Assessing the function of house sparrows' bib size using a flexible meta-analysis method

Shinichi Nakagawa, Nancy Ockendon, Duncan O.S. Gillespie, Ben J. Hatchwell, and Terry Burke
Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

The black throat patch or bib of male house sparrows, *Passer domesticus*, is often referred to as a "badge of status" or a "badge" because previous studies have shown bib size to be correlated with the social status of males. Yet, little is known about how strong and robust this relationship is and how the strength of this relationship compares with that of other associations. We conducted a meta-analysis for 6 well-studied correlates of bib size: fighting ability, parental ability (egg incubation and food provisioning), age, body condition, cuckoldry, and reproductive success. We introduce a flexible meta-analysis method in this study that is better suited in the biological sciences than the methods usually employed in popular meta-analysis software because our method accounts for a common form of nonindependence of the data. The relationship between fighting ability and bib size was found to be strong and robust, and the relationship between age and bib size was moderate and robust. Also, body condition was weakly but significantly correlated with badge size. The other parameters showed nonsignificant small effects and/or large confidence intervals. Therefore, we conclude that bib size signals dominance and to a lesser extent age and possibly reflects body condition in house sparrows. There was weak evidence that bib size is currently under sexual selection because there was little association between reproductive success and bib size. This is surprising as the bib size probably affects the outcome of male–male competition. Empirical data on sparrow bib size could not be reconciled with sexual selection theory, although there is ample evidence that it is a condition-dependent trait. *Key words:* age, badge of status, condition-dependent trait, house sparrow, meta-analysis, sexual selection. *[Behav Ecol 18:831–840 (2007)]*

Understanding phenotypic variation in traits has been a focus of evolutionary biology because phenotypic variation, which is the product of complex interactions between genetic and environmental variation, provides substrates on which natural and sexual selection can act (Roff 1997; Fox and Wolf 2006). Among the countless phenotypic variations investigated, avian plumage variation both within and between species appears to have attracted a disproportionate amount of research (Hill 2002; Owens 2006). Numerous studies have demonstrated that avian plumage, such as elongated tail feathers and colorful patches, play an important role both in aggressive interactions and in mate choice (reviewed in Andersson 1994; Andersson and Simmons 2006).

Although the plumage of male house sparrows (*Passer domesticus*) is far from extravagant, the house sparrow is one of the most intensively studied species in relation to plumage variation (reviewed in Jawor and Breitwisch 2003; Anderson 2006; Griffith et al. 2006). House sparrows have been a model organism in studies of sexual selection, in which correlations have been investigated between numerous life-history traits and the size of the black bib (a secondary sexual trait, e.g., Griffith et al. 1999a; reviewed in Anderson 2006). The best-documented association is that between dominance hierarchy and bib size (e.g., Møller 1987a, 1987b; Liker and Barta 2001; Gonzalez et al. 2002; see Table 1). Therefore, the bib is often referred to as a "badge of status" or just a "badge" (Dawkins and Krebs 1978).

Rohwer (1975) first suggested that conspicuous size variation in plumage patches within species evolved to signal differences in fighting ability to avoid wasteful fights. However, we have little idea about how strong and robust the association

between status and bib size actually is and how the strength of this association compares with that of other associations (e.g., between age and bib size or between parental care and bib size). For such estimations and comparisons, we require a meta-analytical approach, that is, effect-size–based reviews of research (Hunt 1997; see also Nakagawa 2004; Nakagawa and Cuthill 2007). With house sparrows being a model species of sexual selection, we now have an opportunity to conduct meta-analyses using this species alone. However, a prevalent problem of meta-analyses, especially in evolutionary biology and behavioral ecology, is that the statistical assumption of independence of the data is violated when effect sizes from studies by the same research groups or of the same populations or species are used as independent data points (e.g., Møller and Ninni 1998; Dubois and Cézilly 2002; see also Palmer 2000). Meta-analysis software that routinely deals with this nonindependence does not seem to be available (see Methods for more details). In this paper, we used a "flexible" meta-analysis method that dealt with this problem.

We examine 6 well-studied associations between male bib size and life-history traits, including fighting ability, parental ability (nestling provisioning and incubation), age, body condition, cuckoldry (loss of paternity to extrapair males), and reproductive success. This paper's aims are 3-fold: 1) to investigate the generality of the alleged associations regarding bib size, 2) to outline and use a method of meta-analysis more suited for the biological sciences, and 3) to explore the current understanding of the bib size of house sparrows as both a sexually selected and a condition-dependent trait, based on previous findings and on theoretical models.

METHODS

Collecting studies and extracting effect sizes

We conducted an extensive search of the literature on the relationship between life-history traits and bib size in house sparrows. We also contacted researchers who have published

---

Address correspondence to S. Nakagawa. E-mail: itchyshin@yahoo.co.nz.

Table 1

Studies used in meta-analyses of the effect of bib size on fighting ability (Status), parental ability (Parental), age (Age), body condition (Condition), Cuckoldry (Cuckoldry), and reproductive success (Success)

