

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/281480713>

Bayesian estimation of competitiveness in male house finches: Small-billed males are more competitive

Article in *Animal Behaviour* · October 2015

Impact Factor: 3.14 · DOI: 10.1016/j.anbehav.2015.07.027

READS

124

5 authors, including:



[Mathieu Giraudeau](#)

Arizona State University

31 PUBLICATIONS 321 CITATIONS

[SEE PROFILE](#)



[Kevin McGraw](#)

Arizona State University

217 PUBLICATIONS 6,560 CITATIONS

[SEE PROFILE](#)



Bayesian estimation of competitiveness in male house finches: small-billed males are more competitive



Masaru Hasegawa ^{a,*}, Mathieu Giraudeau ^{b,1}, Nobuyuki Kutsukake ^a, Mamoru Watanabe ^c, Kevin J. McGraw ^b

^a Department of Evolutionary Studies of Biosystems, The Graduate University for Advanced Studies, Hayama-machi, Japan

^b School of Life Sciences, Arizona State University, Tempe, AZ, U.S.A.

^c Graduate School of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Japan

ARTICLE INFO

Article history:

Received 9 March 2015

Initial acceptance 7 May 2015

Final acceptance 7 July 2015

Available online

MS. number: 15-00185R

Keywords:

aggression

Bayesian approach

bill size

dominance

Haemorhous mexicanus

latent variable

A common method for assessing the competitiveness of animals, especially in birds, is to pit pairs of unfamiliar individuals against each other in contests for limited resources under controlled conditions. Although this approach can clarify dominant–subordinate relationships within dyads, it is often difficult to determine competitiveness for a large group of individuals. Here, by using Bayesian statistical inference and 'hypothetical competition groups', which are formed when individuals experience a series of paired contests, we estimated social competitiveness of male house finches, *Haemorhous mexicanus*. First, Bayesian competitiveness estimates from paired contests successfully predicted future contest outcomes among four unfamiliar individuals (i.e. social dominance). When data of all rank combinations were pooled, future dominant males had, on average, higher competitiveness estimates than future subordinate males. Similarly, Bayesian statistical inference and hypothetical competition groups identified accurately the relative competitiveness of four subgroups of males (i.e. colourful and drab males from urban and rural sites), which matched the result of direct contests when they were all put into the same cage. This consistency reinforces the validity of Bayesian competitiveness estimation based on hypothetical competition groups. Moreover, we found that the competitiveness estimate was negatively linked to male beak size in the Bayesian framework. Males with smaller bills were more competitive than those with larger bills, perhaps due to their elevated foraging motivation (i.e. limited ability to consume or husk large, valuable seeds). We argue that Bayesian competitiveness estimations, together with a series of paired contests, is a sophisticated approach for acquiring a broad understanding of social and individual competitiveness.

© 2015 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

A classic laboratory method for studying animal competition is to pair two individuals in a contest and let them vie for access to a limited resource. Such dyadic interactions are widely used, especially in birds, for studying dominant–subordinate relationships within a sex (e.g. Koivula, Lahti, Orreli, & Rytönen, 1993), in the use of status signals (e.g. McGraw & Hill, 2000a,b; Senar, 2006) and during intersexual (e.g. Belthoff & Gowaty, 1996) or interspecific interactions (e.g. Grava et al., 2012). This method is straightforward since each contestant has only one opponent and the number of aggressive encounters won and lost reveals the competitiveness of the focal individual relative to the opponent. However, this relative measure within dyads cannot be interpreted as individual

competitiveness in a larger group (i.e. in social species in which group membership is dynamic and multiple individuals can compete). Some studies (e.g. Gonzalez, Sorci, Smith, & de Lope, 2002; McGraw & Hill, 2000b) have instead investigated within-group competitions, in which possible combinations of contestants increase exponentially with group size. However, given the frequency of interactions and the (often ignored) potential for polyadic contests, including information sharing and eavesdropping (e.g. Amy & Leboucher, 2007; Loretto, Fraser, & Bugnyar, 2012; reviewed in Sherratt & Mesterton-Gibbons, 2013), competitiveness of each individual can be more challenging to estimate in these studies.

The question then remains whether competitiveness of individuals in groups/societies can be estimated from dyadic contests. In most cases, running all possible pairwise combinations of individuals through a tournament, which can clarify relative competitiveness of all group members, is unrealistic, as the number of combinations scales with $(N^2 - N)/2$; for example it would require 1225 trials in a study of 50 animals. Bayesian statistical inference may be an

* Correspondence and present address: M. Hasegawa, 1560-35 Kamiyamaguchi, Hayama-machi, Miura-gun, Kanagawa 240-0115, Japan.

E-mail address: perorobomusadiobe@gmail.com (M. Hasegawa).

¹ Present address: Biological Sciences A08, University of Sydney, Sydney, Australia.

alternative solution to this problem. This approach is more accurate and informative for estimating relative competitiveness of individuals within a group, compared to classic nonparametric statistical methods such as I&SI methods (Adams, 2005; Romero & Castellanos, 2010). Moreover, Bayesian competitiveness estimates can be compared even without aggressive encounters when a series of contests takes place. In fact, based on dominant–subordinate relationships of wild male chimpanzees, *Pan troglodytes*, across two decades, Hasegawa and Kutsukake (2015) showed that competitiveness estimates could be compared across years because of the predictable link between male age and competitiveness. Likewise, in principle, a series of paired contest trials in captivity may determine the relative competitiveness of individuals within groups. This is because a series of paired contests forms a 'hypothetical competition group', which can be defined as a set of individuals for which competitiveness can be compared to the opponents, to the opponents' opponents, and so on (Fig. 1). Because previous applications of these Bayesian approaches have focused on outdoor groups, in which most of the dyadic combinations could be observed (i.e. with few blank cells in the interaction matrix; see Adams, 2005) or involved an additional predictor (i.e. age in Hasegawa & Kutsukake, 2015), the applicability of Bayesian competitiveness to a series of captive experiments remains unknown. Such an approach may also be beneficial when investigating how phenotypic variation predicts competitiveness, particularly when the focal traits are hard to manipulate experimentally (e.g. a bird's bill, other tissues containing sensory cells; Freire, Eastwood, & Joyce, 2011).

