

# <sup>1</sup> Validating morphological condition indices and their relationship <sup>2</sup> with reproductive success in great-tailed grackles

Berens JM<sup>1</sup>      Folsom M<sup>2</sup>      Sevchik A<sup>1</sup>      Bergeron L<sup>3</sup>      Logan CJ<sup>2</sup>  
McCune KB<sup>3\*</sup>

2021-07-19

## **6 Affiliations:**

- 7      1) Arizona State University School of Life Sciences  
8      2) Max Planck Institute for Evolutionary Anthropology  
9      3) University of California Santa Barbara

<sup>10</sup> \*Corresponding author: KB McCune (kelseybmccune@gmail.com)

**Cite as:** Berens JM, Logan CJ, Folsom M, Sevchik A, Bergeron L, McCune KB. Submitted to PCI Ecology Nov 2020. Validating morphological condition indices and their relationship with reproductive success in great-tailed grackles.

<sup>14</sup> This preregistration has been pre-study peer reviewed and received an In Principle Acceptance  
<sup>15</sup> by:

<sup>16</sup> Marcos Mendez (2019 In Principle Acceptance) Are condition indices positively related to each other and to  
<sup>17</sup> fitness?: a test with grackles. *Peer Community in Ecology*, 100035. 10.24072/pci.ecology.100035

- <sup>18</sup> • Reviewers: Javier Seoane and Isabel López-Rull

## <sup>19</sup> ABSTRACT

Morphological and physiological variation among individuals has the potential to influence multiple life history characteristics such as dispersal, migration, reproductive success, and survival. Individuals that are in better “condition” can disperse or migrate further or more successfully, have greater reproductive success, and survive longer, particularly in years where environmental conditions are harsh. Condition is defined in various ways, but is most often measured using an individual’s energetic state. These traits are difficult to measure directly, therefore a variety of morphological proxies to quantify energetic condition are used instead, including fat score, weight, ratio of weight to tarsus length, and a scaled mass index. However, there is mixed support regarding whether these energetic condition indices relate to life history characteristics, and whether the relationship is linear. Additionally, although some investigations use multiple morphological proxies for energetic condition, rarely have there been direct comparisons among proxies to validate that they measure the same trait. In this investigation, we define condition as an energetic state and we attempt to measure it by comparing two morphological indices (fat score and the scaled mass index) to validate whether they measure the same trait and whether they correlate with measures of reproductive success in our study system, the great-tailed grackle (*Quiscalus mexicanus*). We found that the morphological proxies did not correlate with each other, indicating that they do not measure the same trait. Further, neither proxy significantly correlated with reproductive success in males, measured as whether a male held a territory containing nests

36 or not. We found that females with a high scaled mass index had a significantly lower probability that  
37 their nest would survive on any given day. However, there was no relationship between female fat score and  
38 nest survival. These results indicate that morphological measures of energetic condition should be validated  
39 before relying on their use as a condition proxy in grackles and birds in general. Future research should  
40 investigate behavioral mechanisms underlying our result that higher scaled mass index correlated with lower  
41 nest survival to better understand the importance of energetic condition for reproductive success - a necessary  
42 component for selection to act.

## 43 INTRODUCTION

44 Morphological and physiological variation among individuals has the potential to influence multiple life  
45 history characteristics such as dispersal, migration, reproductive fitness, and survival (Wilder et al., 2016).  
46 One trait that might be particularly likely to influence these life history characteristics is energetic condition.  
47 Individuals that are in better “condition” can disperse or migrate further or more successfully, have greater  
48 reproductive success, and survive longer (Heidinger et al., 2010; Liao et al., 2011; Wilder et al., 2016),  
49 particularly in years where environmental conditions are harsh (Milenkaya et al., 2015). For example, a  
50 study conducted on vipers showed that while the level of fat reserves in males was not related to their sexual  
51 activity, females with low fat reserves engaged in sexual interactions less frequently than those with higher  
52 fat reserves (Aubret et al., 2002). In contrast, mantids showed conflicting results regarding the relationship  
53 between fat reserves and reproductive success (Barry & Wilder, 2013). Female mantids were fed either a  
54 high protein, low lipid diet, or a high lipid, low protein diet. The females that received the high lipid diet  
55 had higher lipid content in most parts of their body compared to that of their high protein diet counterparts.  
56 However, they were not able to produce even half as many eggs as the females fed the high protein, low  
57 lipid diet. This led to lower male attraction, measured by the number of copulation events, thus negatively  
58 impacting further reproductive success.

59 A variety of morphological proxies have been used to quantify energetic condition [i.e., fat score, weight,  
60 ratio of mass to structural size, residuals from a linear regression of mass as a function of structural body  
61 size; Labocha et al. (2014); Jacobs et al. (2012)]. However, there is mixed support regarding whether and  
62 how these proxies relate to life history characteristics (Labocha et al., 2014; Wilder et al., 2016). A review  
63 conducted by Barnett (2015) shows that, while mass or body size measures of energetic condition are often  
64 assumed to have a positive linear relationship with fitness, this is not always the case, and the relationship  
65 should first be empirically validated before being used as a proxy (Barnett et al., 2015). In some instances, the  
66 proxy might relate to life history characteristics, but in a non-linear way. For example, theoretical simulations  
67 of small birds show that survival does not increase linearly with energy (i.e., fat) reserves (McNamara et  
68 al., 2005). If the reserves are too low, the individual is at risk of starvation. However, once the reserves get  
69 too high, the individual is at an increased risk of predation (McNamara et al., 2005). Thus, fat reserves can  
70 relate to a life history variable (survival), but in a U-shaped relationship rather than a linear one.

71 Although some studies use multiple morphological proxies for energetic condition (e.g., Warnock & Bishop,  
72 1998), rarely are these variables directly compared. Multiple proxies should correlate with each other if  
73 they measure the same trait (energetic condition). Furthermore, there is still confusion about what trait  
74 some proxies actually measure. For example, a study conducted on two species of crickets showed that  
75 three estimates of energetic condition based on fat content or on the relationship between body mass and  
76 body length (scaled mass index or ordinary least squares regression) did not correlate with each other (Kelly  
77 et al., 2014). A similar lack of a relationship was found in flying animals such as birds (A. G. Gosler  
78 et al., 1995; Jacobs et al., 2012) and bats (McGuire et al., 2018) thus indicating cross-taxon support that  
79 morphological proxy measures do not always measure the same trait. This is an example of the jingle fallacy  
80 (Block, 1995; Carter et al., 2013), where a single trait label (“energetic condition”) actually encompasses  
81 more than one distinct trait. In this case, two investigations using different proxies can be conducted on the  
82 same research question, using the same species, but may end up with different results. This is problematic  
83 because inconsistency in results among researchers can result in potentially misleading interpretations of the  
84 impact of variation in morphology in relation to life history and population variables (Stevenson & Woods  
85 Jr, 2006).

86 Here we compare two indices (fat score and the scaled mass index) of an individual's energetic state to  
87 validate whether they correlate with each other, which would indicate that they both measure energetic  
88 condition. Fat score, as described by Kaiser (1993), is a numerical estimate of the amount of fat visible  
89 under the skin (Fig. 1). The score ranges from 0 to 8 depending on the size and appearance of the fat  
90 located in the individual's abdomen and interclavicular depression, with 0 indicating no visible fat and 8  
91 indicating extensive fat covering the ventral surface such that no muscle tissue is visible. For example, a  
92 score of 1 corresponds to sparse traces of fat visible in the interclavicular depression and abdomen. This  
93 measure is frequently used in birds (Cornelius Ruhs et al., 2019; Erciyas et al., 2010; Merilä & Svensson,  
94 1997), and is a straightforward, non-invasive method for estimating energetic condition. However, previous  
95 research found that it does not always positively relate with life history variables. For example, Haas (1998)  
96 found no difference between fat scores in individuals that had successful or failed nests in American robins  
97 and brown thrashers, indicating that fat score may not explain much of the variation in nest success in  
98 some species. Further research is needed to understand the relationship between fat score measures and life  
99 history characteristics.

100 In contrast, the scaled mass index (SMI) is more difficult to calculate than the fat score, but it has become  
101 the predominant ratio method for quantifying energetic condition within and among populations (Delciellos  
102 et al., 2018; English et al., 2018; Maceda-Veiga et al., 2014). The SMI is an individual's mass scaled  
103 by skeletal body size (Peig & Green, 2009). Unlike the common alternative which uses a simple ratio of  
104 tarsus (lower leg) length to body mass, the SMI accounts for the tendency towards allometric scaling where  
105 the relationship between body mass and structural size increases by a power law (Huxley, 1932). When  
106 individuals with different structural body sizes can be standardized to the population average structural  
107 body size, then energetic condition (the amount of mass not explained by structural body size) can be more  
108 directly compared within and across populations. That is, the SMI calculates the energetic condition as  
109 the mass of an individual relative to the population by first computing the mass that the individual would  
110 have at the population average of a specific body measurement (e.g., tarsus length). Next, structural body  
111 size of the individual is standardized by scaling the individual's structural body length by the population  
112 average of that body measurement, which accounts for population differences. The SMI is calculated as:

113  $Mass_i \left[ \frac{AvgLength_p}{Length_i} \right]^{slope_p}$  where  $Mass_i$  is each individual's weight in grams,  $Length_i$  is the value of the  
114 chosen measure of structural body length for each bird,  $AvgLength_p$  is the average structural body length in  
115 the population, and  $slope_p$  is calculated from the standard major axis regression (which is used to compare  
116 variables that were both directly measured and thus have residual error) of a structural body size measure,  
117 like tarsus length on mass (Peig & Green, 2009), and is interpreted as the expected change in structural  
118 length for a one unit increase in mass. Therefore, individuals with superior energetic condition (larger weight  
119 for their structural body size) will have a higher SMI compared to individuals that with inferior energetic  
120 condition. Studies across taxa found that the SMI relates positively to reproductive success and survival. For  
121 example, mallards with a lower SMI had lower rates of survival compared to their higher SMI counterparts  
122 (Champagnon et al., 2012), while in crimson finches SMI was positively related to the number of young that  
123 survived to independence (Milenkaya et al., 2015).