Data ID	Population	Effect size (r)	n	Original statistics	Reference	Note (method of bib measurement)
Status 1	Richmond, KY, USA	0.53	25	$\chi^2 = 7.03$ (df = 1)	Ritchison (1985)	Aggressive interactions in outside aviaries in winter (Ritchison 1985)
Status 2	Hollensted, Denmark	0.57	13	Spearman's r	Møller (1987b)	Aggressive interactions in outside aviaries in winter (Møller 1987b)
Status 3	Ø. Brøndwerslev, Denmark	0.52	10	Spearman's r		
Status 4	Pandrup, Denmark	0.89	14	Spearman's r		
Status 5	Hestmona, Norway	0.32	9	Spearman's r	Solberg and Ringsby (1997)	Aggressive interactions in indoor aviaries in winter. Both visual and total bib size were presented; the results from visual bib size were used. Data from different years were presented separately (Møller 1987b)
Status 6		0.04	6	Spearman's r		
Status 7	Stoselsøy, Norway	0.33	11	Spearman's r		
Status 8		0.54	9	Spearman's r		
Status 9	Ytre-Kvarøy, Norway	0.33	9	Spearman's r		
Status 10	Budapest, Hungary	0.88	10	Spearman's r	Liker and Barta (2001)	Aggressive interactions in an indoor aviary in winter (Veiga 1993)
Status 11		0.483	19	η^2 (R^2) = 0.274	Bókony et al. (2006)	Aggressive interactions in an indoor aviary in winter (Bókony et al. 2006)
Status 12	Badajoz, Spain	0.488	41	Pearson's r	Gonzalez et al. (2002)	Aggressive interactions in outside aviaries in breeding season (Møller 1987b)
Status 13	Madison, WI, USA	0.147	20	Spearman's r	Riters et al. (2004)	Aggressive interactions in indoor aviaries in breeding season; the authors provided r_s (Møller 1987b)
Status 14	Lexington, KY, USA	0.37	22	Kendall's τ	Hein et al. (2003)	Aggressive interactions over outside feeders; 2 different estimations of bib size used (Møller 1987b; Hein et al. 2003)
Status 15	Princeton, NJ, USA	0.10	28	Spearman's r	Lindström et al. (2005)	Aggressive interactions in outside aviaries in winter (Lindström et al. 2005)
Parental 1	Albuquerque, NM, USA	0.49	13	Pearson's r	Kimball (1995)	Feeding rate (visits/chick h) (Møller 1987b)
Parental 2	Norman, OK, USA	0.36	27	Pearson's r	Voltura et al. (2002)	Feeding rate (visits/chick h) (Voltura et al. 2002)
Parental 3	Norman, OK, USA	0.03	45	Pearson's r	Bartlett et al. (2005)	Incubation time (min/h; %) (Bartlett et al. 2005)
Parental 4	Vienna, Austria	0.06	16	Spearman's r	Václav and Hoi (2002)	Incubation contribution %; males with medium bib size found to provide the most care (Møller 1987b)
Parental 5	Lundy, UK	ca. −0.001	29	$R^2 = 0.01$	Griffith (1998)	Feeding rate (visits/chick h) (Griffith et al. 1999a, 1999b)
Parental 6		−0.45	19	Pearson's r	Ockendon (2003)	Feeding rate (visits/chick h) (Griffith et al. 1999a, 1999b)
Parental 7		0.003	74	Pearson's r	Nakagawa et al. (2007)	Feeding rate (visits/chick h) (Griffith et al. 1999a, 1999b)
Parental 8		−0.056	88	Pearson's r		Incubation time (min/h; %) (Griffith et al. 1999a, 1999b)
Parental 9	Deux Sèvres, France	0.458	20	Spearman's r	Mazuc J, Sorci G (unpublished data)	Feeding rate (visits/chick h) (Møller 1987b)
Age 1	Ø. Brøndwerslev, Denmark	0.14	10	Spearman's r	Møller (1987b)	2 classes: 1 or 2 or more years old (Møller 1987b)
Age 2	Pandrup, Denmark	−0.19	14	Spearman's r		
Age 3	Hollensted, Denmark	0.04	13	Spearman's r		
Age 4	Collado Villalba, Spain	ca. 0.62	128	$F = 9.37$ (df = 2, 125)	Veiga (1993)	3 classes: 1, 2, and 3 or more years old (Veiga 1993)

Table 1, continued

Data ID	Population	Effect size (r)	n	Original statistics	Reference	Note (method of bib measurement)
Age 5	Albuquerque, NM, USA	0.506	17	$t = 2.27$ (df = 15)	Kimball (1996)	2 classes: 1 or 2 or more years old (Møller 1987b)
Age 6	Nottingham, UK	0.364	30	$t = 2.09$ (df = 28)	Cordero et al. (1999)	2 classes: 1 or 2 or more years old (Møller 1987b; Cordero et al. 1999)
Age 7	Barcelona, Spain	0.562	18	$t = 2.72$ (df = 16)		
Age 8	Lexington, KY, USA	0.23	99	R^2 ($r = 0.25$)	Hein et al. (2003)	"Minimum age" was used (alternative effect size: $r = 0.23$, $n = 95$) (Møller 1987b; Hein et al. 2003)
Age 9	Princeton, NJ, USA	0.42	28	Spearman's r	Lindström et al. (2005)	2 classes: 1 or 2 or more years old; the authors provided r_s (Møller 1987b)
Age 10	Helgeland, Norway	0.252	50	paired $t = 3.671$	Jensen et al. (2006)	Bib size of the same sparrows in 2 consecutive years; the authors provided t and r values (Møller 1987b)
Age 11	Lundy, UK	0.312	195	Pearson's r	Nakagawa et al. (2007)	Actual age was used (Griffith et al. 1999a, 1999b)
Condition 1	Ø. Brøndwerslev, Denmark	0.14	10	Spearman's r	Møller (1987b)	Body condition = weight/(wing length)^3 (Møller 1987b)
Condition 2	Pandrup, Denmark	−0.19	14	Spearman's r		
Condition 3	Hollensted, Denmark	−0.22	13	Spearman's r		
Condition 4	Copenhagen, Denmark	0.11	149	Pearson's r	Møller and Erritzøe (1988)	Body condition = weight/(tarsus length)^3 (Møller 1987b)
Condition 5	Hollensted, Denmark	−0.05	83	Pearson's r	Møller (1989)	Body condition = weight/(tarsus length)^3; measurements were taken in autumn, breeding season, and winter, respectively (Møller 1987b)
Condition 6	Hollensted, Denmark	−0.23	33	Pearson's r		
Condition 7	Hollensted, Denmark	0.50	60	Pearson's r		
Condition 8	Collado Villalba, Spain	0.209	130	Pearson's r	Veiga (1993)	Physical condition (residual weight regressed on tarsus length) used instead of social rank as a correlate (Veiga 1993)
Condition 9		0.15	178	Pearson's r	Veiga (1996)	The third principal component of 10 morphological measurements including weight (Veiga 1993)
Condition 10	Hestmona, Norway	0.33	9	Spearman's r	Solberg and Ringsby[a] (1997)	Residual of weight regressed on tarsus length (Møller 1987b)
Condition 11		0.13	6	Spearman's r		
Condition 12	Stoselsøy, Norway	−0.25	11	Spearman's r		
Condition 13		0.05	9	Spearman's r		
Condition 14	Ytre-Kvarøy, Norway	0.13	9	Spearman's r		
Condition 15	Nottingham, UK	0.52	17	Pearson's r	Cordero et al. (1999)	Residual of weight regressed on tarsus length (Møller, 1987b; Cordero et al. 1999)
Condition 16	Barcelona, Spain	0.32	18	Pearson's r		
Condition 17	Lexington, KY, USA	0.161	46	Pearson's r	Stewart (1999)	Residual of weight regressed on tarsus length (Møller, 1987b; Stewart 1999)
Condition 18	Extremadura, Spain	0.05	115	$R^2 = 0.03$	Gonzalez et al. (1999)	Residual of weight regressed on tarsus length (Møller 1987b)
Condition 19	Helgeland, Norway	0.031	81	Pearson's r	Jensen et al. (2006)[a]	Residual of weight regressed on tarsus length (Møller 1987b)
Condition 20	Lundy, UK	0.027	195	Pearson's r	Nakagawa et al. (2007)	Residual of weight regressed on tarsus length (Griffith et al. 1999a, 1999b)
Cuckoldry 1	Norman, OK, USA	−0.18	33	Spearman's r	Whitekiller et al. (2000)	Correlation between bib size and % of extrapair offspring (Whitekiller et al. 2000)
Cuckoldry 2	Collado Villalba, Spain	0.104	33	Mean ± standard deviation ($n_1 = 5$, $n_2 = 28$)	Veiga and Boto (2000)	Comparing bib size between cuckolded and noncuckolded males (Veiga 1993)