Here, we used Bayesian inference to study the relative competitiveness of male house finches, *Haemorrhous mexicanus*, in captivity. This is a suitable study system, as there are no strong winner/loser carryover effects (Hasegawa, Ligon, Giraudeau, Watanabe, & McGraw, 2014; Hsu, Earley, & Wolf, 2006) and trial order should thus rarely affect contest outcome. Using the data of Hasegawa et al. (2014), we studied whether aggressive outcomes from paired contests can be used to estimate individual competitiveness across cages based on the hypothetical competition groups. Because our experiments were conducted over a relatively short period of time (ca. 7-day interval), temporal changes in male competitiveness should be minimal.

Given the competitiveness estimates from a set of pairwise trials, we tested whether and how well the estimates predicted outcomes of future contests among unfamiliar individuals within groups. Using colourful and drab males captured from urban and rural habitats, Hasegawa et al. (2014) conducted two dyadic contest experiments, a 'within-site experiment', in which we compared aggression between colourful and drab males from the same habitat type, and a 'between-site experiment', in which we staged agonistic encounters between urban and rural males that were matched for colour type, and one within-group contest experiment (a 'group competition experiment') in which we studied the relative aggressiveness of all four groups of birds (urban colourful, urban drab, rural colourful, and rural drab) by putting them into the same cage simultaneously. We used the first two experiments to estimate competitiveness of individual birds and then investigated whether the estimate differed between

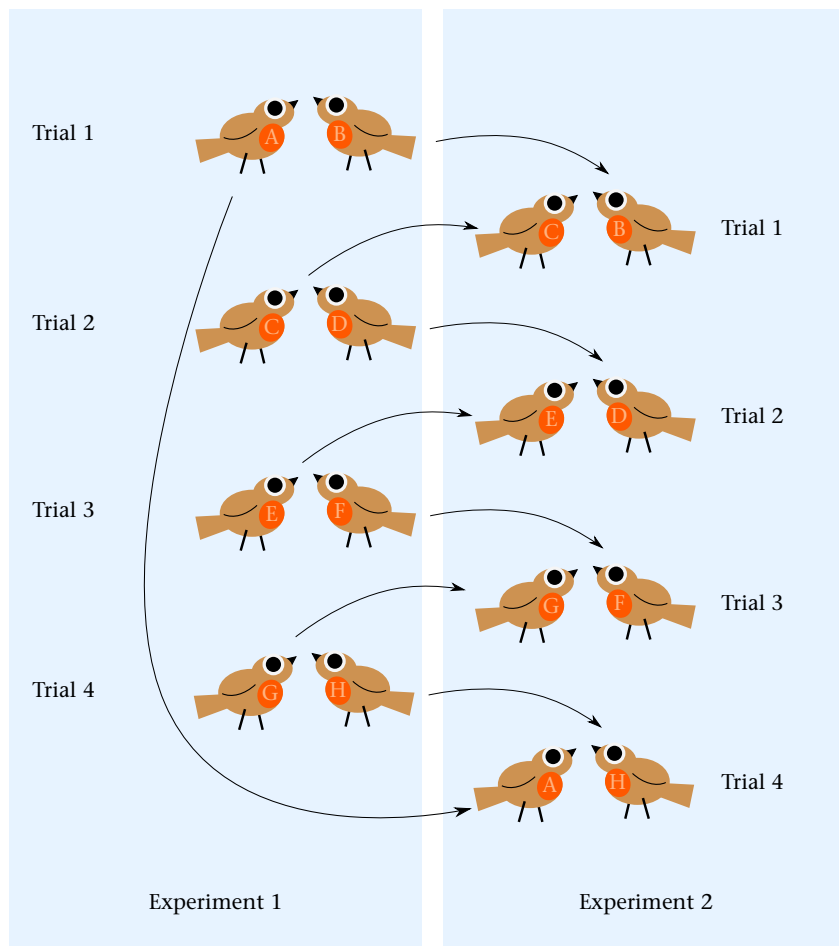


Figure 1. Competitiveness can be estimated within a 'hypothetical competition group', in which each individual competes against two different individuals, at least in principle. The same individuals are indicated by shared letters and arrows.

dominant and subordinate males in the third (group competition) experiment. As four males were pitted against each other in the group competition experiment, we also predicted that the difference in the competitiveness estimate between dominant and subordinate males should be more prominent when the rank difference is large (e.g. top-ranking versus lowest-ranking males). In addition, we predicted that our Bayesian competitiveness estimates based on two trials should mirror the pattern found in group competition experiments, in which urban colourful males were weaker competitors than the other three male categories (urban drab, rural colourful, rural drab; Fig. 4 in Hasegawa et al., 2014). Finally, because males of this species frequently use their bills during fights (McGraw, Medina-Jerez, & Adams, 2007) and because larger bills produce greater bite force (Badyaev, Young, Oh, & Addison, 2008), we predicted a positive relationship between bill size and competitiveness.