124 Our research will determine whether these two indices of energetic condition measure the same trait, and  
125 whether this trait relates to an important life history characteristic: reproductive success. Measuring repro-  
126 ductive success in birds involves finding and monitoring nests (Mayfield, 1961). However, nests are usually  
127 built in cryptic locations and parents behave secretly (Gill, 1995), thus making it difficult to quantify the  
128 number of eggs and nestlings inside the nest over time. Additionally, it is difficult and time-consuming to  
129 track the survival of offspring once they leave the nest. Therefore, we will use the predominant method in  
130 this field for quantifying reproductive success: whether a nest fledged offspring (Mayfield, 1961).

131 Our study system is a population of great-tailed grackles (*Quiscalus mexicanus*), hereafter "grackles," in  
132 Tempe, Arizona. This system is ideal for this investigation because grackles are native to the tropical  
133 climates of Central America (Johnson & Peer, 2001), but have rapidly expanded their geographic range into  
134 new areas (Wehtje, 2003). Because grackles are a water-associated species, the desert habitat of Tempe  
135 presents physiological challenges that could lead to an increased likelihood of a tradeoff between survival  
136 and reproductive attempts (Henderson et al., 2017). Deserts are characterized by a scarcity of water and  
137 extreme temperature fluctuations, which require behavioral and physiological adaptations (Costa, 2012).

138 Wide variation in energetic condition and reproductive success is possible if grackle physiology requires more  
139 water than is present in the environment, and some individuals may cope with physiological stress, or find  
140 hidden sources of water, better than others (Henderson et al., 2017).

141 **Hypotheses** We measured two morphological proxy variables of energetic condition and observed repro-  
142 ductive success in grackles to test two hypotheses. The first examined the relationship between two morpho-  
143 logical proxies of energetic condition to validate that they measure the same inherent trait. Secondly, we  
144 hypothesized that energetic condition, as measured by either or both of the morphological proxy variables,  
145 would relate to reproductive success in male and female grackles.

## 146 METHODS

147 The methods below are based on the preregistration, with small changes summarized in the Deviations from  
148 the planned methods section and further explained in the preregistration (in italics).

149 **Preregistration details** The preregistration used secondary data that were collected as part of other  
150 ongoing investigations (tarsus length in [http://corinalogan.com/Preregistrations/g\\_flexgenes.html](http://corinalogan.com/Preregistrations/g_flexgenes.html); tarsus  
151 length, body weight, number of fledglings, and whether a male holds a territory in [http://corinalogan.com/Preregistrations/g\\_withinpop.html](http://corinalogan.com/Preregistrations/g_withinpop.html); and tarsus length in [http://corinalogan.com/Preregistrations/g\\_expansion.html](http://corinalogan.com/Preregistrations/g_expansion.html)). The preregistration, containing the hypotheses, methods, and analysis plan, was written  
152 (July 2019) and submitted to Peer Community In Ecology for pre-study peer review (August 2019) before any  
153 analyses were conducted. We revised according to reviewer comments and received in principle acceptance  
154 by PCI Ecology of the version on 8 Nov 2019. After that, we conducted the analyses in the preregistration.  
155 Our final methods, results, and discussion, including all data and code, are listed below.



### 158 Summary of methods

159 Figure 1: A male grackle showing the yellow/orange tint of fat under the skin in the intraclavicular depression  
160 (left); and a female grackle showing *no* fat under the skin of the intraclavicular region, but significant fat  
161 deposits under the skin of the abdomen (right).

162 Great-tailed grackles are caught year-round in the wild in Tempe, Arizona using a variety of methods  
163 (e.g., walk-in trap, bownet, mist net). After capture we immediately processed birds by attaching colored  
164 leg bands in unique combinations for individual identification, conducted morphological measurements of  
165 weight, tarsus length, flattened wing length, tail length, skull length, bill length and fat score (Fig. 1, the  
166 amount of visible fat under the skin in the clavicle and abdomen as in Kaiser, 1993). Most grackles were  
167 released after completion of color band marking, measurements, and acquiring a blood sample. A subset  
168 of grackles were held in aviaries for up to 6 months for behavioral testing, and then released back to the  
169 wild at their location of capture. We tested the repeatability of our structural size measures on this subset

170 of individuals by measuring them again before release. The second measures were collected by the same  
171 experimenter in 11 out of 17 females and 10 out of the 18 males that were repeatedly sampled.

172 From March - August, we monitor the behavior of all color-marked grackles to determine their nesting status.  
173 We follow females carrying nesting materials to find their nest. We determine whether the male territory  
174 owner is color-marked as well. Then we check each nest approximately every day to determine the status  
175 based on the female's behavior (building, incubation, feeding nestlings, feeding fledglings, failed).

176 Individuals included in our sample were those for which we have measures of energetic condition when they  
177 were adults. We did not include individuals whose data were collected as juveniles. We also excluded data  
178 that was collected from the grackles when they were released from the aviaries to avoid any confounds due  
179 to their time in the aviary (e.g., perhaps unlimited nutritious food in the aviaries affected their fat score).  
180 However, to validate that our measures of structural body size (tarsus length or wing length) are precise  
181 and accurate, we measured twice the subset of grackles brought into aviaries - once when they were initially  
182 caught, and again up to 6 months later when we released them. We calculated the repeatability of these  
183 multiple measures. All other data included in this study came from wild-caught grackles (including the data  
184 from the birds that were brought into the aviaries on their first capture).

185 We first used logistic mixed-effect models to determine whether SMI and fat score are correlated. We also  
186 tested whether SMI and fat score varied by season because grackles are difficult to catch such that we were  
187 unable to structure our data collection to coincide with the breeding season and instead caught and measured  
188 grackles as often as possible. Previous research found a non-linear relationship between reproductive success  
189 and energetic condition variables (Milenkaya et al., 2015). To check whether this is occurring in our data,  
190 we visually examined our raw data to determine if we need to include a non-linear energetic condition  
191 independent variable into our models (i.e. FatScore<sup>2</sup>). Then we used two types of logistic mixed-  
192 effect models to determine the relationship between energetic condition and reproductive success. Both types  
193 are supported in the literature, but are slightly different in the way in which the link function is specified.  
194 First, we modeled the effect of energetic condition on reproductive success using a generalized linear mixed  
195 model framework with a logit link function (i.e. Milenkaya et al., 2015). We then also used a logistic exposure  
196 model that has a link function which accounts for the time interval between nest checks when estimating  
197 the probability of daily nest survival (Bolker, 2014; Shaffer, 2004).

## 198 After pre-study peer review: Deviations from the planned methods

199 1) We realized that the sexual dimorphism of male and female body sizes necessitates separate analyses.  
200 Therefore, we calculated SMI for males and females separately, and ran separate models for each sex  
201 for the repeatability analysis (P1 and P2).

202 2) Fat score data were distributed such that the majority of scores were 0, with some 1's and very few  
203 higher numbers. Specifically, of the 21 males, 15 had fat scores at 0, 5 scored 1, and a single male had  
204 a fat score of 2. Out of 47 females, 26 scored 0, 18 scored 1, 2 scored 2, and a single female scored 3.  
205 This lack of variance in the response variable led to problems when we ran the models: it was difficult  
206 to fit models using an ordinal regression. The function "simulateResiduals," which we used to check  
207 our data, does not work with data in the ordinal family. Consequently, we modified the model to use  
208 a logistic regression where the dependent variable FatScore is categorized as individuals that showed  
209 no visible fat ( $y = 0$ ), or some fat was present ( $y = 1$ ) where we combined all individuals that had fat  
210 score values of 1 or greater. Subsequent data checking indicated that these data were not zero-inflated  
211 or overdispersed.

## 212 Deviations when testing hypothesis 1: correlation between SMI and Fat score

213 3) Warning messages occurred during the repeatability analysis using the "rptR" package in R (Stoffel et  
214 al., 2017) indicating that the fit was singular, likely because the variance for the Experimenter random  
215 effect in the model for both female and male wing length was 0.001. We thus conducted an unregistered

216 analysis where we confirmed that our repeatability values from the repeatability models were valid,  
217 despite the warning, by hand calculating repeatability following Nakagawa & Schielzeth (2010). The  
218 hand-calculated repeatabilities were nearly identical (female R = 0.5, male R = 0.71) to the output  
219 from the rpt function.

- 220 4) Despite the data checking which indicated our model was not overdispersed or zero inflated, we could  
221 not get the fixed effects or random effect to converge using the Bayesian package in R “MCMCglmm.”  
222 We found no improvement in model fit by tweaking the priors or iterations/burnin/thin options. There-  
223 fore, we fit these models using the function glmer, a frequentist framework.  
224 5) The Season variable only includes 2 males in the breeding season category, thus we do not have a  
225 large enough sample to produce reliable estimates. We removed the Season variable from the model  
226 for males.