Table 1, continued

Data ID	Population	Effect size (r)	n	Original statistics	Reference	Note (method of bib measurement)
Cuckoldry 3	Vienna, Austria	–0.54	16	Spearman's r	Václav et al. (2002)	Correlation between residual bib size and % of extrapair offspring (Møller 1987b)
Cuckoldry 4	Nottingham, UK	–0.355	12	$t = 1.20$ (df = 10)	Cordero et al. (1999)	Comparing bib size between cuckolded and noncuckolded males (Møller, 1987b; Cordero et al. 1999)
Cuckoldry 5	Barcelona, Spain	–0.133	17	$t = 0.52$ (df = 15)		
Cuckoldry 6	Albuquerque, NM, USA	0.037	11	$t = 0.112$ (df = 9)	Kimball (1995)	Comparing bib size between cuckolded and noncuckolded males (Møller 1987b)
Cuckoldry 7	Lundy, UK	0.74	25	$R^2 = 0.54$	Ockendon (2003)	Regression analysis of bib size and % of extrapair offspring (Griffith et al. 1999a, 1999b)
Cuckoldry 8	Lexington, KY, USA	–0.174	43	$\chi^2 = 1.30$ (df = 1)	Stewart et al. (2006)	Comparing bib size between cuckolded and noncuckolded males (Stewart et al. 2006)
Cuckoldry 9	Helgeland, Norway	–0.029	26	$t = -0.141$ (df = 9)	Stewart (1999)	Comparing bib size between cuckolded and noncuckolded males (Møller, 1987b)
Success 1	Collado Villalba, Spain	–0.318	31	Mean $\pm$ standard deviation ($n_1 = 19$, $n_2 = 12$)	Veiga (1993)	No. of fledglings; only 2 experimental groups are used (experimental study) (Veiga 1993)
Success 2	Lexington, KY, USA	0.147	16	Pearson's r	Stewart (1999)	No. of fledglings (Stewart 1999)
Success 3		0.434	25	Pearson's r		No. of fledglings (Møller 1987b; Stewart 1999)
Success 4	Norman, OK, USA	0.465	27	Wald $\chi^2 = 7.18$ (df = 1)	Voltura et al. (2002)	No. of fledglings (Voltura et al. 2002)
Success 5	Vienna, Austria	0.26	30	Spearman's r	Václav and Hoi (2002)	No. of fledglings (Møller 1987b)
Success 6	Helgeland, Norway	–0.115	26	Pearson's r	Stewart (1999)	No. of fledglings (Møller 1987b)
Success 7		0.22	83	Wald $\chi^2 = 4.18$ (df = 1)	Jensen et al. (2004)	Lifetime reproductive success (Jensen et al. 2004; Møller 1987b)
Success 8	Lundy, UK	–0.38	40	$R^2 = 0.17$	Griffith et al. (1999b)	No. of fledglings (Griffith et al. 1999a, 1999b)
Success 9		0.19	36	$R^2 = 0.06$	Ockendon (2003)	No. of fledglings; data from 2000–2001 and data from 2002, respectively (Griffith et al. 1999a, 1999b)
Success 10		0.53	33	$R^2 = 0.30$		
Success 11		–0.013	95	Pearson's r	Nakagawa et al. (2007)	No. of fledglings (Griffith et al. 1999a, 1999b)
Success 12	Deux Sèvres, France	0.033	31	Spearman's r	Mazuc J, Sorci G (unpublished data)	No. of fledglings (Møller 1987b)

[a] Effect sizes from Solberg and Ringsby (1997) and Jensen et al. (2006) in condition are not independent in a strict sense but we used them as independent studies in our analysis.

studies on relevant topics to obtain unpublished information. We considered that, for a meta-analysis, at least 8 studies investigating the same traits in relation to bib size were necessary and that these studies required the inclusion of quantitative information that could be translated into effect size statistics. Our criteria resulted in 6 life-history traits that could be investigated: fighting ability, parental ability, age, body condition (i.e., standardized weight), cuckoldry (the rate or absence/presence of lost paternity due to extrapair paternity [EPP]), and reproductive success (details of studies used in the meta-analysis are summarized in Table 1). In the analysis of parental behavior, we used studies investigating both incubation and feeding behavior. A recent study by Kopisch et al.

(2005) showed incubation time to be a significant predictor of subsequent nestling provisioning in a population, so that treating incubation and provisioning together may be justifiable. Also, this type of integration or grouping is a common practice in meta-analysis. However, we carried out separate meta-analyses on incubation and feeding to further justify our procedure, although, by doing so, we violated our criteria (as there were only 6 studies of feeding and 3 studies of incubation).

Effect size values from different populations were treated as independent data points; also, within the same population, effect size estimates from replicated studies were used for meta-analysis without merging them (see below for how we dealt with pseudoreplication). We employed correlation

Table 2
The results of meta-analyses using LMMs with REML for effects of bib size on fighting ability (Status), parental ability (Parental: Feeding and Incubation), age (Age), condition (Condition), cuckoldry (Cuckoldry), and reproductive success (Success)

Meta-analysis ID	k	m	n	Effect size r (Z_r)	t value (p, df)	95% CI for r (95% CI for Z_r)	Heterogeneity Q_{REML} (p, df) [Q_T (p, df)]
Status	15	12	246	**0.463 (0.502)**	5.382 (0.0002, 12)	0.290 to 0.608 (0.299 to 0.705)	17.792 (0.164, 14) [22.760 (0.064, 14)]
Parental	9	5	331	0.142 (0.143)	1.246 (0.268, 5)	−0.151 to 0.412 (−0.152 to 0.438)	7.421 (0.492, 8) [14.04 (0.081, 8)]
Feeding	6	4	182	0.250 (0.255)	1.505 (0.207, 4)	−0.212 to 0.621 (−0.216 to 0.726)	4.191 (0.522, 5) [12.786, (0.025, 5)]
Incubation	3	3	149	0.011 (0.011)	0.326 (0.775, 2)	−0.138 to 0.160 (−0.138 to 0.161)	<0.0001 (0.999, 2) [0.299, (0.861, 2)]
Age	11	11	602	**0.346 (0.361)**	4.841 (0.001, 10)	0.192 to 0.483 (0.195 to 0.527)	3.011 (0.981, 10) [24.804 (0.006, 10)]
Condition	20	14	1236	**0.093 (0.093)**	2.404 (0.031, 14)	0.010 to 0.174 (0.010 to 0.176)	33.466 (0.021, 19) [33.466 (0.021, 19)]
Cuckoldry	9	9	216	−0.045 (−0.045)	−1.238 (0.762, 8)	−0.359 to 0.279 (−0.376 to 0.286)	<0.0001 (0.999, 8) [28.678 (0.0003, 8)]
Success	12	7	473	0.106 (0.106)	1.280 (0.241, 7)	−0.090 to 0.294 (−0.090 to 0.302)	33.093 (0.001, 11) [33.093 (0.001, 11)]

