S1 Appendix: Information about data used in the study.

S1 Table. Summary of key differences in methodology among all studies (published and unpublished) testing the relationship between dominance rank and bib size in male house sparrows (N = 19 studies).

Variable	Levels	Number of studies	Order of preference[1]
Group composition	Males and females	11	-
	Males only	8	-
Resource competed for	Food only	12	-
	Food, water and roosting place	6	-
	Female	1	-
Type of interactions	Aggressive only	12	-
	Aggressive and non-aggressive	7	-
Interactions recording protocol	Live observations	11	-
	Video	6	-
	Live and video observations	2	-
Type of bib size measured	Visible	14	1
	Hidden	2	2
	Both	3	-
Beak angle during measurement	90°	8	1
	180°	3	2
	Both	1	-
	Unknown	7	-
Season	Non-breeding	13	-
	Breeding	5	-
	Both	1	-
Study location	Captive	16	-
	Wild	2	-
	Both	1	-

[1] Order of preference used for the analyses (see main text). The order of preference was determined based on how frequently the method was used in previous studies.

S2 Table. List of the different methods used to estimate bib size in published and unpublished studies testing the relationship between dominance rank and bib size in male house sparrows (N = 19 studies). Note that some studies used more than one method to estimate bib size.

Method to estimate bib size	Number of times used	Order of preference[1]
Area[#]	8	1
Møller 1987's equation	6	2
Length and width*	3	2
Length only	2	3
Møller 1987's drawings	1	4
Veiga 1993's equation	1	5

[#]Area was measured from pictures (N = 5 studies), by tracing and weighing (N = 2 studies), and by tracing and ranking (N = 1 study). *If length and width were available, we estimated bib area using the equation in Møller [2]. [1] Order of preference used for the analyses (see main text). The order of preference was determined based on how frequently the method was used in previous studies.

S3 Table. List of the different methods used to infer dominance rank from dyadic interactions in published studies that tested the relationship between dominance rank and bib size in male house sparrows (N = 13 published studies, 11 different methods). Note that some studies used more than one method to estimate dominance rank and that unpublished studies are not included in this summary.

Method to infer dominance rank	Number of times used	Order of preference[1]
Proportion of contests won	4	4
Proportion of initiated contests	3	5
Kendall's linearity index	2	3
Proportion of contests won per dyad	2	6
Proportion of initiated contests won	2	6
David's score	1	1
I&SI: de Vries 1998	1	2
Landau's linearity index	1	3
Proportion of the received attacks won	1	7
Proportion of birds dominated	1	7
Proportion of contests won per dyad + linear assumption	1	7

[1] Order of preference used for the analyses (see main text). The order of preference was determined based on both how frequently the method was used in previous studies and by taking into account the (expected) performance of each of the methods. First, higher order of preference was assigned to methods specifically designed for inferring linear dominance hierarchies (i.e. David's score, I&SI, Landau's and Kendall's linearity indices). We used the information available in [4] to rank David's score and I&SI as first and second methods in preference, respectively. Second, we ranked the remaining (proportion-based) methods based on how frequently they were used in previous studies. Importantly, the order of preference was chosen prior to conducting any statistical analysis, and thus, method selection was blind to the outcome of the analyses.

S4 Table. Additional comments on some of the published studies included in the meta-analysis.

Reference	Comments
Ritchison (1985) [5]	According to the original publication, the total number of birds studied was 35, as opposed to the 25 individuals used in the meta-analyses of Nakagawa *et al.* [6] and Santos *et al.* [7].
Hein *et al.* (2003) [8]	The total number of birds included in our re-analysis of the primary data is smaller than that presented in the original publication. This is because our re-analysis only included fully identified individuals (e.g. birds missing rings could not be included).
Dolnik & Hoi (2010) [9]	32 males were selected for the experiment, but one bird was excluded before the start of the experiment. Thus, *n* was set to 31 individuals for this study.
Buchanan *et al.* (2010) [10]	96 birds were separated in 24 aviaries of four individuals each. The final *n* of several aviaries was less than four individuals, and therefore, these aviaries were not included in our meta-analyses (see main text, section “Materials and Methods”).
Rojas Mora *et al.* (2016) [11]	According to the primary data, one male did not interact, and thus, *n* was set to 59 individuals for Appendix S3.