METHODS

General Description

We reanalysed three of the data sets ('within-site experiment', 'between-site experiment' and 'group competition experiment') from Hasegawa et al. (2014) using a Bayesian predictive framework and by including bill size as a potential predictor of competitiveness. We conducted this study under appropriate university (no. 09-1054R), state (no. SP575267) and federal (no. 23362 for banding; no. MB088806-0 for collecting) permits. Detailed information on bird capture sites and dates, housing procedures and aggression trials is described therein. In short, we captured six colourful and six drab males from each of two urban and two rural sites in the Phoenix metropolitan area (Arizona, U.S.A., see Giraudeau, Mousel, Earl, & McGraw, 2014 for a description of the sites). Birds were housed individually in an indoor, IACUC-approved room in small wire cages (0.21×0.29 m and 0.40 m high) on the Arizona State University campus. Because the colour of one bird was misclassified, our final sample sizes were 11 urban colourful, 12 urban drab, 12 rural colourful and 12 rural drab males. From the data of Hasegawa et al. (2014), we obtained four hypothetical competition groups across two experiments, in which 8, 8, 12 and 20 individuals were included (Appendix 1, Table A1). Estimates of the misclassified individual were excluded afterwards. Because we arranged the dyads to match them for similar rank of coloration (e.g. brighter versus brighter and drabber versus drabber) and body mass at capture (Hasegawa et al., 2014), the number and members of the hypothetical competition groups were automatically determined in response. This can be understood using the example in Fig. 1: there, if we exchange the positions of A and E in experiment 2, two hypothetical groups (A, B, C, D and E, F, G, H) emerge instead of a single large group.

All types of trials were conducted for 30 min each, were completed between 0600 and 0900 hours, and were run in a large wire cage (0.50×0.59 m and 0.77 m high) containing multiple perches and a single food dish of sunflower seeds. We first ran the within-site experiment (24–29 May 2011), then the between-site experiment (1–6 June 2011) and finally the group competition experiment (8–10 June 2011; see Hasegawa et al., 2014 for details). No bird was ever grouped with another male against which he had previously competed.

At capture, we measured keel length, bill length, bill height (all to the nearest 0.01 mm) and body mass (to the nearest 0.1 g) for all birds. Bill length was measured from the anterior end of the nostril to the tip of the upper mandible, and bill height was measured in a vertical plane at the anterior end of the nostrils over both mandibles (Giraudeau, Nolan, et al., 2014; note that bill width was not measured in the current data set). To investigate bill size, a principal component analysis based on the correlation matrix using the 'prcomp' function

in R software was used. Although our main focus was overall bill size (PC1), we also studied bill shape (PC2) in relation to competitiveness, as bill shape (i.e. sharpness) can influence contest outcomes (e.g. Rico-Guevara & Araya-Salas, 2015; see also Sustaita & Rubega, 2014). PC1 and PC2 explained 56% and 44 % of the variance, respectively. Rotation values of bill length and bill height were 0.70 each for PC1 and 0.70 and –0.70 for PC2, and thus larger PC1 values indicate larger bills, while larger PC2 values indicate pointier bills. Although previous studies with larger sample sizes found significant relationships between the degree of urbanization and bill morphology in house finches (Badyaev et al., 2008; Giraudeau, Nolan, et al., 2014; see also Badyaev, 2014), PCs were not significantly explained by site in our study (see Results).

Statistical Analysis

A hierarchical Bayesian inferential approach was used to estimate the competitiveness of each male based on the win/loss outcomes of dyadic interactions (Adams, 2005; Hasegawa & Kutsukake, 2015; Romero & Castellanos, 2010). The model was hierarchically constructed so that observable outcomes were modelled conditionally on a prior distribution, which itself is given a probabilistic specification in terms of further parameters (i.e. a hyper-prior distribution; Gelman et al., 2013). The hyper-prior distribution was a non-informative, flat distribution (i.e. uniform distribution, $U(0, 1000)$; see Appendix 2). To analyse the data, we applied the Markov chain Monte Carlo (MCMC) method to compute the likely posterior distributions (McCarthy, 2007), using software developed specifically for Bayesian estimation (WinBUGS 1.4; Lunn, Best, Thomas, & Spiegelhalter, 2000; Spiegelhalter, Thomas, Best, & Lunn, 2003). The reproducibility of the MCMC simulation was assessed by calculating the Brooks–Gelman–Rubin statistic (Rhat), which should be <1.2 for all parameters (Kass, Carlin, Gelman, & Neal, 1998).

We used a Bayesian extension of the Bradley–Terry model proposed by Adams (2005) for estimating the attained competitiveness values, or 'dominance index' (d in the model), which is estimated relative to all other group members (Adams, 2005; see also Shev, Hsieh, Beisner, & McCowan, 2012). The Bradley–Terry model is widely used in the context of animal contests and assumes the existence of a transitive linear hierarchy in which the probability that one individual prevails over another in a competition is a logistic function of the difference in dominance indices (i.e. competitiveness differences; Bradley & Terry, 1952; see also Boyd & Silk, 1983). Let P_{ij} be the probability that individual i dominates individual j , then the probability that individual i wins against individual j is given by logit (P_{ij}) = $d_i - d_j$. Competitiveness in each trial (attained competitiveness: d) would be determined by $d = \mu + \epsilon$, where μ = inherent competitiveness of each bird and ϵ denotes the residual competitiveness, representing additional unquantified within-individual variation. The term ϵ is not directly specified in the model and was drawn from a normal distribution with mean = 0 and variance estimated by the model (Adams, 2005; Jackson, Inger, Bearhop, & Parnell, 2009; Mutshinda, O'Hara, & Woiwod, 2011). Male identity was included as a random intercept; thus, each male was assumed to have his own μ value, all of which come from a normal distribution with mean and variance estimated by the model.

A total of 240 000 Monte Carlo iterations per chain, including 80 000 burn-in iterations, was performed, and one of every 40 steps was sampled from the remaining 160 000 steps, yielding 4000 samples per calculation trial. The calculation was repeated three times, so in total 12 000 samples were obtained for each estimate. Model estimates are presented as posterior means, 95% credible intervals (CIs) and MCMC-based P values (P_{MCMC} : `mcmc.pval` in R package 'MCMC.qpcr'). Because the posterior distribution is unimodal, average values (the posterior expectation of the parameters) could be

close to the posterior mode (the Bayesian maximum likelihood estimate; Gelman et al., 2013; Shev et al., 2012). All analyses were carried out in R 3.0.2 (R Development Core Team, 2013) using the coda package (Plummer, Best, Cowles, & Vines, 2006).