227 **Deviations when testing hypothesis 2: energetic condition and reproductive success**

- 228 6) Only two females had reproductive success data from more than one year in our study (2019 and 2020).  
229 Consequently, there were very few repeated measures in this sample and our random effect of bird ID  
230 accounted for zero variance. This led to a warning that our model fit was singular. Therefore, we  
231 removed the data for these females for 2020 so we could remove ID as a random effect from the model,  
232 which resulted in the model running without warnings. We removed the 2020 data for these females  
233 because their energetic condition data was collected in 2019 and these measures were more likely to  
234 relate to their 2019 reproductive success data than to their reproductive success in 2020.  
235 7) The fit of the model analyzing the relationship between energetic condition and male reproductive  
236 success (ability to hold a territory containing female nests) was singular. The Year random effect  
237 accounted for zero variance in the data, so we removed it. The fit was still singular, but we retained  
238 the ID random effect (although it also explained zero variance) to account for repeated measures in  
239 this sample.  
240 8) The model fit was again singular in our logistic exposure model because the Year random effect ex-  
241 plained zero variance in the data. We removed this random effect from the analysis.

242 **RESULTS**

243 **Prediction 1: correlation between SMI and Fat Score**

244 We were able to calculate SMI for 24 males and 62 females, and fat score values were available for 22 males  
245 and 47 females (Table 1).

246 We found that wing length was more tightly correlated with body mass than tarsus length in both sexes,  
247 therefore we used wing length in our SMI calculations (female n = 62, r = 0.26, p = 0.03; male n = 24, r =  
248 0.35, p = 0.08). This allows us to account for as much variation in body mass as possible that is associated  
249 with skeletal body size because leftover variation in body mass is more likely to relate to energetic condition.  
250 Consequently, we used wing length in our calculation of SMI.

251 To validate that we were measuring structural body size consistently across experimenters, we analyzed the  
252 repeatability of wing length in the birds in our sample that were measured more than once. We found that  
253 average wing length was repeatable (n = 17 females, Repeatability  $\pm$  standard error =  $0.53 \pm 0.18$ ; n = 18  
254 males, Repeatability  $\pm$  SE =  $0.75 \pm 0.11$ ). Data permutations and a likelihood ratio test both confirmed  
255 that these repeatability values were statistically significant at  $p < 0.01$ .

256 In females, we found that for every one unit increase in SMI, the bird is 1.3 times more likely to have some  
257 fat (a 30% increase in the odds of having fat), which is not a statistically significant relationship (female  $p$   
258 = 0.81; Table 2). In males, a one unit increase in SMI corresponds to an odds ratio of 1.6, or a 60% increase

Table 1: Table 1. Sample sizes for P1 and P2. The \*Breeding\* and \*Non-breeding season\* categories refer to the number of individuals measured in each season. The \*Reprod. success\* category represents the total number of individuals in each year observed engaging in breeding behaviors. Note that the 2019 and 2020 reproductive success sample sizes include some of the same individuals that were observed in both years. Whereas, the \*Prop. successful\* category represents the proportion of the total individuals observed engaging in breeding behaviors in each year that held a territory containing nests (males) or fledged young (females).

Category	Males	Females
Breeding Season Fat	2	12
Non-breeding fat	20	35
Breeding season SMI	6	24
Non-breeding SMI	18	38
Aviaries	16	9
Repro. success 2019	8	9
Repro. success 2020	17	13
Prop. successful 2019	0.63	0.22
Prop. successful 2020	0.47	0.54

Table 2: Table 2. Results from the logistic mixed-effect regression for 47 females and fixed-effect regression for 21 males to determine whether fat score and scaled mass index (SMI) are correlated. Estimates are presented with the standard error in parentheses. Our sample size was too small to test for a season effect in males.

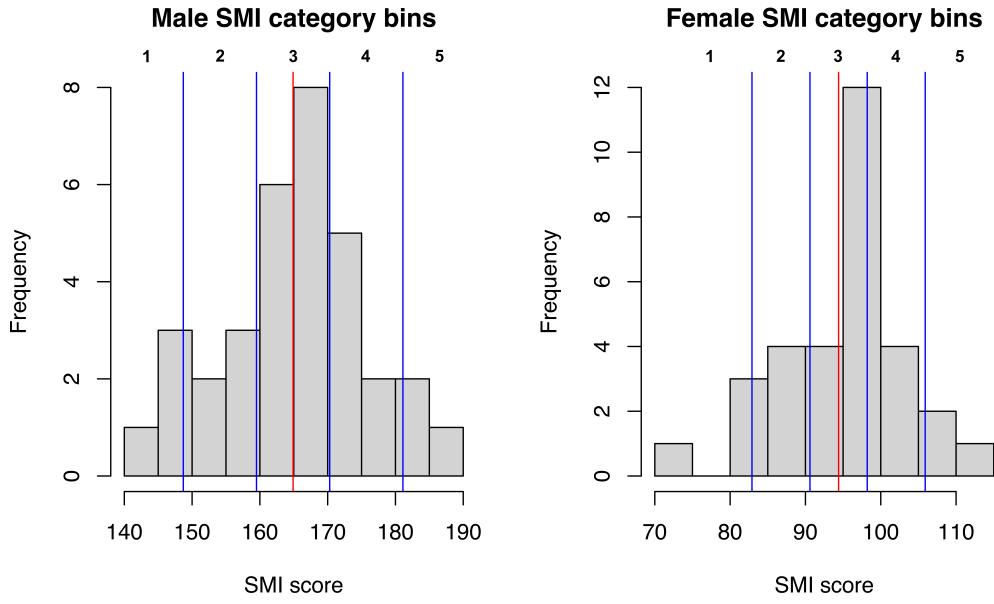
Parameter	Females		Males	
	Estimate (SE)	p-value	Estimate (SE)	p-value
Intercept	-0.20 (0.74)	0.79	-0.82 (0.64)	0.21
SMI	0.07 (0.30)	0.81	0.46 (0.62)	0.46
Season	0.27 (0.71)	0.70	NA	NA

in the odds of having some fat, which is also not a statistically significant relationship ( $p = 0.50$ ; Table 2). Together, this indicates that SMI and fat score are not equally measuring energetic condition. There was also no relationship between season (breeding or non-breeding) and female fat score ( $p = 0.71$ ). Only 2 males were measured during the breeding season, therefore we omitted season as an independent variable in the male model (Table 1).

#### Prediction 2: energetic condition and reproductive success

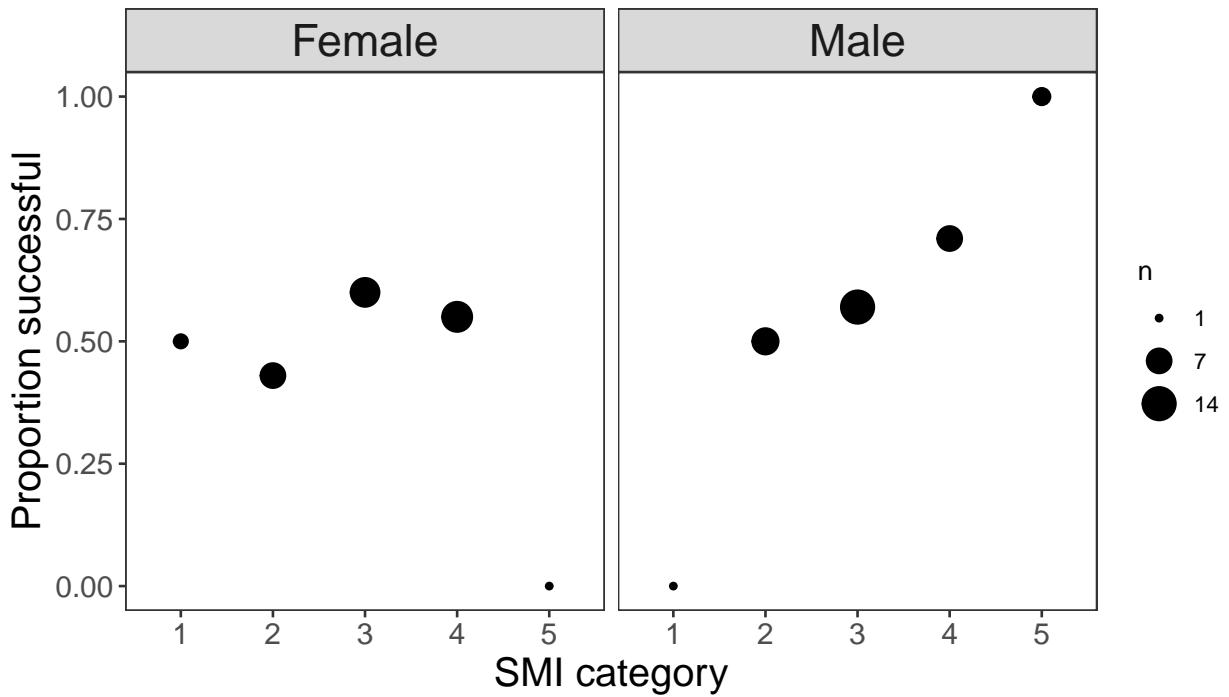
Our sample size for P2, where individuals had measures of reproductive success, SMI, and fat scores, was 20 for females and 20 for males.

To determine whether we should include any non-linear effects of SMI in our models (A. G. Gosler et al., 1995; Milenkaya et al., 2015), we visually evaluated whether individuals in any of 5 categories, ranging from low to high SMI, were more likely to be reproductively successful (Fig. 2). We found no visual evidence for a non-linear relationship between reproductive success and SMI for males or females (Fig. 3). Consequently, we did not include non-linear terms in subsequent models.