S5 Table. Data descriptions for the unpublished data analysed in the meta-analysis.

Study ID*	Data description
14	88 individuals were separated into 4 captive mixed-sex groups. Live observations after mild deprivation were conducted to record agonistic dyadic interactions (i.e. fights) over (mostly) food for around one week in Feb 2003 (total = 1,563 fights). Bib length and bib width were measured for each male before the dominance observations using a ruler. More information can be found in [12] and [13].
15	61 individuals were separated into 3 captive mixed-sex groups. Live observations after mild deprivation were conducted to record agonistic dyadic interactions (i.e. fights) over (mostly) food between Oct and Dec 2005 (two groups) and 2006 (one group; total = 2,003 fights). Bib area was measured for each male using standardized pictures taken after the dominance observations. More information can be found in [14] and [13].
16	60 individuals were separated into 4 captive mixed-sex groups. Live and video observations after mild deprivation were conducted to record agonistic dyadic interactions (i.e. fights) over (mostly) food for around two weeks per group between Oct 2007 and Feb 2008 (total = 6,641 fights). Bib length and bib width were measured for each male before the dominance observations using a ruler. More information can be found in [15] and [13].
17	96 males were separated into 4 captive male-only groups. Videos after mild deprivation were taken to record agonistic dyadic interactions (i.e. fights) over food for 10 days between Oct and Dec 2014 (total = 3,776 fights). Bib area was measured several times for each male (median = 3 times/male, range = 2 to 6) using standardized pictures taken from Oct to Dec 2014, and the mean bib area of each individual was used in the analyses.
18	453 individuals (215 females and 238 males) were observed in seven discrete sampling events in a wild population of house sparrows. Videos were taken to record agonistic dyadic interactions (i.e. fights)

	over food for 20 days between Nov 2013 and Dec 2016 (total = 11,063 fights). Bib length was measured several times for each male (median = 1 time/male, range = 1 to 6) from Nov 2013 to Dec 2016 using a calliper, and the mean bib area of each individual in each sampling event was used in the analyses.
19	128 individuals were separated into 16 captive mixed-sex groups. Live observations after mild deprivation were conducted to record agonistic dyadic interactions (i.e. supplants and hold-offs) over food between March and April 2005 (total = 5,496 fights). Bib length and bib width were measured for each male before the dominance observations using a calliper as in [16].

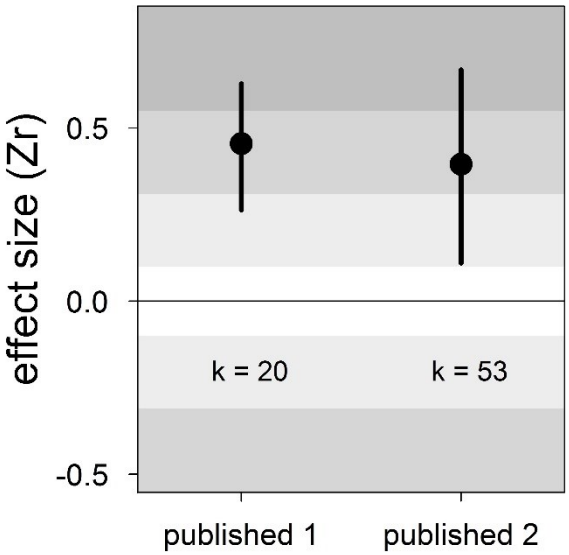
* Study ID corresponding to Table 1 in main text.