To study whether future dominant males have higher competitiveness estimates than future subordinate males in each group, we estimated the posterior probability that the difference in the competitiveness estimates between them would be higher than zero [i.e. $\Pr(\mu_{\text{dominant}} - \mu_{\text{subordinate}} > 0)$] based on the proportion of the total of 12 000 simulations for which it was true (i.e. P_{MCMC} ; Gelman et al., 2013). Male rank in the group competition experiments (i.e. top-ranking, second-highest, third-highest and lowest ranking) was estimated following Adams (2005). The assigned rank was highly correlated with rank judged from the observation based on the number of aggressive encounters won and lost in each group ($0.93 < r \leq 1.00$), indicating clear rank structure within groups independent of assignment method. Competitiveness estimates were compared only within each hypothetical competition group, to exclude potential differences in mean competitiveness across groups (see Hasegawa & Kutsukake, 2015 for similar usage) when studying competitiveness in relation to future contest outcomes. When we analysed Bayesian competitiveness estimates of each male category (urban colourful, urban drab, rural colourful and rural drab), we studied competitiveness relative to the mean competitiveness in each hypothetical competition group, to study whether the male had higher or lower competitiveness than the mean value of each group. Here, we calculated $\Pr(\text{mean}(\mu_{\text{urban-colourful}}) - \text{mean}(\mu_{\text{group}})) > 0$, and so on (i.e. we calculated it for urban drab, rural colourful and rural drab males as well).

RESULTS

Contest Outcomes and Estimation of Future Competitiveness

In the same hypothetical competition group, future dominant males tended to have higher competitiveness estimates than future subordinate males (Fig. 2a), especially when the rank difference was large (Fig. 2a). After excluding duplicate individuals while maintaining similar sample sizes for each rank difference (four pairs for one rank difference, two rank differences and three rank differences, each), we found that dominant males had higher competitiveness estimates than subordinate males (Fig. 2b), indicating that Bayesian

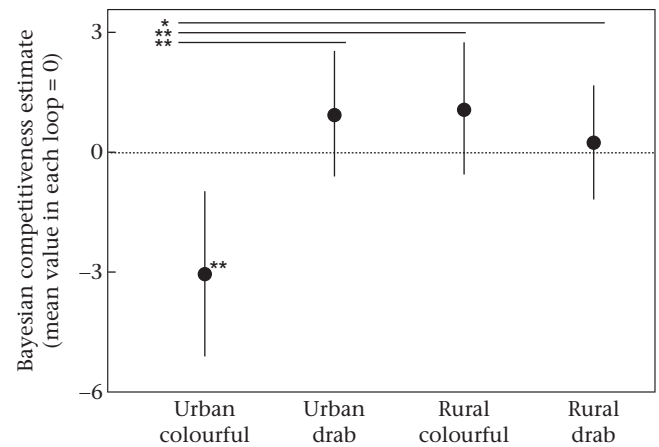


Figure 3. Bayesian competitiveness estimates for each category of males. The mean value in each hypothetical competition group was subtracted from each competitiveness estimate (and thus zero equals mean competitiveness estimate in each hypothetical group). Posterior means and 95% credible intervals are presented. * $P_{\text{MCMC}} < 0.05$; ** $P_{\text{MCMC}} < 0.01$ (see text for detailed information).

competitiveness estimates from dyadic contests, on average, explained future dominant–subordinate relationships during competitions within groups (with four birds).

Competitiveness and Other Male Attributes

After subtracting mean competitiveness in each hypothetical competition group, the mean competitiveness estimate for urban colourful males was lower than zero, indicating that they had lower competitiveness than mean competitiveness in each group (see Methods). In contrast, this measure did not differ from zero for the other three male categories (Fig. 3). Moreover, mean competitiveness estimates for urban colourful males were much lower than those for the other three male categories (Fig. 3). In addition, the competitiveness estimate decreased with increasing bill size ($r = -0.24$, $\text{CI} = -0.42, -0.05$, $P_{\text{MCMC}} = 0.01$; Fig. 4), whereas the correlation coefficient between competitiveness estimate and bill shape PC2 did not deviate from zero ($r = 0.03$, $\text{CI} = -0.16, 0.20$, $P_{\text{MCMC}} = 0.77$).

Bill size PCs were not significantly explained by site (urban versus rural), plumage colour (colourful versus drab), the interaction between them (PC1: $F < 0.50$, $P > 0.50$; PC2: $F < 2.72$, $P > 0.11$)

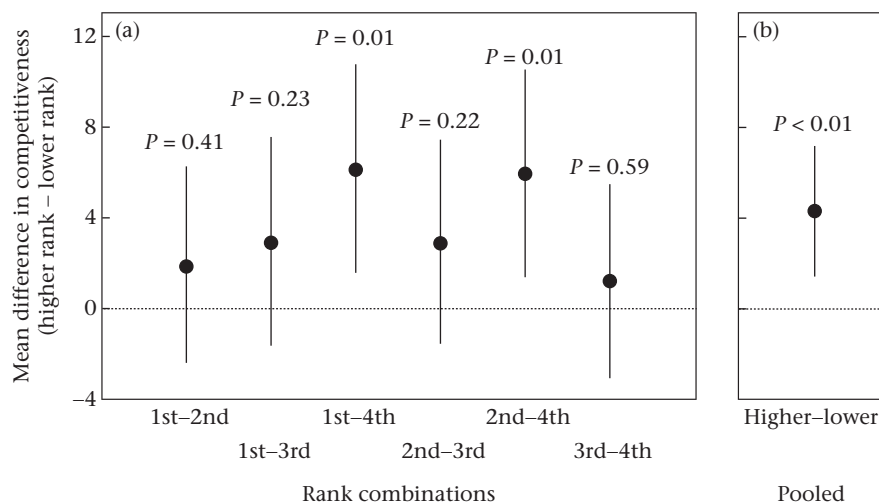


Figure 2. Mean differences in competitiveness estimates obtained from a series of two paired contests between future dominant (higher-ranking) males and future subordinate (lower-ranking) males (a) when each combination was analysed separately and (b) when all the combinations were pooled while excluding duplicate males. Posterior means, 95% credible intervals and P values based on MCMC are presented.