272

273 Figure 2: Frequency histogram of the SMI scores, illustrating the SMI categories, for the 33 males and 31  
 274 females for which we also had reproductive success data. The mean SMI value is indicated by a red vertical  
 275 line. We created SMI category bins (indicated with vertical blue lines) in 1 standard deviation increments,  
 276 centered on the mean. Category 3 indicates the SMI value is close to the population mean value. Categories  
 277 1 and 2 are individuals that have SMI scores that are low, and moderately low, respectively, compared to  
 278 the population mean value. Similarly, categories 4 and 5 contain individuals that have SMI scores that are  
 279 moderately high and high, respectively, compared to the population mean value.

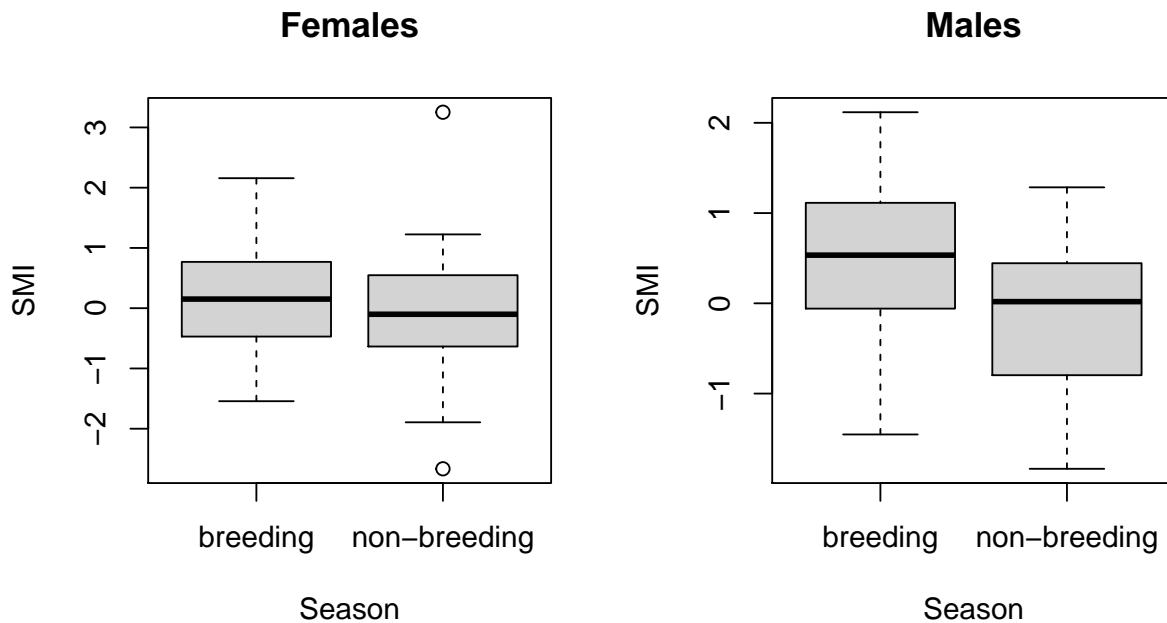


280

281 Figure 3: The proportion of individuals that successfully fledged nests (females: left) or held a territory  
 282 (males: right) in low (1), moderately low (2), moderate (3), moderately high (4) and high (5) scaled mass

283 index (SMI) categories. Dots are sized according to the number (n) of individuals in that category. There is  
284 no evidence of a non-linear relationship.

285 We used linear models to determine whether season would be important to include in our models testing  
286 whether body condition relates to reproductive success. We found that SMI did not differ by season for  
287 females (Estimate (SE):  $\beta = -0.30$  (0.26),  $p = 0.26$ ) or males ( $\beta = -0.65$  (0.43),  $p = 0.15$ ). Similarly, fat  
288 score for females ( $\beta = 0.28$  (0.68),  $p = 0.68$ ) and males ( $\beta = 17.08$  (2797.4),  $p = 0.99$ ) did not differ by  
289 season (Fig. 4). Although we note that, as stated above and indicated in the standard error value, we lack  
290 sufficient fat score data from males in the breeding season so these results should be interpreted with caution.  
291 Consequently, we did not include season as an independent variable in our subsequent models testing the  
292 relationship between our body condition proxies and reproductive success.



293  
294 Figure 4: Scaled mass index (SMI) was not significantly different between the breeding and non-breeding  
295 seasons for either sex.

296 Because fat score and SMI did not correlate, we included both as independent variables in our models testing  
297 prediction 2. For both males and females, we found no statistically significant relationships between either  
298 proxy of energetic condition and reproductive success (Table 3). Of note, the inconsistent direction of the  
299 effects for the parameter estimates further supports that SMI and fat score do not measure the same trait.

300 For females, our SMI parameter estimate of -0.92 (exponentiated to get the log odds = 0.40) indicates that  
301 a one unit increase in SMI corresponded to a 60% decrease in the odds a female would fledge an offspring  
302 ( $p = 0.13$ ). Whereas an increase from no visible fat to showing some fat corresponded to a 16% increase  
303 in the odds a female would fledge an offspring (log odds = 1.16,  $p = 0.82$ ). There was also no evidence of  
304 a significant relationship between the ability of a female to produce fledglings and having previously spent  
305 time in the aviaries (log odds = 0.25,  $p = 0.22$ ), where the odds that a female would fledge an offspring were  
306 75% lower if females spent time in the aviaries.

307 For males, there was also no statistically significant support for a relationship between whether a male  
308 defended a territory and SMI (log odds = 3.25,  $p = 0.13$ ). Nevertheless, this relationship may be biologically  
309 important because a one unit increase in SMI corresponded to a more than 300% increase in the odds a  
310 male will hold a territory containing nests. Fat score was also statistically unrelated to male reproductive  
311 success where an increase from showing no visible fat to showing some fat corresponded to a 28% decrease  
312 in territory holding (log odds = 0.72,  $p = 0.76$ ). Lastly, we found that those males who spent time in the

Table 3: Table 3. Results from the logistic regression for 20 females and 20 males to test whether reproductive success relates to condition. Estimates are presented with the standard error in parentheses.

Parameter	Females		Males	
	Estimate (SE)	p-value	Estimate (SE)	p-value
Intercept	-0.02 (0.73)	0.98	3.05 (1.40)	0.03
FatScore	0.15 (1.02)	0.89	-0.33 (1.10)	0.77
SMI	-0.92 (0.61)	0.13	1.18 (0.78)	0.13
Aviary	-1.38 (1.14)	0.23	-3.62 (1.56)	0.02*

313 aviaries were statistically less likely (97% decrease in the odds) to hold a territory compared with males who  
 314 were never in the aviaries ( $\log \text{odds} = 0.03$ ,  $p = 0.02$ ). However, we stress that our sample size was relatively  
 315 small (20 males), and we did not have a balanced sample because there were no males that did not defend a  
 316 territory and were never in the aviaries. Additionally, only five males had data from more than one breeding  
 317 season, which resulted in our model fit being singular because the random effect for bird ID accounted for  
 318 essentially zero variance. However, we kept ID in the model to account for the repeated samples.

319 **Prediction 2: energetic condition and probability of daily nest survival**

320 Logistic regression analyses to determine reproductive success from nests discovered in different stages will  
321 be systematically biased (Shaffer, 2004). Nests discovered at a more progressed stage (i.e., nestling stage  
322 compared to building stage) are statistically more likely to succeed and nests with frequent and prolonged  
323 adult visits (such as those that occur when nests survive longer) are more likely to be discovered. Therefore,  
324 nests that fail early are less likely to be detected (Shaffer, 2004). Consequently, we analyzed female repro-  
325 ductive success using a logistic exposure model (Bolker, 2014), which uses survival analysis to determine the  
326 factors affecting the probability of daily nest survival, while accounting for incomplete nest observations.

327 We found that the probability of daily nest survival was significantly negatively related to SMI (log odds  
328 = 0.50,  $p = 0.03$ ; Table 4), where, for every unit increase in SMI, the odds of daily nest survival decreased  
329 by half. This indicates that a female with a larger SMI (more mass for her structural body size) was less  
330 likely to have her nest survive each day (Fig. 5). There was no statistically significant relationship between  
331 the probability of daily nest survival and fat score (log odds = 2.48,  $p = 0.06$ ), day of the year (log odds =  
332 0.81,  $p = 0.16$ ), or time spent in the aviaries (log odds = 0.63,  $p = 0.44$ , Table 4). Although not statistically  
333 significant, the effect size for the relationship between fat score and daily nest survival is large (Fig. 5) and  
334 potentially biologically meaningful. The odds of a nest surviving on a given day are almost 2.5 times greater  
335 (248%) for birds with some fat (a score of 1) compared to no fat (a score of 0).