S2 Appendix: Meta-analyses based on published studies only.

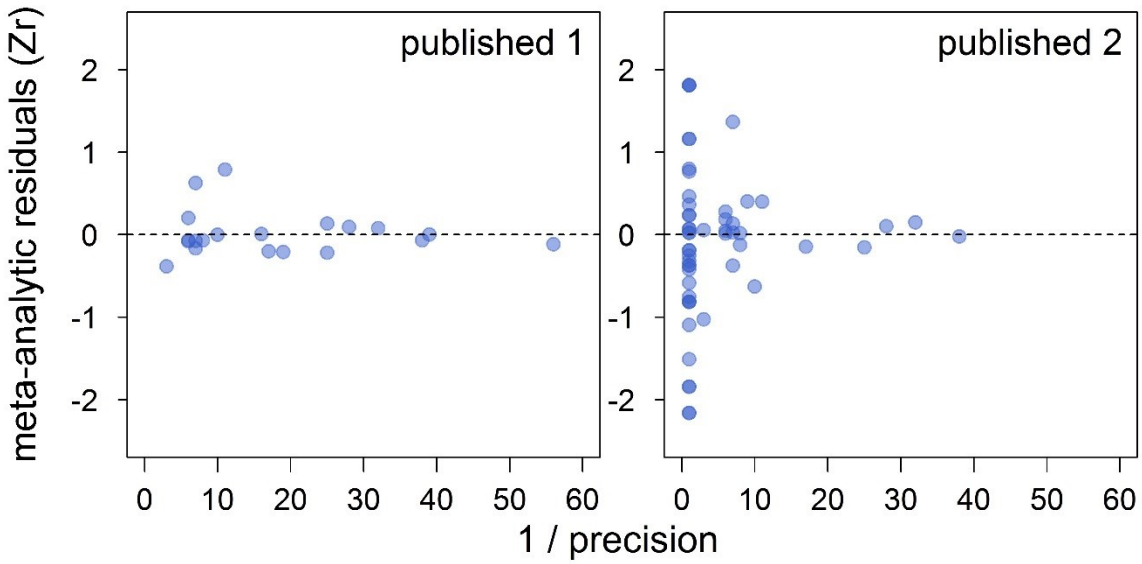
S6 Table. Results of two multilevel meta-analyses to test the relationship between dominance rank and bib size in male house sparrows based on published studies only. Published 1 includes published effect sizes obtained from summary data, whereas published 2 includes published re-analysed effect sizes together with the remaining published effect sizes obtained from summary data. Additionally, the results of the Egger's regressions are shown. Estimates are presented as standardized effect sizes using Fisher's transformation (*Zr*). Credible intervals not overlapping zero are highlighted in bold.

Meta-analysis	***k***	**Meta-analytic mean [95% CrI]**	$I^2_{\text{population ID}}$ **[95% CrI] (%)**	$I^2_{\text{study ID}}$ **[95% CrI] (%)**	I^2_{overall} **[95% CrI] (%)**	**Egger's regression [95% CrI]**
published 1	20	**0.45 [0.26,0.63]**	17 [0,51]	17 [0,53]	46 [15,78]	0.42 [-0.73,1.48]
published 2	53	**0.40 [0.11,0.67]**	14 [0,46]	13 [0,42]	46 [17,72]	-0.25 [-0.73,0.26]

k = number of estimates; CrI = credible intervals; I^2 = heterogeneity.



S1 Figure. Forest plot showing the overall effect size of the relationship between dominance rank and bib size in male house sparrows based on published studies only. Published 1 includes published effect sizes obtained from summary data, whereas published 2 includes published re-analysed effect sizes together with the remaining published effect sizes obtained from summary data. We show posterior means and 95% credible intervals from multilevel meta-analyses. Estimates are presented as standardized effect sizes using Fisher's transformation (*Zr*). Light, medium and dark grey show small, medium and large effect sizes, respectively [24]. *k* is the number of estimates.