The numbers of effect sizes (k), the populations (m), and the individuals or observations (n) used for meta-analyses are indicated. Statistically significant effect sizes (Status, Age, and Condition) are in bold.

coefficients, which were subsequently converted into Fisher's Z (Z_r) for meta-analysis, as a standardized effect size statistic. The correlation coefficient (r including Pearson's r, Spearman's r, Kendall's τ, point-biserial r, and phi) was extracted from each study. In the cases where studies did not include correlation coefficients, the reported statistics (t, F, χ^2, z, and means and standard deviations) were used to estimate r following Rosenthal (1994) and Lipsey and Wilson (2001). When R^2 was the only source for estimating effect size, we used the formula

$$r = \sqrt{R^2 - \frac{p(1-R^2)}{n-p-1}}, \qquad (1)$$

where p is the number of predictors in the model and n is the sample size (the root square of R^2 is a biased estimate of effect size; Ezekiel 1930 cited in Thompson 2002). When related statistics were not immediately appropriate for estimating the correlation coefficient (e.g., F with a numerator degrees of freedom of more than 2), we used an approximation, which is indicated as ca. in Table 1. If a study included more than one effect size for the same or related traits, we chose values that made more biological sense (e.g., Solberg and Ringsby 1997; Jensen et al. 2006; see Table 1); if this criterion did not decide which effect size to use, we used the one with the largest absolute value (e.g., Hein et al. 2003; note that we also conducted meta-analyses using the alternative effect sizes reported in our collection of studies but none of these changed our main conclusions).

Flexible meta-analytic procedures

All statistical analyses were conducted in the R environment (version 2.4.1; R Development Core Team 2006; the R script and data sets used in this study are available from the corresponding author by request). As far as we are aware, there is no popular meta-analysis software that incorporates a grouping random factor when one or more groups include only a single data point. Incorporation of a grouping factor allows us to use nonindependent effect sizes without merging them prior to a meta-analysis (combining effect size estimates reduces statistical power and loses information). To deal with this, we used a linear mixed-effects model (LMM) approach with the restricted maximum likelihood method (REML, nlme package in R; Pinheiro and Bates 2000). The R function lme in the nlme package enabled us to use a grouping random factor even when each group had only a single data point

(one effect size value). Meta-analysis is basically a weighted linear regression analysis with the response (dependent) variable consisting of effect size values and with the weighting being the inverse variance of the corresponding effect size statistic; this type of meta-analysis is usually referred to as a fixed-effects model (Lipsey and Wilson 2001). The weight used for a fixed-effects model with Z_r is written as

$$w_i = n - 3, \qquad (2)$$

where w_i is the weight for a particular effect size and n is the sample size (the number of observations) used to estimate that effect size. The random-effects model in meta-analysis is characterized by the use of a modified weight; for Z_r, the formula is

$$w_i^* = n - 3 + \frac{1}{\tau^2}, \qquad (3)$$

where τ^2 is the between-study variance (in our case, between-population variance; Raudenbush 1994). A random-effects model (sometimes referred to as a mixed-effects model) in meta-analysis is usually a 3-step process (Rosenberg et al. 2000): in the case of Z_r, 1) running a fixed-effects model using weights from Equation 2 to obtain the summary statistics, 2) using these statistics to find τ^2, and then 3) running a random-effects model using weights from Equation 3 (for more details, see Rosenberg et al. 2000; Lipsey and Wilson 2001). The use of the R function lme with REML achieves an equivalent of this process simultaneously, providing a mean effect size estimate as the intercept of an LMM (see Pinheiro and Bates 2000).

In meta-analysis, the total heterogeneity of a sample, Q_T, is usually calculated to determine whether a set of effect sizes is homogeneous (Hedges and Olkin 1985). Q_T is found from Z_r by

$$Q_T = \sum_{i=1}^{k} w_i (Z_{r_i} - \overline{Z_r})^2, \qquad (4)$$

where k is the number of effect size estimates and $\overline{Z_r}$ is the weighted mean of Z_r effect sizes, which is given by

$$\overline{Z_r} = \frac{\sum_{i=1}^{k} w_i Z_{r_i}}{\sum_{i=1}^{k} w_i}. \qquad (5)$$

The heterogeneity Q_T is tested against a χ^2-distribution with $k - 1$ degrees of freedom. When Q_T is significant (in a
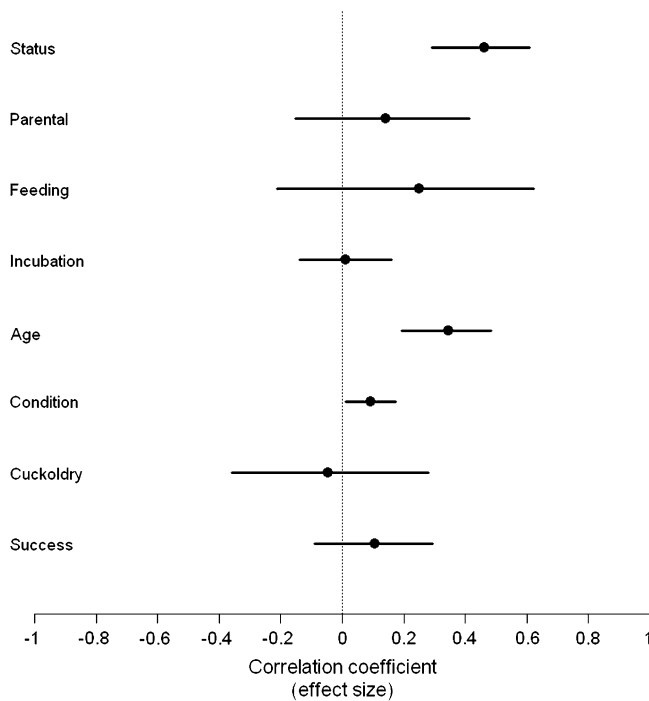
Figure 1
Visual presentation of the results from meta-analyses on the correlates of bib size (see Table 2 for details); effect size (correlation coefficient) and 95% confidence intervals are shown for each result.