336 \begin{table}

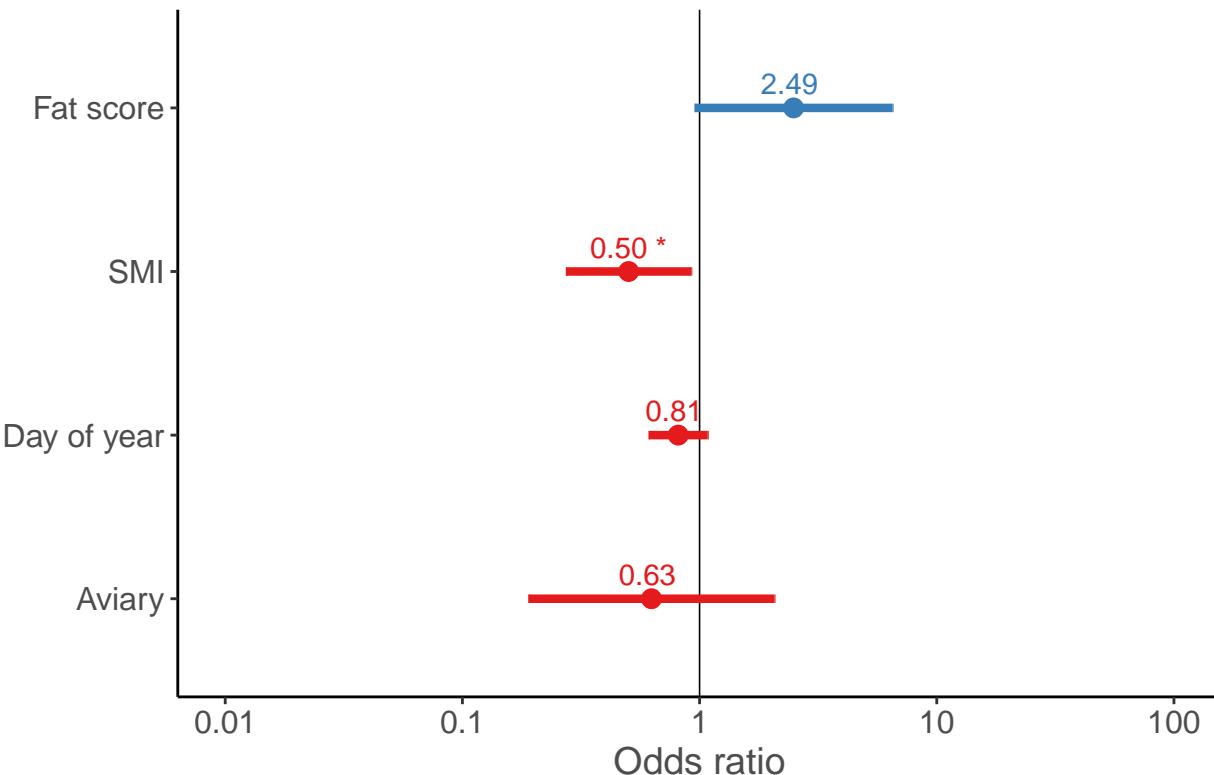
337 \caption{Table 4. Results of the logistic exposure model showing the relationship between the probability  
338 of daily nest survival and scaled mass index (SMI), fat score, the amount of time spent in the aviaries, and  
339 the day of the year. Parameter estimates are presented with the standard error in parentheses. Odds ratios  
340 (OR) represent the exponentiated estimates, and are presented to increase interpretability with 95%  
341 confidence intervals in parentheses.}

Parameter	Estimate (SE)	OR (CI)	p-value
Intercept	1.99 (0.40)	7.32 (3.3-16.0)	<0.001
Fat score	0.91 (0.49)	0.50 (0.27-0.92)	0.06
SMI	-0.69 (0.31)	2.48 (0.95-6.49)	0.03*
Day of year	-0.21 (0.15)	0.63 (0.19-2.10)	0.16
Aviary	-0.47 (0.61)	0.81 (0.60-1.10)	0.44

342 \end{table}

343

## Probability of daily nest survival



344  
345 Figure 5: Odds ratios for independent variables affecting the probability of a nest surviving a given day.  
346 The dots and corresponding values represent the odds ratio values, and lines represent the confidence  
347 intervals around the odds ratio value. The vertical line at  $x = 1$  delineates the odds ratio value for no  
348 relationship between the estimates and the probability of daily nest survival. The asterisk indicates an  
349 odds ratio value that is statistically significant.

## DISCUSSION

350 Energetic condition is not directly observable, but variation can affect life history characteristics (Barnett  
351 et al., 2015; Labocha et al., 2014). Consequently, a large corpus of research attempts to measure energetic  
352 condition using various proxy measures (Labocha et al., 2014) and largely assumes that the chosen proxy  
353 accurately reflects energetic condition as a singular trait. Although it is often implicitly assumed that all  
354 proxy measures for energetic condition reflect the same inherent trait, it is rare for one study to compare  
355 multiple proxies. However, if all proxy measures are affected similarly by a singular energetic condition  
356 phenotype, then multiple proxy measures should produce correlated results. The aim of the current study  
357 was therefore to test the idea that multiple commonly used morphological proxies equally measure  
358 energetic condition (by correlating with each other), and that these measures can explain variation in  
359 reproductive success.

360 Here we found that two morphological proxies of energetic condition, fat score and SMI, did not correlate  
361 with each other in the great-tailed grackle, regardless of whether it was the breeding or non-breeding  
362 season. While both proxies are well supported in previous research as measures of energetic condition, our  
363 results indicate that they may not be measuring the same trait. This has also been found in studies on  
364 bats (McGuire et al., 2018), which are species that similarly experience distinct demands on body structure  
365 to facilitate flight. There are several potential reasons why grackle fat score and SMI did not correlate.

367 First, it is possible that we were unable to accurately measure the amount of fat the birds actually stored.  
368 In addition to storing fat under their skin, birds may also store fat intraperitoneally (Musacchia, 1953),  
369 which would not have been detected with our fat score measure. Second, SMI and fat score may measure  
370 different components of energetic condition because variation in mass among grackles could be attributable  
371 to muscle or body water content, whereas fat score only accounts for subcutaneous fat (Labocha & Hayes,  
372 2012). Research shows that stored fat is the primary source of energy in many taxa (Walsberg, 1988),  
373 especially in birds (Blem, 1990; Pond, 1981) because the energy per ounce from fat is much higher than  
374 from proteins or carbohydrates (Gessaman, 1999). However, because desert birds, such as the grackles in  
375 our investigation, have inconsistent access to water sources, variation in body water content may obscure  
376 variation in lipid content. Measuring muscle content often requires destructive methods [i.e. sacrificing the  
377 birds; Zhang et al. (2015)] or less objective assessments such as keel prominence or breast muscle shape  
378 (Abolins-Abols & Ketterson, 2017; A. Gosler, 1991), which was beyond the scope of the current research  
379 program. Third, it is possible that fat score and SMI did not correlate due to temporal variation at a fine  
380 scale that we were unable to capture. Although we found no evidence that SMI or fat score varied by  
381 season, there is evidence from other studies that avian mass changes with time of day (Nip et al., 2019)  
382 and stage of breeding (Milenkaya et al., 2013). It was logistically impossible in our project (and in many  
383 avian research programs) to capture birds multiple times within a season or at several times per day,  
384 therefore temporal variation in data collection could obscure the correlation between these two proxies, if  
385 such a correlation exists. However, the stage of breeding is unlikely to introduce additional variance to our  
386 study because we did not catch any females that were actively engaged in any stage of the breeding  
387 process. Finally, our sample sizes might have been too small to detect an effect, but the effect size for the  
388 relationship between fat score and SMI was essentially zero (0.001), therefore it is unlikely that a larger  
389 sample size would find a biologically informative relationship between these two proxies.

390 Energetic condition can have a large impact on reproductive success in birds (Drent & Daan, 1980;  
391 Montreuil-Spencer, 2017) and in flying mammals (Welbergen, 2011). For example, female chickadees with  
392 higher winter fat scores are more likely to lay eggs earlier in the subsequent breeding season, as well as go  
393 on to feed those offspring more frequently (Montreuil-Spencer, 2017). Energetic condition is likely a factor  
394 in reproductive success in our system because previous research in great-tailed grackles found that larger  
395 and heavier males were more likely to hold territories, have more social mates, and sire more offspring  
396 (Johnson et al., 2000). Our study additionally considered female morphology and reproductive success,  
397 subcutaneous fat, and controlled for the impact of structural body size on mass. However, we found  
398 reproductive success, measured as the ability to produce fledglings (females) or to hold a territory  
399 containing nests (males), did not significantly correlate with fat score or SMI. Although our results were  
400 not statistically significant, in some cases the parameter estimates revealed log-odds that may be large  
401 enough to be biologically significant. Notably, a one unit increase in SMI corresponded to a more than  
402 300% increase in the odds a male will hold a territory containing nests, but a 60% decrease in the odds a  
403 female would fledge an offspring.

404 We additionally used logistic exposure models to determine whether the energetic condition of females  
405 related to the probability of daily nest survival. We only included females in this analysis because males  
406 were never observed contributing to nest building, incubation, or feeding nestlings in our population and so  
407 will not have a direct effect on daily nest survival. We found a negative relationship between female SMI  
408 and the likelihood of daily nest survival. This could be due to larger females actually carrying  
409 proportionally smaller energetic reserves than their smaller female counterparts (Jacobs et al., 2012), as  
410 seen in red-winged blackbirds (Langston et al., 1990). In some species, females with smaller body sizes are  
411 able to initiate breeding earlier because they can allocate more resources to reproduction compared to  
412 larger individuals that have higher bodily energy demands and therefore fewer excess energetic resources  
413 (Barbraud et al., 2000; Langston et al., 1990; Murphy, 1986). This indirectly affects reproductive success  
414 because nesting earlier increases the probability of nesting success and multiple nesting attempts (Johnson  
415 & Peer, 2001; Perrins, 1970). Yet, in our study we found no relationship between the probability of daily  
416 nest survival and day of the year, therefore this is unlikely to explain the negative relationship between  
417 SMI and nest survival. Alternatively, it is possible that larger females are unable to build a more concealed  
418 nest in the most dense vegetation, or that larger females are unable to build nests in delicate vegetation  
419 structure that is more likely to be inaccessible to predators. Moreover, the parameter estimate for the

420 relationship between fat score and the daily probability of nest survival indicates that females with some  
421 visible fat are more than twice as likely to have a nest survive a given day. Because the direction of this  
422 effect is opposite to the relationship between SMI and nest survival, this is further evidence that these two  
423 proxies represent different traits.