S2 Figure. Funnel plots of the meta-analytic residuals against their precision for the meta-analyses based on published studies only. Published 1 includes published effect sizes obtained from summary data, whereas published 2 includes published re-analysed effect sizes together with the remaining published effect sizes obtained from summary data. Estimates are presented as standardized effect sizes using Fisher's transformation (*Zr*). Precision = square root of the inverse of the variance.

Supplementary File 1: Meta-analysis challenges a textbook example of status signaling and demonstrates publication bias. Sánchez-Tójar et al. More in https://osf.io/cwkxb/ [1]

S3 Appendix: Power analysis based on meta-analytic mean.

R code used and explanations:

First, we need to clear up the memory and load the pwr library.

```
# clear memory
rm(list=ls())

# package needed
library(pwr)
```

Furthermore, we create a function to transform Zr values into r values. This is because our meta-analysis was based on Zr values, but the power analysis is based on r values.

```
# function to convert Zr to r
Zr.to.r<-function(Zr){
  r<-(exp(2*Zr)-1)/(exp(2*Zr)+1)
}
```

Power analysis

Next, we estimated the sample size necessary to find an effect size as small as the one estimated by our meta-analysis (Zr = 0.20). We used a significance level of 0.05, and the recommended 80% statistical power [24].

```
pwr.r.test(r = Zr.to.r(0.20), sig.level = 0.05, power = 0.8)

##
##      approximate correlation power calculation (arctangh transformation)
##
##              n = 198.3401
##              r = 0.1973753
##      sig.level = 0.05
##          power = 0.8
##    alternative = two.sided
```

This shows that we would need the dominance rank and bib size of 198 individuals to find a significant r correlation of 0.20 with an 80% statistical power.

Additionally, we estimated the across-study statistical power of the tests on status signaling in house sparrows to compare it the overall statistical power found in the behavioural ecology literature [25].

```
pwr.r.test(n = 10, r = Zr.to.r(0.20), sig.level = 0.05)

##
##      approximate correlation power calculation (arctangh transformation)
##
##              n = 10
##              r = 0.1973753
##      sig.level = 0.05
```

```
##          power = 0.08474157
##    alternative = two.sided
```

This shows that the statistical power of the sparrow literature on status signaling is as low as 8.5%, which is alarming.

References

1. Sánchez-Tójar A, Nakagawa S, Sánchez-Fortún M, Martin DA, Ramani S, Girndt A, et al. Supporting information for "Meta-analysis challenges a textbook example of status signalling: evidence for publication bias" [Internet]. 2018. doi:10.17605/OSF.IO/CWKXB

2. Møller A. Variation in badge size in male house sparrows *Passer domesticus*: evidence for status signaling. Anim Behav. 1987;35: 1637–1644. doi:10.1016/S0003-3472(87)80056-8

3. Veiga JP. Badge size, phenotypic quality, and reproductive success in the house sparrow: a study on honest advertisement. Evolution. 1993;47: 1161–1170. doi:10.1111/j.1558-5646.1993.tb02143.x

4. Sánchez-Tójar A, Schroeder J, Farine DR. A practical guide for inferring reliable dominance hierarchies and estimating their uncertainty. J Anim Ecol. 2018;87: 594–608. doi:10.1111/1365-2656.12776

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

6. Nakagawa S, Ockendon N, Gillespie DOS, Hatchwell BJ, Burke T. Assessing the function of house sparrows' bib size using a flexible meta-analysis method. Behav Ecol. 2007;18: 831–840. doi:10.1093/beheco/arm050

7. Santos ESA, Scheck D, Nakagawa S. Dominance and plumage traits: meta-analysis and metaregression analysis. Anim Behav. 2011;82: 3–19. doi:10.1016/j.anbehav.2011.03.022