fixed-effects model), the variance among effect sizes is greater than expected from sampling error, indicating that some explanatory variables that contribute to significant heterogeneity may exist (Rosenberg et al. 2000). In meta-analysis, explanatory (independent or predictor) variables are often referred to as moderators or moderator variables. For our meta-analyses, we calculated Q_T using Equations 4 and 5. However, Q_T may not portray heterogeneity in LMMs well because Equations 4 and 5 treat each effect size as independent, and Q_T does not take account of a random component of variation between effect sizes. The heterogeneity Q_T in weighted linear regression analysis equates to a "weighted" residual sum-of-squares of an LMM (calculated as in Equation 4) when the model does not have any fixed factors (fixed moderators) or only has an intercept (see Hedges and Olkin 1985; Lipsey and Wilson 2001). Therefore, we used these residual sum-of-squares as our main measure of heterogeneity (Q_{REML}; more precisely, it is an equivalent of residual heterogeneity in random-effects models) and tested Q_{REML} against a χ^2-distribution with $k - 1$ degrees of freedom (Table 2). For each meta-analysis, we provide a t value to test if the intercept in the respective LMM (i.e., mean effect size estimates; Raudenbush 1994) was significantly greater than zero using the corresponding degrees of freedom from the LMM: $k - 1$ when no group (i.e., population) had more than one effect size and the number of populations (m) when one or more populations had more than one effect size. Even when the heterogeneity was significant, we did not try to fit any possible moderator variables; for all the traits used in the meta-analysis, the sample size of the effect sizes was small ($k = 3–20$), so that adding moderators (fixed factors) into our meta-analyses would not have led to any meaningful and/or interpretable results.

Meta-analysis is often accompanied by one or more types of estimator of publication bias (reviewed in Møller and Jennions 2001). Popular estimators include fail-safe numbers and rank

correlation tests. Due to our small sample size (the number of studies, k in Table 2), both of these publication bias estimators were of little use here. Instead, we used funnel plots (i.e., plots of effect size against the sample size on which each effect size was based) to enable a visual inspection for bias (see Palmer 2000; Cassey et al. 2004).

RESULTS

The results of our 8 meta-analyses are summarized in Table 2, and Figure 1 presents a visual comparison of effect size estimates and their 95% confidence intervals (CIs). Bib size showed a strong positive association with status or fighting ability, a moderate association with age, and a weak association with body condition (all associations were statistically significant; sensu Cohen 1988; $r = 0.1$, 0.3, and 0.5 for "small," "moderate," and "large" effects). Parental ability (both provisioning and incubation), cuckoldry, and reproductive success all showed nonsignificant small effects and/or large CIs. Funnel plots for each variable are shown (Figure 2). The degree of asymmetry about the mean effect size in Figure 2 is reflected in the extent of the respective CI (Figure 1), which in turn seems to be influenced by the effect size sample size (k). As far as we can see, there is little indication of publication bias for fighting ability (status), age, and body condition, each of which showed a significant relationship with bib size. The heterogeneity Q_{REML}, which incorporated a random component of variation, was significant for body condition and reproductive success (Table 2; note that Q_T values are shown for comparison; these values were not interpreted as they do not account for random variation between populations). These significant heterogeneities indicated the possible presence of moderator variables for body condition and reproductive success, and, therefore, the estimated effect size values for each factor should be interpreted with care.

DISCUSSION

What is the function of the bib size?

Our meta-analysis revealed that the relationship between fighting ability and bib size is robust and strong in house sparrows ($r = 0.464$), confirming that the black throat patch indeed acts as a badge of status. Also, the relationship between bib size and age is robust and moderate ($r = 0.346$), suggesting that bib size also acts as an age indicator, at least to some extent (see Brooks and Kemp 2001). Anderson (2006) concluded in his monograph on house sparrows that the only consistent finding concerning bib size was the association with age and that other associations, including fighting ability, were equivocal and unconvincing. However, his conclusion was based on a vote-counting–type review, where the overall conclusion was based on the number of papers with significant and nonsignificant results, regardless of sample size (Cooper and Hedges 1994; Hunt 1997). Anderson (2006) is probably correct that age accounts for some component of the badge of status in house sparrows. In many badge-of-status systems, the extent or even existence of badges is sex and/or age specific (Searcy and Nowicki 2005). Our meta-analytic results indicate that the effects of fighting ability and age may be largely dependent on each other because the CIs around the effect sizes of these 2 factors overlap to a large extent (Figure 1). Two experimental studies provide further support for our finding of a strong effect on fighting ability. Gonzalez et al. (2002) showed that males increased the proportion of agonistic interactions won after their bib sizes were experimentally enlarged. Veiga (1993) also reported that males
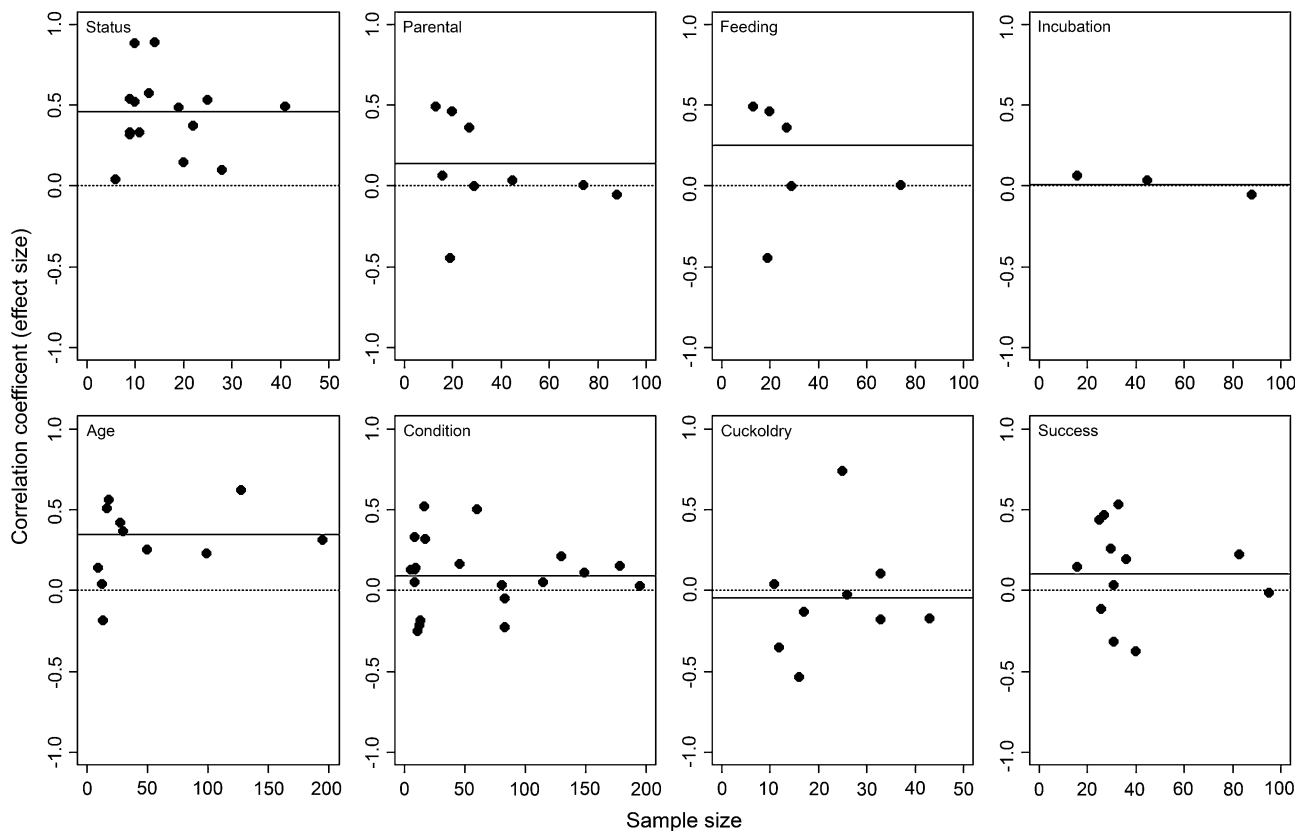
Figure 2
Funnel plots for the 8 variables used in meta-analyses. Dashed lines represent zero effects and solid lines indicate mean effect size estimates from the meta-analyses.