424 Great-tailed grackles are an interesting system to study energetic condition and reproductive success  
425 because they recently expanded their range into Arizona, where the climate and habitat are distinct from  
426 that in Central America where the species originally evolved (Wehtje, 2003). The increase in temperature  
427 variation and decrease in available water at our desert study site are both environmental stressors that  
428 have previously been found to negatively affect energetic condition (Pendlebury et al., 2004). Although our  
429 study spanned only two years, our data are likely representative of reproductive success in this environment  
430 because the temperatures during our study were in line with those from the previous three years (Center,  
431 2020). Reproductive success is vital to species persistence and abundance in novel environments (Maspons  
432 et al., 2019). Therefore, an understanding of energetic condition and its relationship with reproductive  
433 success in grackles outside of their original range could broadly inform conservation research in invasive  
434 and non-native species. While reproductive success of certain avian species may be easier to monitor at a  
435 more fine scale (i.e. cavity nesters), the predominant measure of reproductive success currently used by  
436 avian ecologists is the ability of adults to fledge offspring (since foundational work by Mayfield, 1961)  
437 because it is financially and logically accessible to more researchers. Therefore, we believe our measure of  
438 reproductive success in grackles is informative, and that research that spans taxa with diverse reproductive  
439 strategies is important for understanding general trends in energetic condition and the appropriate proxies.

440 The results of this study highlight the need to better understand proxy measures of energetic condition,  
441 not only in grackles, but for birds in general. Most studies on avian energetic condition only use one proxy  
442 variable, but because energetic condition is difficult to measure directly, it is important to compare  
443 multiple proxy variables to determine whether the proxy is measuring the intended trait (Block, 1995;  
444 Carter et al., 2013). If financially and logically feasible, future research could measure total body  
445 composition and relative mass of fat using the relatively new and promising method of quantitative  
446 magnetic resonance (Guglielmo et al., 2011), or researchers could incorporate additional physiological  
447 methods to measure energetic condition, for example, blood hematocrit levels (Dawson & Bortolotti, 1997;  
448 but see Fair et al., 2007). Additionally, studying traits that could relate to variation in energy stores, such  
449 as dispersal (Ellers et al., 1998), migratory endurance (Deppe et al., 2015), or survival (Liao et al., 2011)  
450 would allow us to disentangle whether morphological proxies like fat score and SMI are poor proxy  
451 measures for energetic condition, or whether fat score and SMI do not affect reproductive success but may  
452 be associated with other life history characteristics. Because SMI can perform poorly in birds with low  
453 lipid mass, future research should also compare several mass by structural body size equations to determine  
454 the most appropriate proxy for a specific study system (Jacobs et al., 2012). Lastly, future research would  
455 benefit from using logistic exposure models to examine the relationship between energetic condition and  
456 reproductive success because these models control for the bias that arises when early nest failures are not  
457 detected, which is not possible in logistic regression models, and it is more sensitive to changes in a bird's  
nest status (Shaffer, 2004).

459           **DETAILED HYPOTHESES AND METHODS FROM THE**  
460           **PREREGISTRATION**

461           **HYPOTHESES**

462           We measured two morphological proxy variables of energetic condition and observed reproductive success  
463           in grackles to test two hypotheses:

464           **H1 - There is a relationship between two different morphological indices of energetic**  
465           **condition: fat score and the scaled mass index.**

466           **Prediction 1:** Fat score and the scaled mass index will be positively correlated. This would indicate that  
467           these two indices measure the same trait, and it is likely they both are proxies for fat content.

468           **Prediction 1 alternative 1:** There is a negative correlation between fat score and the scaled mass index.  
469           This would indicate that there may be a tradeoff between the two indices where a larger value of the scaled  
470           mass index may measure muscle content rather than fat, and individuals with more muscle have less visible  
471           fat.

472           **Prediction 1 alternative 2:** There is no correlation between fat score and the scaled mass index. This  
473           indicates that these two variables do not measure the same trait. Fat score may not adequately capture a  
474           bird's energetic condition because birds may be selected to only store the minimal fat necessary to prevent  
475           starvation, while also minimizing the weight gain that would make them easier targets for predators  
476           (Barnett et al., 2015). Similarly, the scaled mass index could be heavily influenced by body size, therefore  
477           reflecting structural size rather than fat storage (Labocha & Hayes, 2012).

478           **H2 - Energetic condition (as measured by fat score and the scaled mass index) relates to**  
479           **reproductive success (measured as a binary variable of whether a female had one or more**  
480           **fledglings (1) or not (0), and whether a male defended a territory containing nests (1) or not**  
481           **(0)).**

482           **Prediction 2:** Morphological indices of energetic condition (fat score and the scaled mass index) will  
483           correlate positively with reproductive success. This would indicate that individuals with more fat, and  
484           therefore higher energy reserves, are better able to acquire the resources necessary for reproduction.

485           **Prediction 2 alternative 1:** Morphological indices of energetic condition (fat score and the scaled mass  
486           index) will correlate negatively with reproductive success. This indicates that individuals may make trade  
487           offs, with some acquiring more food and increasing their energy reserves, and others prioritizing  
488           reproductive activities over increasing energy reserves.

489           **Prediction 2 alternative 2:** Morphological indices of energetic condition (fat score and the scaled mass  
490           index) do not correlate with reproductive success. This indicates that other, potentially non-morphological,  
491           individual characteristics relate to reproductive success (i.e., cognition, nest site selection, breeding  
492           experience, predator vigilance, etc.).

493           **METHODS**

494           The methods below are based on the preregistration, with small changes as described in the Deviations  
495           from the planned methods section above.

496           **Planned Sample** Great-tailed grackles are caught in the wild in Tempe, Arizona using a variety of  
497           methods (e.g., walk-in trap, bownet, mist net). After capture we immediately process birds by attaching  
498           colored leg bands in unique combinations for individual identification, conducting morphological  
499           measurements of weight, tarsus length, flattened wing length, tail length, skull length, bill length and fat  
500           score (the amount of visible fat under the skin in the clavicle and abdomen as in Kaiser, 1993). Most  
501           grackles are released after completion of color band marking, measurements, and acquiring a blood sample.

502 A subset of grackles are held in aviaries for up to 6 months for behavioral testing, and then released back  
503 to the wild at their location of capture.

504 From March - August, we monitor the behavior of all color-marked grackles to determine their nesting  
505 status. We follow females carrying nesting materials to find their nest. We determine whether the male  
506 territory owner is color-marked as well. Then we check each nest approximately every day to determine the  
507 status based on the female's behavior (building, incubation, feeding nestlings, feeding fledglings, failed).

508 Individuals included in this sample will be those for which we have measures of energetic condition when  
509 they were adults. We will not include individuals whose data were collected as juveniles. As of 30 July  
510 2019, we have fledgling data for 14 females that exhibited breeding behavior (5 had 1+ fledgling, 9 had no  
511 fledglings) and breeding territory status for 10 males (7 territory holders, 3 non-territory holders, 2 not  
512 observed so not part of this sample). Therefore, the minimum sample size for H2 will be 24. The minimum  
513 sample size for H1 will be 72, because that is how many marked individuals we have biometric data for so  
514 far. However, we expect to be able to add to the sample size for both H1 and H2 before the end of this  
515 investigation in Tempe, Arizona. *UPDATE Oct 2020: In the second breeding season we had 20 females and*  
516 *20 males with reproductive success and energetic condition data.*

517 **Sample size rationale** We will continue to color mark as many grackles as possible, and collect  
518 biometric data and fat scores. Our current sample of reproductive success is small because the grackles in  
519 Tempe nest in very tall palms, making it difficult to determine nest status. However, we plan to collect  
520 additional reproductive success data during the breeding season in summer 2020. *UPDATE Oct 2020: In*  
521 *the second breeding season we had 20 females and 20 males with reproductive success and energetic*  
522 *condition data.*

523 **Data collection stopping rule** We will stop collecting data for this project in early August 2020 when  
524 research at the Tempe, Arizona field site will be finished.

525 **Open materials** Biometric measurement protocol:  
526 <https://gitlab.com/corinalogan/the-grackle-project/blob/master/protocolBiometrics.pdf>  
527 Nest check protocol:  
528 <https://gitlab.com/corinalogan/the-grackle-project/blob/master/protocolNestCheck.pdf>

529 **Open data** All data (Berens et al., 2020) are available at  
530 <https://knb.ecoinformatics.org/view/doi:10.5063/F1NZ862D> and at github (the provided code will load  
531 these files directly from github).

532 **Randomization and counterbalancing** There is no randomization or counterbalancing in this  
533 investigation.

534 **Blinding of conditions during analysis** No blinding is involved in this investigation.