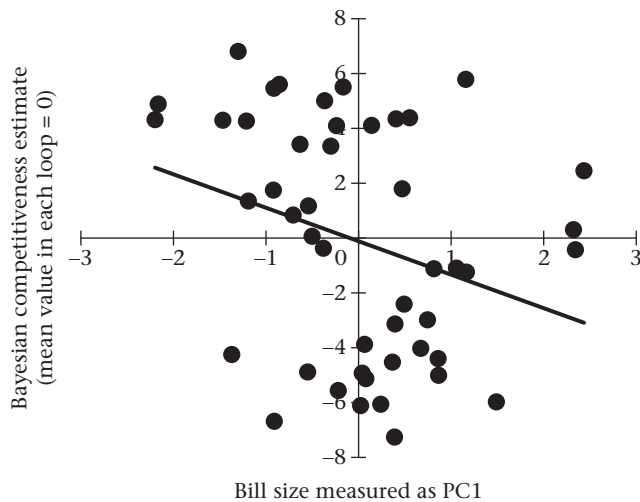


Figure 4. Bayesian competitiveness estimate in relation to male bill size (PC1). The mean value in each hypothetical competition group was subtracted from each competitiveness estimate (and thus zero equals mean competitiveness estimate in each hypothetical competition group). For illustrative purpose, only posterior mean values together with the simple regression line for these values are represented (note that actual statistics in the text are based on each of 12 000 iterations).

or keel length (an index of body size; PC1: $r = 0.17$, $P = 0.24$; PC2: $r = -0.11$, $P = 0.46$). Thus, the relationship between bill size PCs and the competitiveness estimates should not be confounded by these variables.

DISCUSSION

Bayesian competitiveness estimates that were derived from paired contests successfully predicted overall future dominant–subordinate relationships when animals were competing in larger groups of four birds. Our competitiveness estimates failed to explain relative competitiveness for some rank combinations, particularly for males of adjacent ranks (Fig. 2). However, this is a predictable outcome because of some uncertainty in the competitiveness measure, and because the competitiveness is situation dependent (e.g. condition and motivation at time of contest). Thus some rank turnover may occur when contestants had similar inherent competitiveness (see Hasegawa & Kutsukake, 2015). We conclude that Bayesian competitiveness estimates can capture relative competitiveness among birds, at least on average, when based on two paired contests.

On average, colourful urban males had much smaller competitiveness estimates (μ in our model) than the other three categories of males. This finding is consistent with the results of our previous group competition experiment (Fig. 3; see also Fig. 4 in Hasegawa et al., 2014), reinforcing the validity of Bayesian competitiveness estimation based on hypothetical competition groups. This consistency further indicates that, at least for short-term contests in our study system, agonistic outcomes of polyadic contests do not differ dramatically from those of dyads. Thus, although polyadic contests have potentially more opportunities for coalition formation and eavesdropping (reviewed in Sherratt & Mesterton-Gibbons, 2013), dyadic contests may be a main determinant of group contest outcomes. Furthermore, the outcomes of aggressive encounters won and lost were explained solely by inherent, consistent competitiveness (μ) of the individual, without considering the additional, situation-dependent component (ϵ). Thus, changing antagonism based on the plumage coloration of the opponents, which is expressed in ϵ together with other situation-dependent competitiveness, should at best be a minor contributor to contest outcomes.

In fact, no detectable relationship was observed between the ϵ value and the opponent's category (see Appendix 3, Fig. A1). This inference on the minor importance of changing antagonism is also consistent with previous experimental studies (McGraw & Hill, 2000b; McGraw et al., 2007), in which colour manipulation did not affect the win/loss outcomes of aggressive encounters in male house finches. Unlike female house finches, which decrease their competitiveness towards colourful males (perhaps due to female mate preference for colourful males; Belthoff & Gowaty, 1996), discrimination of colourful versus drab males may be trivial in males.

Bayesian analysis of animal contests yields individual competitiveness estimates (in this case, inherent competitiveness as a latent variable; see Appendix 2), unlike the number of aggressive encounters won/lost, which depends on the combination of winner and loser competitiveness in each contest (Briffa et al., 2013). As mentioned above, an advantage of using inherent competitiveness is that this measure is less dependent on changing antagonism based on the phenotype of specific opponents than the outcomes of single paired contests. Another advantage of this method is that the results are less influenced by trial conditions (e.g. body condition at each trial, random fluctuation of trial environment), which is particularly useful when studying traits that predict competitiveness (e.g. badges of status; Senar, 2006), rather than using situation-dependent, fluctuating contest outcomes based on each trial. In fact, this may explain why we found a negative relationship between competitiveness and bill size using the Bayesian framework, but not in a conventional statistical analysis of contest outcomes (Appendix 4, Table A2).