with enlarged bibs acquired more nest sites than did control males or males with reduced bibs. Although the exact proximate mechanisms determining bib expression are still not clear, several studies support the idea that a melanin-based badge of status in house sparrows is a condition-dependent trait (Jawor and Breitwisch 2003; Griffith et al. 2006; see below for more discussion).

Body condition (mass-related condition indices) showed weak but statistical significant association with bib size ($r = 0.093$, Table 2, Figure 1). Green (2001) emphasized the potential inappropriateness of mass-related body condition indices as they may not capture the size of a bird's real energy store. It is possible, however, that mass-related body condition indices may reflect some component of condition that, in turn, correlates with bib size, supporting the idea that bib size is a condition-dependent trait (see Is the bib size sexually selected?). However, it should be noted that "body condition" here is, at best, a weak correlate of bib size. Furthermore, the significant heterogeneity suggests that a covariate may exist, which may have influenced the estimate of effect size.

The meta-analytical results of the other associations (parental ability, cuckoldry, and reproductive success) revealed nonsignificant small effects and/or large CIs, thus providing little evidence that any of these factors is directly or indirectly related to bib size (see Is the bib size sexually selected? for more on cuckoldry and reproductive success). There is little evidence of an association between parental care (either nestling provisioning or incubation) and bib size. Theoretical models predict the evolution of traits that indicate parental quality in males of species with biparental care (Wolf et al. 1997; Kokko 1998). At the same time, theory also predicts that

if the costs of advertising a direct benefit outweigh the benefits, then the signal of parental care will not evolve (Kokko 1998). It seems that, on current evidence, bib size does not signal parental care in house sparrows.

Taken together, we conclude that the function of bib size variation is to signal the fighting ability and also, to lesser extent, the age of individuals. It may also reflect some aspect of the condition of males, but any such effect is weak.

Is the bib size sexually selected?

Sexual selection can act on a trait when there is variation in reproductive success in relation to that trait (Andersson 1994). Our meta-analytical review provided little evidence for sexual selection acting on bib size primarily because the relationship between reproductive success and bib size was not significantly different from zero (note, however, that one or more unidentified moderator variables may have influenced this estimate). This is surprising given the evidence that the male's bib size is used to establish a social hierarchy and, therefore, is likely to be important in male–male competition. However, it does not seem that this effect of bib size is translated into variability in reproductive success. This finding parallels what has been found in studies of the size of the red-and-yellow epaulet in red-winged blackbirds *Agelaius phoeniceus*, another model species of sexual selection (Searcy and Yasukawa 1995). There is evidence that epaulet size affects the outcome of male–male competition over territories, whereas there is little evidence that variation in reproductive success occurs in relation to epaulet size (reviewed in Searcy and Yasukawa 1995; see also Westneat 2006). There

may be substantial costs that cancel out the advantages of large bib size in male–male competition, which we have yet to identify.

Evidence for the other component of sexual selection, female choice, in house sparrows is mixed, and most studies have been observational or correlational (Møller 1988, 1989; Veiga 1993; Kimball 1996; Griffith et al. 1999b; Ockendon 2003). Furthermore, a pair of house sparrows usually stays together in subsequent breeding seasons if both of them are still alive (Anderson 2006; Nakagawa S, personal observation), even though male bib size changes every year, sometimes to a great extent (Griffith 2000). Therefore, female choice in relation to bib size in this species may be a weak selective force.

Another aspect of sexual selection, which is not explored properly in our analysis, is "hidden" sexual selection through EPP. Very few studies have investigated the relationship between gaining EPP and bib size; the information on paternity gain is as important as paternity loss (cuckoldry) to reveal the direction of sexual selection on bib size through EPP (see Griffith et al. 2002). We would predict a positive relationship between gaining EPP gain and bib size, based on two findings: 1) older males gain more EPP (Wetton et al. 1995; Ockendon 2003) and 2) older males tend to have larger bibs. However, Ockendon (2003) did not find significant differences in bib size between males that gained EPP and those that did not. It may be that the effect size of EPP gain in relation to bib size is small (the results were replicated, Nakagawa S, unpublished data). Therefore, there is currently little evidence for hidden sexual selection through EPP, as is also the case for apparent sexual selection through fledging success or recruits.

A real problem in determining the role of the bib size in sexual selection is that its size may have very low or no heritability, meaning that the additive genetic variance underpinning bib size is zero or close to zero (Griffith et al. 1999a; Ockendon 2003; cf. Møller 1989). The variation we observe in the bib size of house sparrows may be largely due to environmental factors during and prior to a molt, such as nutritional conditions (Veiga and Puerta 1996; Poston et al. 2005), rearing conditions (Griffith et al. 1999a), social conditions (McGraw et al. 2003; Jensen et al. 2006), climatic conditions (Jensen et al. 2006), and preceding parental effort (Griffith 2000). Therefore, there is ample evidence that the bib size is a condition-dependent trait, a feature in common with most sexually selected traits (reviewed in Andersson 1994; Jennions et al. 2001).