#### 535 **Dependent Variables**

##### 536 **P1: correlation between fat and the scaled mass index**

- 537 1) Fat score [the amount of visible fat under the skin in the clavicle and abdomen reported as a score  
538 from 0 (no fat) to 8 (fat completely covers muscles and underside of the bird); Kaiser (1993)] *UPDATE*  
539 *Oct 2020: Fat score was heavily 0 skewed with few scores greater than one. To increase model fit we*  
540 *used a binomial response variable instead, where 0 is no fat and 1 is some fat observed under the skin.*

##### 541 **P2: energetic condition and reproductive success**

- 542 1) Female had one or more fledglings (yes, no)  
543 2) Male held a territory consisting of 1 to 3 clumped palms containing nests (yes, no)

544 **Independent Variables**

545 **P1: correlation between fat and the scaled mass index**

- 546 1) Scaled mass index using measures of body weight and tarsus length or flattened wing length (average  
547 of left and right as in Bleeker et al., 2005). We will choose the measure that is most correlated with  
548 body weight (Peig & Green, 2009).  
549 2) Season (non-breeding [Sep-Feb], breeding [Mar-Aug]). *UPDATE Oct 2020: The Season variable only*  
550 *includes 2 males in the breeding season category, thus we do not have a large enough sample to produce*  
551 *reliable estimates. We removed the Season variable from the model for males.*  
552 3) Random effect: Experimenter (because several different experimenters measure dependent variables on  
553 multiple different birds)

554 **P2: energetic condition and reproductive success**

- 555 1) Fat score
  - 556 • Note 1: if the fat score and the scaled mass index are positively correlated, then we will use only  
557 fat score in the model for P2. If they are not positively correlated, then we will add the scaled  
558 mass index as an independent variable in the P2 analysis
  - 559 • Note 2: if fat score and/or the scaled mass index vary by season (breeding or non-breeding), then  
560 we will only use the data from the breeding season to ensure that less time has elapsed between  
561 the collection of energetic condition and reproductive success variables
- 562 2) Temporarily held in aviaries for behavioral testing at any point during this study, because this may  
563 affect breeding behaviors (yes, no)
- 564 3) Random effect: Year (to determine whether conditions in a given breeding season similarly affected all  
565 grackle behavior and nest success)
- 566 4) Random effect: Bird ID (because there may be multiple measures of reproductive success for each bird)

567 **ANALYSIS PLAN**

568 *UPDATE Oct 2020:*

- 569 1) *We realized that the sexual dimorphism of male and female body sizes necessitates separate analyses.*  
570 *Therefore, we calculated SMI for males and females separately, ran separate models for each sex for the*  
571 *repeatability analysis, P1 and P2.*
- 572 2) *Fat score data were distributed such that the majority of scores were 0, with some 1's and very few*  
573 *higher numbers. This made it difficult to fit models using an ordinal regression. The function*  
574 *simulateResiduals, which we used to check our data, does not work with data in the ordinal family.*  
575 *Consequently, we used logistic regression where the dependent variable FatScore represents no fat (score =*  
576 *0), or some fat (score = 1)*
- 577 3) *Despite the data checking which indicated our model was not overdispersed or zero inflated, we could not*  
578 *get the fixed effects or random effect to converge using the Bayesian MCMCglmm. We found no*  
579 *improvement in model fit by tweaking the priors or iterations/burnin/thin options. Therefore, we fit these*  
580 *models using the function glmer, a frequentist framework.*

581 4) The Season variable only includes 2 males in the breeding season category, thus we do not have a large  
582 enough sample to produce reliable estimates. We removed the Season variable from the model for males.

583 We will **exclude** data that was collected from the grackles when they were released from the aviaries to  
584 avoid any confounds due to their time in the aviary (e.g., perhaps unlimited nutritious food in the aviaries  
585 affected their fat score). However, to validate that our measures of structural body size (tarsus length or  
586 wing length) are precise and accurate, we will measure twice a subset of grackles brought into aviaries -  
587 once when they are initially caught, and again up to 6 months later when we release them. We will then  
588 calculate the repeatability of these multiple measures. All other data included in this study will come only  
589 from wild-caught grackles (including the birds that were brought into the aviaries on their first capture).  
590 When **missing data** occur, the existing data for that individual will be included in the analyses for which  
591 their data exist. Analyses will be conducted in R [current version 4.0.5; R Core Team (2017)].

592 **Ability to detect actual effects** To begin to understand what kinds of effect sizes we will be able to  
593 detect given our sample size limitations, we used G\*Power Faul et al. (2009) to conduct power analyses  
594 based on confidence intervals. G\*Power uses pre-set drop down menus and we chose the options that were  
595 as close to our analysis methods as possible (listed in each analysis below). Note that there were no explicit  
596 options for GLMMs, thus the power analyses are only an approximation of the kinds of effect sizes we can  
597 detect. We realize that these power analyses are not fully aligned with our study design and that these  
598 kinds of analyses are not appropriate for Bayesian statistics (e.g., our MCMCglmm below), however we are  
599 unaware of better options at this time. Additionally, it is difficult to run power analyses because it is  
600 unclear what kinds of effect sizes we should expect due to the lack of data on this species for these  
601 particular research questions.

602 **Data checking** The data will be checked for overdispersion, underdispersion, zero-inflation, and  
603 heteroscedasticity with the DHARMA R package (Hartig, 2019) following methods by Hartig.

604 *P1 analysis: correlation between fat and the scaled mass index*

605 We will calculate the scaled mass index as described by Peig & Green (2009) using either tarsus or  
606 flattened wing length - whichever measure is most correlated with body weight (Peig & Green, 2009).

607 We use a Generalized Linear Mixed Model (GLMM; MCMCglmm function, MCMCglmm package;  
608 (Hadfield 2010)) with an ordinal distribution (for categorical variables in MCMCglmm) and probit link  
609 using 130,000 iterations with a thinning interval of 10, a burnin of 30,000, and minimal priors (V=1, nu=0)  
610 (Hadfield, 2014). We will ensure the GLMM shows acceptable convergence [lag time autocorrelation values  
611 <0.01; Hadfield (2010)], and adjust parameters if necessary to meet this criterion. We will determine  
612 whether an independent variable had an effect or not using the Estimate in the full model.

613 Where we have multiple measures of tarsus or flattened wing length, we will check that our measurements  
614 are repeatable using the rptR package (Stoffel et al., 2017).

615 To roughly estimate our ability to detect actual effects (because these power analyses are designed for  
616 frequentist statistics, not Bayesian statistics), we ran a power analysis in G\*Power with the following  
617 settings: test family=F tests, statistical test=linear multiple regression: Fixed model ( $R^2$  deviation from  
618 zero), type of power analysis=a priori, alpha error probability=0.05. We changed the power and the effect  
619 size until we reached an output that we project our sample size will be (n=90). The number of predictor  
620 variables was restricted to only the fixed effects because this test was not designed for mixed models. The  
621 protocol of the power analysis is here:

622 *Input:*

623 Effect size  $f^2 = 0.15$

624 err prob = 0.05

625 Power (1- err prob) = 0.86

626 Number of predictors = 3

This means that, with a sample size of 89, we would have an 86% chance of detecting a medium effect (approximated at  $f^2=0.15$  by Cohen, 1988).

636 *code shown in .rmd*

## P2 analysis: energetic condition and reproductive success

To model the effect of energetic condition on reproductive success, we will use two types of logistic mixed-effect models. Both types are supported in the literature, but are slightly different in the way in which the link function is specified. First, we will model reproductive success using a generalized linear mixed model framework with a logit link function (i.e. Milenkaya et al., 2015). We will also use a logistic exposure model that has a link function which accounts for the time interval between nest checks when estimating the probability of daily nest survival (Bolker, 2014; Shaffer, 2004). If fat score and the scaled mass index are positively correlated in P1, then we will use only fat score as the independent variable in this GLMM. If they are not positively correlated, we will include both as independent variables.

646 Previous research found a non-linear relationship between reproductive success and energetic condition  
647 variables (Milenkaya et al., 2015). To check whether this is occurring in our data, we will first plot our raw  
648 data to determine if we need to include a non-linear energetic condition independent variable into our  
649 model (i.e. FatScore<sup>2</sup>). Our dependent variable is binary, so to more clearly see the trends in the data, on  
650 the x-axis we will bin our energetic condition scores into 5 categories based on standard deviations (sd)  
651 around the mean (low = < 2 sd, moderately low = -2 sd to -1 sd, moderate = -1 sd to +1 sd, moderately  
652 high = +1 sd to +2 sd, high = > 2 sd). Then on the y-axis we will use the proportion of individuals in  
653 each category that had successful nests. *UPDATE Oct 2020: Because most individuals fell within the*  
654 *medium category when we grouped data using 1 standard deviation around the mean, we switched to using*  
655 *half standard deviation increments around the mean.*

656 A power analysis was conducted as above for P1 and the protocol reported here:

657 *Input*

Effect size  $f^2 = 0.15$

659 err prob = 0.05

660 Power (1- err prob) = 0.90

661 Number of predictors = 2

662 *Output:*

Noncentrality parameter = 13.2000000

664 Critical F = 3.1038387

665 Numerator df = 2

Denominator df = 85

667 Total sample size = 88

Actual power = 0.9020264

669 This means that, with a sample size of 88, we would have a 90% chance of detecting a medium effect  
670 (approximated at  $f^2=0.15$  by Cohen, 1988).

671 *code shown in .rmd*

672 **Do energetic condition variables vary by season?** *code shown in .rmd*

673 **Does energetic condition relate to reproductive success?** *code shown in .rmd*

674 **Does female energetic condition relate to the probability of daily nest survival?** Our measure  
675 of female nest success could be systematically biased against nests that failed early (Shaffer, 2004).  
676 Consequently, we also analyzed female reproductive success using a logistic exposure model. This type of  
677 model determines the factors affecting daily nest survival probability.

678 *code shown in .rmd*

## 679 ETHICS

680 This research is carried out in accordance with permits from the:

- 681 1) US Fish and Wildlife Service (scientific collecting permit number MB76700A-0,1,2)  
682 2) US Geological Survey Bird Banding Laboratory (federal bird banding permit number 23872)  
683 3) Arizona Game and Fish Department (scientific collecting license number SP594338 [2017], SP606267  
684 [2018], and SP639866 [2019])  
685 4) Institutional Animal Care and Use Committee at Arizona State University (protocol number 17-1594R)

## 686 AUTHOR CONTRIBUTIONS

687 **Berens:** Hypothesis development, data collection, revising/editing.