Although the negative relationship between competitiveness and bill size is contrary to our prediction, it is possible that smaller bills in agile songbirds such as house finches are easier to manoeuvre and therefore more suitable for rapid, short-range jousting, which is among the most aggressive forms of physical interactions in this species (e.g. McGraw et al., 2007; see also Herrel, Podos, Vanhooydonck, & Hendry, 2009). Measurements of manoeuvrability or experimental manipulations of bill size would shed light on this possibility. An alternative explanation is that birds with smaller bills have a disadvantage during scramble competition for food, due to lower ability to handle large, nutrition-rich seeds (e.g. Badyaev et al., 2008; Spenckle & Blem, 1984), and thus invest more in contest competition for food as a 'best-of-a-bad-job' strategy (i.e. negatively correlated handicap: Hasegawa et al., 2014; McGraw & Hill, 2000a,b; McGraw et al., 2007). This explanation may also help explain the negative link between dominance, body size and nutritional state in female house finches (Belthoff & Gauthreaux, 1991) and other species (body condition; Edler & Fiedl, 2010; Santos, Maia, & Macedo, 2009). Differential mating success may also drive a motivational asymmetry between larger- and smaller-billed males, as bill morphology affects vocal performance (e.g. fast trill rate and high-pitch song of small-billed males: Badyaev et al., 2008; Badyaev, 2014; Christensen, Kleindorfer, & Robertson, 2006; Giraudeau, Nolan, et al., 2014), which can be an intersexually selected trait (Mennill, Badyaev, Jonart, & Hill, 2006; Nolan & Hill, 2004; also see Cardoso, 2012; Gil & Brumm, 2014 for female preference for higher frequency vocalizations). These advantages of having small bill size may override the disadvantage (e.g. low bite force; see above) and contribute to the observed weak negative (rather than positive) relationship between bill size and competitiveness.

In summary, we found that Bayesian competitiveness estimation is a valuable approach for using acquired information on paired aggressiveness to predict future social competition outcomes and to examine the potential phenotypic predictors of competitiveness without testing all combinations of individuals. The relative competitiveness of males based on paired contests was comparable to the outcomes of group contests, suggesting that house finches have conserved competitive tendencies, both in one-on-one and in

larger group competitions, at least within a limited timeframe (1 week in our study). The goal is now to evaluate the predictive power of the Bayesian approach in other species, over longer timeframes, and in more complex social settings.

Acknowledgments

We thank S. Shanmugam, A. Callaghan and M. Weaver for their assistance with experiments. We are grateful to the members of the McGraw Lab from the School of Life Sciences at ASU and N. Sasaki for meaningful discussion. We also thank Professor R. Rutowski, his lab members at ASU, T. Union, A. Fields and L. Zarroli at ASU for their great help and daily support during Masaru's stay. We thank S. Neill-Eastwood, D. McAndrew and their colleagues for animal husbandry. We are also grateful to the members of the laboratory of the Graduate University for Advanced Studies, especially to Drs A. Kawamori and T. Harano, for their technical advice on statistics.