According to the model by Rowe and Houle (1996), condition-dependent sexually selected traits harbor large amounts of additive genetic variance. In their model, condition is equated to the total pool of resources acquired by an individual, so that condition reflects an internal property that contributes to individual fitness. If the condition revealed by a secondary sexual trait is determined purely by environmental factors, as in the model by Price et al. (1993), the secondary sexual trait is expected to indicate direct benefits. However, the house sparrow's bib does not seem to fit either of these models because 1) in the former model, we still expect heritability in bib size and 2) in the latter model, bib size has to be an indicator of parental quality. There may be heritable indirect genetic effects that influence the expression of the bib (sensu Wolf et al. 1997, reviewed in Qvarnström and Price 2001), but this is unlikely because indirect genetic effects should act on offspring traits that indicate the extent of indirect genetic effects (i.e., the bib size has to be an indicator of parental quality for this model to work). It may be that wearing an honest signal (an appropriate bib size for an individual's condition) is optimal and that males with different bib sizes, which are subject to change every year, achieve similar fitness over their lifetime.

Finally, our meta-analysis on the available data and our examination of current theory suggest that the bib size of male house sparrows is currently under little or, at most, weak sexual selection, although bib size in male house sparrows has been the focus of investigations in sexual selection for many years (Anderson 2006). It is noteworthy that a comparable conclusion concerning the relationship between sexually dimorphic traits and sexual selection theory has been reached in another model bird species, the red-winged blackbird (Searcy and Yasukawa 1995; Westneat 2006).

REFERENCES

Anderson TR. 2006. Biology of the ubiquitous house sparrow. Oxford: Oxford University Press.

Andersson M. 1994. Sexual selection. Princeton (NJ): Princeton University Press.

Andersson V, Simmons LW. 2006. Sexual selection and mate choice. Trends Ecol Evol. 21:296–302.

Bartlett TL, Mock DW, Schwagmeyer PL. 2005. Division of labor: incubation and biparental care in house sparrows (*Passer domesticus*). Auk. 122:835–842.

Bókony V, Lendvai AZ, Liker A. 2006. Multiple cues in status signalling: the role of wingbars in aggressive interactions of male house sparrows. Ethology. 112:947–954.

Brooks R, Kemp DJ. 2001. Can older males deliver the good genes? Trends Ecol Evol. 16:308.

Cassey P, Ewen J, Blackburn T, Møller A. 2004. A survey of publication bias within evolutionary ecology. Proc R Soc Lond B. 271:S451–S454.

Cohen J. 1988. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates.

Cooper H, Hedges LV, editors. 1994. The handbook of research synthesis. New York: Russell Sage Foundation.

Cordero PJ, Wetton JH, Parkin DT. 1999. Extra-pair paternity and male badge size in the house sparrow. J Avian Biol. 30:97–102.

Dawkins R, Krebs JR. 1978. Animal signals: information or manipulation. In: Krebs JR, Davies NB, editors. Behavioural ecology: an evolutionary approach. Oxford: Blackwell. p. 282–309.

Dubois F, Cézilly F. 2002. Breeding success and mate retention in birds: a meta-analysis. Behav Ecol Sociobiol. 52:357–364.

Fox CW, Wolf JB, editors. 2006. Evolutionary genetics: concepts and case studies. Oxford: Oxford University Press.

Gonzalez G, Sorci G, Møller AP, Ninni P, Haussy C, de Lope F. 1999. Immunocompetence and condition-dependent sexual advertisement in male house sparrows (*Passer domesticus*). J Anim Ecol. 68:1225–1234.

Gonzalez G, Sorci G, Smith LC, de Lope F. 2002. Social control and physiological cost of cheating in status signalling male house sparrows (*Passer domesticus*). Ethology. 108:289–302.

Green AJ. 2001. Mass/length residuals: measures of body condition or generators of spurious results? Ecology. 82:1473–1483.

Griffith S. 1998. Sexual selection in the house sparrow, *Passer domesticus* [PhD thesis]. [Leicester]: University of Leicester.

Griffith SC. 2000. A trade-off between reproduction and a condition-dependent sexually selected ornament in the house sparrow *Passer domesticus*. Proc R Soc Lond B. 267:1115–1119.

Griffith SC, Owens IPF, Burke T. 1999a. Environmental determination of a sexually selected trait. Nature. 400:358–360.

Griffith SC, Owens IPF, Burke T. 1999b. Female choice and annual reproductive success favour less-ornamented male house sparrows. Proc R Soc Lond B. 266:765–770.

Griffith SC, Owens IPF, Thuman KA. 2002. Extra pair paternity in birds: a review of interspecific variation and adaptive function. Mol Ecol. 11:2195–2212.

Griffith SC, Parker TH, Olson VA. 2006. Melanin- versus carotenoid-based sexual signals: is the difference really so black and red? Anim Behav. 71:749–763.

Hedges L, Olkin I. 1985. Statistical methods for meta-analysis. New York: Academic Press.

Hein WK, Westneat DF, Poston JP. 2003. Sex of opponent influences response to a potential status signal in house sparrows. Anim Behav. 65:1211–1221.

Hill GE. 2002. A red bird in a brown bag: the function and evolution of colorful plumage in the house finch. Oxford: Oxford University Press.

Hunt M. 1997. How science takes stock: the story of meta-analysis. New York: Russell Sage Foundation.

Jawor JM, Breitwisch R. 2003. Melanin ornaments, honesty, and sexual selection. Auk. 120:249–265.

Jennions MD, Møller AP, Petrie M. 2001. Sexually selected traits and adult survival: a meta-analysis. Q Rev Biol. 76:3–36.

Jensen H, Sæther BE, Ringsby TH, Tufto J, Griffith SC, Ellegren H. 2004. Lifetime reproductive success in relation to morphology in the house sparrow, *Passer domesticus*. J Anim Ecol. 73:599–611.

Jensen H, Svorkmo-Lundberg T, Ringsby TH, Sæther B-E. 2006. Environmental influence and cohort effects in a sexual ornament in the house sparrow, *Passer domesticus*. Oikos. 114:212–224.

Kimball RT. 1995. Sexual selection in house sparrows [PhD thesis]. [Albuquerque (NM)]: University of New Mexico.

Kimball RT. 1996. Female choice for male morphological traits in house sparrows, *Passer domesticus*. Ethology. 102:639–648.

Kokko H. 1998. Should advertising parental care be honest? Proc R Soc Lond B. 265:1871–1878.

Kopisch AD, Schwagmeyer PL, Mock DW. 2005. Individual consistency in parental effort across multiple stages of care in the house sparrow, *Passer domesticus*. Ethology. 111:1062–1070.

Liker A, Barta Z. 2001. Male badge size predicts dominance against females in house sparrows. Condor. 103:151–157.

Lindström KM, Hasselquist D, Wikelski M. 2005. House sparrows (*Passer domesticus*) adjust their social status position to their physiological costs. Horm Behav. 48:311–320.

Lipsey MW, Wilson DB. 2001. Practical meta-analysis. Beverly Hills (CA): Sage.

