Salmon pink bills were classed as female, and grey-blue as male, [2] Score of <3 classed as female, ≥3 as male, [3] Score of ≤3 classed as female, >3 as male, [4] Score of <3 classed as female, ≥3 as male, [5] Birds under 78g classed as female, [6] Birds with tarsi under 29.3 mm classed as female

In the 'eyes alone' model, two birds were misclassified. One was a genetic female, which was firmly misclassified as a male, with a 'male typical' score of four for both iris colour and feather shape. As there was no ambiguity, the combined model also misclassified this animal. The other was a genetic male, assigned an ambiguous eye score of 2.5, yet a 'male typical' feather length score of five. He was subsequently correctly classified by the combined model that considered both factors. However, due to the inclusion of feather shape as a predictor, this model also misclassified a genetic male which had been correctly categorised by the 'eyes alone' model. He was a borderline case, having a 'male typical' score of three for iris colour, but a 'female typical' score of two for feather shape.

Sexing of first-winter birds

Iris colour and feather length remained useful indicators of sex (Table 2). However, the best combination of variables for predicting the sex of first-winter birds was iris colour, feather length and mass, entered in that order, which together produced a perfect fit to the data (100% correctly classified, effect of removing mass from model: change in –2 Log Likelihood, $\chi^2 = 7.75$, df = 1, $P < 0.005$). The probability of a bird being male at this stage can be calculated using the formula: P = e^(-521.877 + 22.981* eye score + 28.051* feather score + 4.626* mass) / (1 + e^(-521.877 + 22.981* eye score + 28.051* feather score + 4.626* mass)). Again, if P is less than 0.5 then the predicted sex is female, if P is greater than 0.5 the predicted sex is male.

Also useful was the breeding season colouration that develops on the base of the lower mandible, as we found that this discriminated the genetic sexes with 100% accuracy (Table 2). Males had grey-blue bases to the bill, and females salmon pink ones.

Inter-observer reliability test

For iris colour, all three observers gave the animal the same eye rank category in only 60% of cases. However, there was agreement on whether or not the animal was male or female according to this character in 94% of cases (coefficient of concordance between scores, Kendall's W = 0.929, df = 62, $P < 0.001$). Where there was disagreement, the mismatched ratings were only one category apart. Some mismatches may have been caused by changeable and variable lighting conditions outdoors; irides appear darker when a cloud passes over and the pupil expands. Also, ICC and JEE both said they were aware of their classification changing slightly as they gained more experience of looking at Starling eyes. For feather length, all three observers gave the same category in only 27% of cases, but agreed on whether the animal was male or female according to this character in 77% of cases (Kendall's W = 0.901, df = 62, $P < 0.001$). Mismatches were never more than two categories apart. This shows that is possible for a naïve observer to achieve a high degree of accuracy in sexing using the figures provided in this paper.

DISCUSSION

There were sex differences in nearly all of the characters we measured in juvenile and first-winter Starlings. However, we found many sexually dimorphic characters, eg tarsus size and mass, are by themselves misleading in a high proportion of cases (Table 2). Although Svensson (1992) recommends combining all possible characters for maximum accuracy, we found that the most successful method of predicting sex relied only on assessing iris colour and the relative length to width of the emerging first-winter chest feathers. This produced 98% accuracy in sexing juveniles in August, and this remained accurate when the birds were in breeding plumage the following February. Even though the iris colour, gives very good discriminatory power, consideration of feather length is vital in highlighting borderline or atypical cases; as Svensson (1992) says, 'few sex categories are neatly defined and there will always be a few birds that do not look as they should'. With first-winter Starlings, if body mass is considered in addition to eye colour this allows perfect discrimination. However, when in breeding plumage, we found that the simpler method of looking at the colour of the base of the bill predicted sex with 100% accuracy. This is in contrast with the findings of Klijn (1975), who concluded from dissection of 74 Starlings that the eye colouration totally agrees with the sex of the animal, and the base of the bill colouration was unreliable. One possible explanation is that in a small proportion of animals, gonadal sex may not match genetic sex; ie the animals may be 'sex reversed' (Lewis & Long 1992). It is also noteworthy that none of the animals in the Klijn study were healthy, as all were in poor condition and died on capture, and thus were an atypical sample of the many birds that were caught.

We conclude that it is possible to accurately sex post-fledging Starlings using external features, even when they are not in breeding plumage. Our reliability test has shown that even completely naïve observers can do this successfully, simply by rating the birds according to the photographs and ranking scales provided in this paper.

ACKNOWLEDGEMENTS

This research was performed under UK Home Office project licence PPL 30/1651, and conformed to the guidelines of the Association for the Study of Animal Behaviour. This study was funded by Universities Federation of Animal Welfare studentship to ELS.

REFERENCES

Baillie, S.R., Crick, H.Q.P., Balmer, D.E., Bashford, R.I., Beaven, L.P., Freeman, S.N., Marchant, J.H., Noble, D.G., Raven, M.J., Siriwardena, G.M., Thewlis, R. & Wernham, C.V. (2001) *Breeding Birds in the Wider Countryside: their conservation status 2000.* BTO, Thetford.

Cramp, S. & Perrins, C.M. (1994) *The birds of the Western Palearctic. Vol VIII: Crows to Finches.* Oxford University Press, Oxford.

Cuthill, I.C., Bennett, A.T.D., Partridge, J.C. & Maier, E.H. (1999) Plumage reflectance and the objective assessment of avian sexual dichromatism. *American Naturalist* **153**, 183-200.

Feare, C.J. (1984) *The Starling.* Oxford University Press, Oxford.

Feare, C.J. (1994) Changes in numbers of Common Starlings and farming practice in Lincolnshire. *British Birds* **87**, 200-204.

Griffiths, R., Double, M., Orr, K. & Dawson, R. (1998) A simple DNA test to sex most birds. *Molecular Ecology* **7**, 1071-1076.

Harrison, J.M. (1928) The colour of the soft parts of the Starling. *British Birds* **22**, 36-37.

Hick, L.E. (1934) Individual and sexual variations in the European Starling. *Bird Banding* **5**, 103-118.

Klijn, H.B. (1975) Quantitative and qualitative age and sex criteria of the Starling, *Sturnus vulgaris*, in March. *Bijdragen tot de Dierkunde* **45**, 39-49.

Lewis, P.D. & Long, S.E. (1992). Incidence of non-laying in domestic hens. *British Poultry Science* **33**, 289-295.

Parks, G.H. (1962) A convenient method of sexing and ageing the Starling. *Bird Banding* **33**, 148-151.

Redfern, C.P.F. & Clark, J.A. (2001) *Ringers' Manual.* BTO, Thetford.

Snow, D.W. & Perrins, C.M. (1998) *The Birds of the Western Palearctic, Concise Edition. Vol 2 Passerines.* Oxford University Press, Oxford.

Svensson, L. (1992) *Identification Guide to European Passerines.* 4th edn. Svensson, Stockholm.

MS received 15 September 2004; MS accepted 12 August 2005