References

- Adams, E. S. (2005). Bayesian analysis of linear dominance hierarchies. *Animal Behaviour*, 69, 1191–1201.
- Amy, M., & Leboucher, G. (2007). Male canaries can visually eavesdrop on conspecific food interactions. *Animal Behaviour*, 74, 57–62.
- Badyaev, A. V. (2014). Reconciling innovation and adaptation during recurrent colonization of urban environments: molecular, genetic, and developmental bases. In D. Gil, & H. Brumm (Eds.), *Avian urban ecology: Behavioural and physiological adaptations* (pp. 155–168). Oxford, U.K.: Oxford University Press.
- Badyaev, A. V., Young, R. L., Oh, K. P., & Addison, C. (2008). Evolution on a local scale: developmental, functional, and genetic bases of divergence in bill form and associated changes in song structure between adjacent habitats. *Evolution*, 62, 1951–1964.
- Belthoff, J. R., & Gauthreaux, S. A., Jr. (1991). Aggression and dominance in house finches. *Condor*, 93, 1010–1013.
- Belthoff, J. R., & Gowaty, P. A. (1996). Male plumage coloration affects dominance and aggression in female house finches. *Bird Behavior*, 11, 1–7.
- Boyd, R., & Silk, J. B. (1983). A method for assigning cardinal dominance ranks. *Animal Behaviour*, 31, 45–58.
- Bradley, R. A., & Terry, M. E. (1952). The rank analysis of incomplete block designs. I. The method of paired comparisons. *Biometrika*, 39, 324–345.
- Briffa, M., Hardy, I. C. W., Gammell, M. P., Jennings, D. J., Clark, D. D., & Goubault, M. D. D. (2013). Analysis of animal contest data. In I. C. W. Hardy, & M. Briffa (Eds.), *Animal contests* (pp. 47–85). Cambridge, U.K.: Cambridge University Press.
- Cardoso, G. C. (2012). Paradoxical calls: the opposite signaling role of sound frequency across bird species. *Behavioral Ecology*, 23, 237–241.
- Christensen, R., Kleindorfer, S., & Robertson, J. (2006). Song is a reliable signal of bill morphology in Darwin's small tree finch *Camarhynchus parvulus*, and vocal performance predicts male pairing success. *Journal of Avian Biology*, 37, 617–624.
- Edler, A. U., & Fiedl, T. W. (2010). Plumage coloration, age, testosterone and dominance in male red bishops (*Euplectes orix*): a laboratory experiment. *Ethology*, 116, 806–820.
- Freire, R., Eastwood, M. A., & Joyce, M. (2011). Minor beak trimming in chickens leads to loss of mechanoreception and magnetoreception. *Journal of Animal Science*, 89, 1201–1206.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. (2013). *Bayesian data analysis* (3rd ed.). London, U.K.: Chapman and Hall.
- Gil, G., & Brumm, H. (2014). *Avian urban ecology: Behavioural and physiological adaptations*. Oxford, U.K.: Oxford University Press.
- Giraudeau, M., Mousel, M., Earl, S., & McGraw, K. (2014). Parasites in the city: degree of urbanization predicts poxvirus and coccidian infections in house finches (*Haemorrhous mexicanus*). *PLoS One*, 9, e86747. <http://dx.doi.org/10.1371/journal.pone.0086747>.
- Giraudeau, M., Nolan, P. M., Black, C., Earl, S., Hasegawa, M., & McGraw, K. J. (2014). Song characteristics track bill morphology along a gradient of urbanization in house finches. *Frontiers in Zoology*, 11, 83.
- Gonzalez, G., Sorci, G., Smith, L., & de Lope, F. (2002). Social control and physiological cost of cheating in status signalling male house sparrow (*Passer domesticus*). *Ethology*, 108, 289–302.
- Grava, A., Grava, T., Didier, R., Lait, L. A., Dosso, J., Koran, E., et al. (2012). Interspecific dominance relationships and hybridization between black-capped and mountain chickadees. *Behavioral Ecology*, 23, 566–572.
- Hasegawa, M., & Kutsukake, N. (2015). Bayesian competitiveness estimation predicts dominance turnover among wild male chimpanzees. *Behavioral Ecology and Sociobiology*, 69, 89–99.
- Hasegawa, M., Ligon, R. A., Giraudeau, M., Watanabe, M., & McGraw, K. J. (2014). Urban and colorful male house finches are less aggressive. *Behavioral Ecology*, 25, 641–649.
- Herrel, A., Podos, J., Vanhooydonck, B., & Hendry, A. P. (2009). Force-velocity trade-off in Darwin's finch jaw function: a biomechanical basis for ecological speciation? *Functional Ecology*, 23, 119–125.
- Hsu, Y., Earley, R. L., & Wolf, L. L. (2006). Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. *Biological Reviews*, 81, 33–74.
- Jackson, A. L., Inger, R., Bearhop, S., & Parnell, A. (2009). Erroneous behaviour of MixSIR, a recently published Bayesian isotope mixing model: a discussion of Moore & Semmens (2008). *Ecology Letters*, 12, E1–E5.
- Kass, R. E., Carlin, B. P., Gelman, A., & Neal, R. (1998). Markov chain Monte Carlo in practice: a round table discussion. *American Statistician*, 52, 93–100.
- Koivula, K., Lahti, K., Orreli, M., & Rytönen, S. (1993). Prior residency as a key determinant of social dominance in the willow tit. *Behavioral Ecology and Sociobiology*, 33, 283–287.
- Loretto, M.-C., Fraser, O. N., & Bugnyar, T. (2012). Ontogeny of social relations and coalition formation in common ravens (*Corvus corax*). *International Journal of Comparative Psychology*, 25, 180–194.
- Lunn, D., Best, N. G., Thomas, A., & Spiegelhalter, D. J. (2000). WinBUGS – a Bayesian modeling framework: concepts, structure, and extensibility. *Statistics and Computing*, 10, 325–337.
- McCarthy, M. A. (2007). *Bayesian methods for ecology*. Cambridge, U.K.: Cambridge University Press.
- McGraw, K. J., & Hill, G. E. (2000a). Plumage brightness and breeding-season dominance in the house finch: a negatively correlated handicap? *Condor*, 102, 456–461.
- McGraw, K. J., & Hill, G. E. (2000b). Carotenoid-based ornamentation and status signaling in the house finch. *Behavioral Ecology*, 11, 520–527.
- McGraw, K. J., Medina-Jerez, W., & Adams, H. (2007). Carotenoid-based plumage coloration and aggression during molt in male house finches. *Behaviour*, 144, 165–178.
- Mennill, D. J., Badyaev, A. V., Jonart, L. M., & Hill, G. E. (2006). Male house finches with elaborate songs have higher reproductive performance. *Ethology*, 112, 174–180.
- Mutshinda, C. M., O'Hara, R. B., & Woiwod, I. P. (2011). A multispecies perspective on ecological impacts of climatic forcing. *Journal of Animal Ecology*, 80, 101–107.
- Nolan, P. M., & Hill, G. E. (2004). Female mate choice for song characteristics in the house finch. *Animal Behaviour*, 67, 403–410.
- Plummer, M., Best, N. G., Cowles, K., & Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. *R News*, 6, 7–11.
- R Development Core Team. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- Rico-Guevara, A., & Araya-Salas, M. (2015). Bills as daggers? A test for sexually dimorphic weapons in a lekking hummingbird. *Behavioral Ecology*, 26, 21–29.
- Romero, T., & Castellanos, M. A. (2010). Dominance relationships among male hamadryas baboons (*Papio hamadryas*). *Journal of Ethology*, 28, 113–121.
- Santos, E. S. A., Maia, R., & Macedo, R. H. (2009). Condition-dependent resource value affects male-male competition in the blue-black grassquit. *Behavioral Ecology*, 20, 353–359.
- Senar, J. C. (2006). Color displays as intrasexual signals of aggression and dominance. In G. E. Hill, & K. J. McGraw (Eds.), *Function and evolution: Vol. II. Bird coloration* (pp. 87–136). Cambridge, MA: Harvard University Press.
- Sheratt, T. N., & Mesterton-Gibbons, M. (2013). Models of group or multi-party contests. In M. Briffa, & I. C. W. Hardy (Eds.), *Animal contests* (pp. 33–46). Cambridge, U.K.: Cambridge University Press.
- Shev, A., Hsieh, F., Beisner, B., & McCowan, B. (2012). Using Markov chain Monte Carlo (MCMC) to visualize and test the linearity assumption of the Bradley–Terry class of models. *Animal Behaviour*, 84, 1523–1531.
- Spence, J. M., & Blem, C. R. (1984). Metabolism and food selection of eastern house finches. *Wilson Bulletin*, 96, 184–195.
- Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2003). *WinBUGS user manual, version 1.4*. Cambridge, U.K.: MRC Biostatistics Unit.
- Sustaita, D., & Rubega, M. A. (2014). The anatomy of a shrike bite: bill shape and bite performance in Loggerhead Shrikes. *Biological Journal of the Linnean Society*, 112, 485–498.