McGraw KJ, Dale J, Mackillop EA. 2003. Social environment during moult and the expression of melanin-based plumage pigmentation in male house sparrows (*Passer domesticus*). Behav Ecol Sociobiol. 53:116–122.

Møller AP. 1987a. Social control of deception among status signalling house sparrows *Passer domesticus*. Behav Ecol Sociobiol. 20: 307–311.

Møller AP. 1987b. Variation in badge size in male house sparrows *Passer domesticus*: evidence for status signalling. Anim Behav. 35: 1637–1644.

Møller AP. 1988. Badge size in the house sparrow *Passer domesticus*: effects of intra- and intersexual selection. Behav Ecol Sociobiol. 22:373–378.

Møller AP. 1989. Natural and sexual selection on a plumage signal of status and on morphology in house sparrows, *Passer domesticus*. J Evol Biol. 2:125–140.

Møller AP, Erritzøe J. 1988. Badge, body and testes size in house sparrows, *Passer domesticus*. Ornis Scand. 19:72–73.

Møller AP, Jennions MD. 2001. Testing and adjusting for publication bias. Trends Ecol Evol. 16:580–586.

Møller AP, Ninni P. 1998. Sperm competition and sexual selection: a meta-analysis of paternity studies of birds. Behav Ecol Sociobiol. 43:345–358.

Nakagawa S. 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. Behav Ecol. 15:1044–1045.

Nakagawa S, Cuthill IC. Forthcoming 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. Biol Rev.

Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T. Forthcoming 2007. Does the badge of status influence parental care and investment in house sparrows? An experimental test. Oecologia. doi: 10.1007/s00442-007-0765-4

Ockendon N. 2003. Genetic diversity and sexual selection in an isolated population of house sparrows [PhD thesis]. [Sheffield]: University of Sheffield.

Owens IPF. 2006. Where is behavioural ecology going? Trends Ecol Evol. 21:356–361.

Palmer AR. 2000. Quasi-replication and the contract of error: lessons from sex ratios, heritabilities and fluctuating asymmetry. Annu Rev Ecol Syst. 31:441–480.

Pinheiro JC, Bates DM. 2000. Mixed-effects models in S and S-Plus. New York: Springer.

Poston JP, Hasselquist D, Stewart IRK, Westneat DF. 2005. Dietary amino acids influence plumage traits and immune responses of male house sparrows, *Passer domesticus*, but not as expected. Anim Behav. 70:1171–1181.

Price T, Schluter D, Heckman NE. 1993. Sexual selection when the female directly benefits. Biol J Linn Soc. 48:187–211.

Qvarnström A, Price TD. 2001. Maternal effects, paternal effects and sexual selection. Trends Ecol Evol. 16:95–100.

R Development Core Team. 2006. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing.

Raudenbush SW. 1994. Random effects models. In: Cooper H, Hedges LV, editors. The handbook of research synthesis. New York: Russell Sage Foundation.

Ritchison G. 1985. Plumage variability and social status in captive house sparrows. Kentucky Warbler. 61:39–42.

Riters LV, Teague DP, Schroeder MB. 2004. Social status interacts with badge size and neuroendocrine physiology to influence sexual behavior in male house sparrows (Passer domesticus). Brain Behav Evol. 63:141–150.

Roff DA. 1997. Evolutionary quantitative genetics. New York: Chapman & Hall.

Rohwer SA. 1975. The social significance of avian winter plumage variability. Evolution. 29:593–610.

Rosenberg MS, Adams DC, Gurevitch J. 2000. Metawin: statistical software for meta-analysis. Sunderland (MA): Sinauer.

Rosenthal R. 1994. Parametric measures of effect size. In: Cooper H, Hedges LV, editors. The handbook of research synthesis. New York: Russell Sage Foundation. p. 231–244.

Rowe L, Houle D. 1996. The lek paradox and the capture of genetic variance by condition dependent traits. Proc R Soc Lond B. 263: 1415–1421.

Searcy WA, Nowicki S. 2005. The evolution of animal communication: reliability and deception in signaling systems. Princeton (NJ): Princeton University Press.

Searcy WA, Yasukawa K. 1995. Polygyny and sexual selection in red-winged blackbirds. Princeton (NJ): Princeton University Press.

Solberg EJ, Ringsby TH. 1997. Does badge size signal status in small island populations of house sparrows, *Passer domesticus*. Ethology. 103:177–186.

Stewart IRK. 1999. Aspects of the breeding ecology of the house sparrow, *Passer domesticus* [PhD thesis]. [Leicester]: University of Leicester.

Stewart IRK, Hanschu RD, Burke T, Westneat DF. 2006. Tests of ecological, phenotypic, and genetic correlates of extra-pair paternity in the house sparrow. Condor. 108:399–413.

Thompson B. 2002. "Statistical," "practical," and "clinical": how many kinds of significance do counselors need to consider? J Couns Dev. 80:64–71.

Václav R, Hoi H. 2002. Different reproductive tactics in house sparrows signalled by badge size: is there a benefit to being average? Ethology. 108:569–582.

Václav R, Hoi H, Blomqvist D. 2002. Badge size, paternity assurance behaviours and paternity losses in male house sparrows. J Avian Biol. 33:315–318.

Veiga JP. 1993. Badge size, phenotypic quality, and reproductive success in the house sparrow: a study on honest advertisement. Evolution. 47:1161–1170.

Veiga JP. 1996. Permanent exposure versus facultative concealment of sexual traits: an experimental study in the house sparrow. Behav Ecol Sociobiol. 39:345–352.

Veiga JP, Boto L. 2000. Low frequency of extra-pair fertilisations in house sparrows breeding at high densities. J Avian Biol. 31:237–244.

Veiga JP, Puerta P. 1996. Nutritional constraints determine the expression of a sexual trait in the house sparrow, *Passer domesticus*. Proc R Soc Lond B. 263:229–234.

Voltura KM, Schwagmeyer PL, Mock DW. 2002. Parental feeding rates in the house sparrow, *Passer domesticus*: are larger-badged males better fathers? Ethology. 108:1011–1022.

Westneat DF. 2006. No evidence of current sexual selection on sexually dimorphic traits in a bird with high variance in mating success. Am Nat. 167:E171–E189.

Wetton JH, Burke T, Parkin DT, Cairns E. 1995. Single-locus DNA fingerprinting reveals that male reproductive success increases with age through extra-pair paternity in the house sparrow (*Passer domesticus*). Proc R Soc Lond B. 260:91–98.

Whitekiller RR, Westneat DF, Schwagmeyer PL, Mock DW. 2000. Badge size and extra-pair fertilizations in the house sparrow. Condor. 102:342–348.

Wolf JB, Moore AJ, Brodie ED 3rd. 1997. The evolution of indicator traits for parental quality: the role of maternal and paternal effects. Am Nat. 150:639–649.