8. Hein WK, Westneat DF, Poston JP. Sex of opponent influences response to a potential status signal in house sparrows. Anim Behav. 2003;65: 1211–1221. doi:http://dx.doi.org/10.1006/anbe.2003.2132

9. Dolnik O V., Hoi H. Honest signalling, dominance hierarchies and body condition in House Sparrows *Passer domesticus* (Aves: *Passeriformes*) during acute coccidiosis. Biol J Linn Soc. 2010;99: 718–726. doi:10.1111/j.1095-8312.2010.01370.x

10. Buchanan KL, Evans MR, Roberts ML, Rowe L, Goldsmith AR. Does testosterone determine dominance in the house sparrow *Passer domesticus*? An experimental test. J Avian Biol. 2010;41: 445–451. doi:10.1111/j.1600-048X.2010.04929.x

11. Rojas Mora A, Meniri M, Glauser G, Vallat A, Helfenstein F. Badge size reflects sperm oxidative status within social groups in the house sparrow *Passer domesticus*. Front Ecol Evol. 2016;4: 67. doi:10.3389/fevo.2016.00067

12. Lendvai ÁZ, Barta Z, Liker A, Bókony V. The effect of energy reserves on social foraging: hungry sparrows scrounge more. Proc R Soc B. 2004;271: 2467–2472. doi:10.1098/rspb.2004.2887

13. Bókony V, Seress G, Nagy S, Lendvai ÁZ, Liker A. Multiple indices of body condition reveal no negative effect of urbanization in adult house sparrows. Landsc Urban Plan. 2012;104: 75–84. doi:10.1016/j.landurbplan.2011.10.006

14. Tóth Z, Bókony V, Lendvai ÁZ, Szabó K, Pénzes Z, Liker A. Kinship and aggression: do house sparrows spare their relatives? Behav Ecol Sociobiol. 2009;63: 1189–1196. doi:10.1007/s00265-009-0768-8

15. Bókony V, Kulcsár A, Liker A. Does urbanization select for weak competitors in house sparrows? Oikos. 2010;119: 437–444. doi:10.1111/j.1600-0706.2009.17848.x

16. Morrison EB, Kinnard TB, Stewart IRK, Poston JP, Hatch MI, Westneat DF. The links between plumage variation and nest site occupancy in male house sparrows. Condor. 2008;110: 345–353. doi:10.1525/cond.2008.8470

17. Rohwer S. The social significance of avian winter plumage variability. Evolution (N Y). 1975;29: 593–610. doi:10.2307/2407071

18. Alarcón-Nieto G, Graving JM, Klarevas-Irby JA, Maldonado-Chaparro AA, Mueller I, Farine DR. An automated barcode tracking system for behavioural studies in birds. Methods Ecol Evol. 2018; doi:10.1111/2041-210X.13005

19. Drews C. The concept and definition of dominance in animal behaviour. Behaviour. 1993;125: 283–313. doi:10.1163/156853993X00290

20. Solberg EJ, Ringsby TH. Does male badge size signal status in small island populations of house sparrows, *Passer domesticus*? Ethology. 1997;103: 177–186. doi:10.1111/j.1439-0310.1997.tb00114.x

21. Andersson S, Åhlund M. Hunger affects dominance among strangers in house sparrows. Anim Behav. 1991;41: 895–897. doi:10.1016/S0003-3472(05)80356-2

22. Riters L V., Teague DP, Schroeder MB. Social status interacts with badge size and neuroendocrine physiology to influence sexual behavior in male house sparrows (*Passer domesticus*). Brain Behav Evol. 2004;63: 141–150. doi:10.1159/000076240

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

doi:10.1016/j.yhbeh.2005.04.002

24. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New Jersey: Taylor & Francis Inc; 1988.

25. Jennions M, Møller A. A survey of the statistical power of research in behavioral ecology and animal behavior. Behav Ecol. 2003;14: 438–445. doi:10.1093/beheco/14.3.438