Appendix 1

Table A1
Four hypothetical competition groups in the current data set

	Within-site experiment	Between-site experiment
Group 1	1 vs 2	2 vs 3
	3 vs 4	4 vs 5
	5 vs 6	6 vs 7
	7 vs 8	8 vs 1
Group 2	9 ^a vs 10	10 vs 11
	11 vs 12	12 vs 13
	13 vs 14	14 vs 15
	15 vs 16	16 vs 9 ^a
Group 3	17 vs 18	18 vs 19
	19 vs 20	20 vs 21

Table A1 (continued)

	Within-site experiment	Between-site experiment
Group 4	21 vs 22	22 vs 23
	23 vs 24	24 vs 25
	25 vs 26	26 vs 27
	27 vs 28	28 vs 17
	29 vs 30	30 vs 31
	31 vs 32	32 vs 33
	33 vs 34	34 vs 35
	35 vs 36	36 vs 37
	37 vs 38	38 vs 39
	39 vs 40	40 vs 41
	41 vs 42	42 vs 43
	43 vs 44	44 vs 45
	45 vs 46	46 vs 47
	47 vs 48	48 vs 29

Note that $4N + 1$, $4N + 2$, $4N + 3$, $4N + 4$ indicate urban colourful, urban drab, rural drab and rural colourful males, respectively (N as a non-negative integer, 0–11; the same numbers indicate the same individuals).

^a The misclassified individual (see text).

Appendix 2

Here we give the WinBUGS specification of the Bradley–Terry model modified from Adams (2005) and Hasegawa and Kutsukake (2015)

```

model {
# For each of 48 dyads, the number of aggressive interactions won by individual 1 is a binomial
random variable. Logit of  $p$  is equal to the difference in the attained dominance  $d$ .
for (i in 1:48) {
n[i] <- win1[i] + win2[i]
win1[i] ~ dbin (p[i], n[i])
logit(p[i]) <- d[ind1[i]] - d[ind2[i]] }

# For each 48 birds, inherent competitiveness  $r$  follows a normal distribution with mean 0.
for (i in 1:48) {
r[i] ~ dnorm(0, tau2) }

# Each opponent's  $d$  is derived from a normal distribution with his  $\mu (= r)$  value as the mean
for (i in 1:96) {
d[i] ~ dnorm(mu[i], tau) I(-15, 15)
mu[i] <- r[ID[i]] }

# Specifications of priors.
tau <- pow(sigma, -2)
tau2 <- pow(sigma2, -2)
sigma ~ dunif(0, 1000)
sigma2 ~ dunif(0, 1000)
}

```

Appendix 3

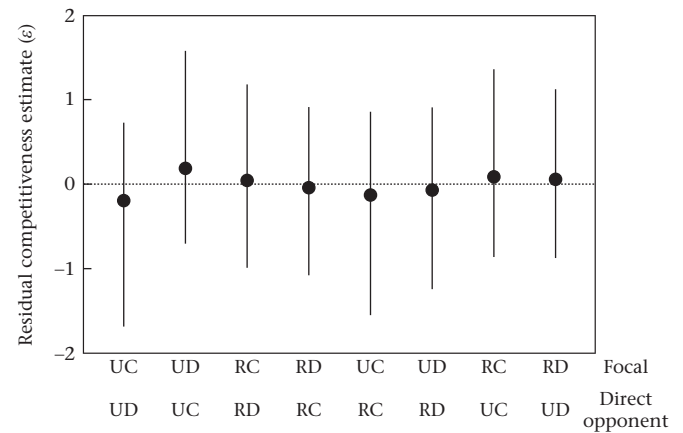


Fig. A1. Residual competitiveness estimates (ϵ) in relation to the combinations of the focal males and the direct opponents. Male categories: UC, urban colourful; UD, urban drab; RC, rural colourful; RD, rural drab. Posterior means and 95% credible intervals are presented.

Appendix 4

Table A2
Generalized linear mixed-effect model (GLMM) predicting the proportion of aggressive encounters won by each individual in relation to bill size PCs

	Slope	CI (lower, upper)	<i>P</i> _{MCMC}
Group competition experiment			
Bill size PC1	−0.64	−1.43, 0.14	0.11
Bill size PC2	0.31	−0.65, 1.31	0.53
Pooled data from within-site and between-site experiments			
Bill size PC1	−3.23	−7.88, 0.65	0.13
Bill size PC2	0.28	−4.73, 5.22	0.91

Each line indicates the results of separate GLMM analysis with binomial error distribution with logit link function using MCMCglmm in package MCMCglmm (posterior mean coefficient, 95% credible intervals, and *P* values based on MCMC are presented;

the group competition experiment: $N_{\text{total}} = 44$, $N_{\text{trial}} = 11$; pooled data from the within-site and between-site experiments: $N_{\text{total}} = 88$, $N_{\text{trial}} = 44$, $N_{\text{male}} = 44$). The response variable was assigned by the proportion of aggressive encounters won by using the cbind function in R [i.e. cbind(number of aggressive encounters won, number of aggressive encounters lost)]. Trial number is included as a random factor in each analysis, and individual ID is also included as a random factor in the analyses of pooled data. For each analysis, a total of 240 000 Monte Carlo iterations per chain (total three chains), including 80 000 burn-in iterations, were performed, and one of every 40 steps was sampled from the remaining 160 000 steps, yielding 4000 samples per single calculation trial (total 12 000), as in our Bayesian Bradley–Terry model (see Methods). The Brooks–Gelman–Rubin statistic (*R*_{hat}) was <1.2 for all parameters.