688 **Logan:** Study design, write up, revising/editing, materials/funding.

689 **Folsom:** Data collection, revising/editing.

690 **Sevchik** Data collection, revising/editing.

691 **Bergeron:** Data collection, revising/editing.

692 **McCune:** Hypothesis development, data collection, data analysis, write up, revising/editing.

## 693 FUNDING

694 This research is funded by the Department of Human Behavior, Ecology and Culture at the Max Planck  
695 Institute for Evolutionary Anthropology.

## ACKNOWLEDGEMENTS

697 We thank Aaron Blackwell and Ken Kosik for being the UCSB sponsors of the Cooperation Agreement  
 698 with the Max Planck Institute for Evolutionary Anthropology; and our research assistants for help with  
 699 trapping the grackles and collecting the biometric and nest/territory data: Aelin Mayer, Nancy Rodriguez,  
 700 Brianna Thomas, Aldora Messinger, Elysia Mamola, Michael Guillen, Rita Barakat, Adriana Boderash,  
 701 Olateju Ojekunle, August Sevchik, Justin Huynh, Amanda Overholt, and Michael Pickett.

## REFERENCES

- 703 Abolins-Abols, M., & Ketterson, E. D. (2017). Condition explains individual variation in mobbing  
 704 behavior. *Ethology*, 123(8), 495–502.
- 705 Aubret, F., Bonnet, X., Shine, R., & Lourdais, O. (2002). Fat is sexy for females but not males: The  
 706 influence of body reserves on reproduction in snakes (*vipera aspis*). *Hormones and Behavior*, 42(2),  
 707 135–147.
- 708 Barbraud, C., Lormée, H., & LeNevé, A. (2000). Body size and determinants of laying date variation in the  
 709 snow petrel *pagodroma nivea*. *Journal of Avian Biology*, 31(3), 295–302.
- 710 Barnett, C. A., Suzuki, T. N., Sakaluk, S. K., & Thompson, C. F. (2015). Mass-based condition measures  
 711 and their relationship with fitness: In what condition is condition? *Journal of Zoology*, 296(1), 1–5.
- 712 Barry, K. L., & Wilder, S. M. (2013). Macronutrient intake affects reproduction of a predatory insect.  
 713 *Oikos*, 122(7), 1058–1064.
- 714 Berens, J., Logan, C., Folsom, M., Sevchik, A., Bergeron, L., & McCune, K. (2020). Validating  
 715 morphological condition indices and their relationship with reproductive success in great-tailed grackles.  
 716 *Knowledge Network for Biocomplexity, Data package*. <https://doi.org/10.5063/7P8WSM>
- 717 Bleeker, M., Kingma, S. A., Szentirmai, I., Székely, T., & Komdeur, J. (2005). Body condition and clutch  
 718 desertion in penduline tit *remiz pendulinus*. *Behaviour*, 142, 1465–1478.
- 719 Blem, C. (1990). Avian energy storage. *Curr Ornithol*, 7, 59–113.
- 720 Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological  
 721 Bulletin*, 117(2), 187.
- 722 Bolker, B. (2014). Logistic regression, accounting for differences in exposure. *Version 09.30. 2014. RPubs*.
- 723 Carter, A. J., Feeney, W. E., Marshall, H. H., Cowlishaw, G., & Heinsohn, R. (2013). Animal personality:  
 724 What are behavioural ecologists measuring? *Biological Reviews*, 88(2), 465–475.
- 725 Center, N. N. C. D. (2020). *Climate data online global summary of the month*.  
 726 <https://www.ncdc.noaa.gov/cdo-web/>
- 727 Champagnon, J., Guillemain, M., Elmberg, J., Massez, G., Cavallo, F., & Gauthier-Clerc, M. (2012). Low  
 728 survival after release into the wild: Assessing “the burden of captivity” on mallard physiology and  
 729 behaviour. *European Journal of Wildlife Research*, 58(1), 255–267.
- 730 Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* 2nd edn. Erlbaum Associates,  
 731 Hillsdale.
- 732 Cornelius Ruhs, E., Vézina, F., & Karasov, W. H. (2019). Physiological and immune responses of  
 733 free-living temperate birds provided a gradient of food supplementation. *Physiological and Biochemical  
 734 Zoology*, 92(1), 106–114.
- 735 Costa, G. (2012). *Behavioural adaptations of desert animals*. Springer Science & Business Media.
- 736 Dawson, R. D., & Bortolotti, G. R. (1997). Are avian hematocrits indicative of condition? American  
 737 kestrels as a model. *The Journal of Wildlife Management*, 1297–1306.

- 738 Delciellos, A. C., Barros, C. dos S. de, Prevedello, J. A., Ferreira, M. S., Cerqueira, R., & Vieira, M. V.  
 739 (2018). Habitat fragmentation affects individual condition: Evidence from small mammals of the  
 740 brazilian atlantic forest. *Journal of Mammalogy*, 99(4), 936–945.
- 741 Deppe, J. L., Ward, M. P., Bolus, R. T., Diehl, R. H., Celis-Murillo, A., Zenjal, T. J., Moore, F. R.,  
 742 Benson, T. J., Smolinsky, J. A., Schofield, L. N., & others. (2015). Fat, weather, and date affect  
 743 migratory songbirds' departure decisions, routes, and time it takes to cross the gulf of mexico.  
 744 *Proceedings of the National Academy of Sciences*, 112(46), E6331–E6338.
- 745 Drent, R., & Daan, S. (1980). The prudent parent: Energetic adjustments in avian breeding 1. *Ardea*,  
 746 55(1–2), 225–252.
- 747 Ellers, J., Van Alphen, J. J., & Sevenster, J. G. (1998). A field study of size–fitness relationships in the  
 748 parasitoid asobara tabida. *Journal of Animal Ecology*, 67(2), 318–324.
- 749 English, M. D., Robertson, G. J., Peck, L. E., Pirie-Hay, D., Roul, S., & Mallory, M. L. (2018). Body  
 750 condition of american black ducks (*anas rubripes*) wintering in atlantic canada using carcass  
 751 composition and a scaled mass index. *Canadian Journal of Zoology*, 96(10), 1137–1144.
- 752 Erciyas, K., Gürsoy, A., Özsemir, A., & Barış, Y. (2010). Body mass and fat score changes in recaptured  
 753 birds during the autumn migration at the cernek ringing station in turkey. *The Ring*, 32(1-2), 3–15.
- 754 Fair, J., Whitaker, S., & Pearson, B. (2007). Sources of variation in haematocrit in birds. *Ibis*, 149(3),  
 755 535–552.
- 756 Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using g\* power 3.1:  
 757 Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160.
- 758 Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\* power 3: A flexible statistical power  
 759 analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2),  
 760 175–191.
- 761 Gessaman, J. (1999). Evaluation of some nonlethal methods of estimating avian body fat and lean mass.  
 762 *Proceedings of the 22nd International Ornithological Congress*. University of Natal Press, Durban, 2–16.
- 763 Gill, F. (1995). *Ornithology*. Freeman; Company.
- 764 Gosler, A. (1991). On the use of greater covert moult and pectoral muscle as measures of condition in  
 765 passerines with data for the great tit *parus major*. *Bird Study*, 38(1), 1–9.
- 766 Gosler, A. G., Greenwood, J. J., & Perrins, C. (1995). Predation risk and the cost of being fat. *Nature*,  
 767 377(6550), 621–623.
- 768 Guglielmo, C. G., McGuire, L. P., Gerson, A. R., & Seewagen, C. L. (2011). Simple, rapid, and  
 769 non-invasive measurement of fat, lean, and total water masses of live birds using quantitative magnetic  
 770 resonance. *Journal of Ornithology*, 152(1), 75.
- 771 Haas, C. A. (1998). Effects of prior nesting success on site fidelity and breeding dispersal: An experimental  
 772 approach. *The Auk*, 115(4), 929–936.
- 773 Hadfield, J. (2010). MCMC methods for multi-response generalized linear mixed models: The  
 774 MCMCglmm r package. *Journal of Statistical Software*, 33(2), 1–22.
- 775 Hadfield, J. (2014). *MCMCglmm course notes*.  
 776 <http://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf>
- 777 Hartig, F. (2019). *DHARMA: Residual diagnostics for hierarchical (multi-level / mixed) regression models*.  
 778 <http://florianhartig.github.io/DHARMA/>
- 779 Heidinger, I. M. M., Hein, S., & Bonte, D. (2010). Patch connectivity and sand dynamics affect  
 780 dispersal-related morphology of the blue-winged grasshopper *oedipoda caerulescens* in coastal grey  
 781 dunes. *Insect Conservation and Diversity*, 3(3), 205–212.

- 782 Henderson, L., Evans, N., Heidinger, B., Herborn, K., & Arnold, K. (2017). Do glucocorticoids predict  
783 fitness? Linking environmental conditions, corticosterone and reproductive success in the blue tit,  
784 *cyanistes caeruleus*. *Royal Society Open Science*, 4(10), 170875.
- 785 Huxley, J. (1932). *Problems of relative growth*. Dover Publications.
- 786 Jacobs, S. R., Elliott, K., Guigueno, M. F., Gaston, A. J., Redman, P., Speakman, J. R., & Weber, J.-M.  
787 (2012). Determining seabird body condition using nonlethal measures. *Physiological and Biochemical  
788 Zoology*, 85(1), 85–95.
- 789 Johnson, K., DuVal, E., Kielt, M., & Hughes, C. (2000). Male mating strategies and the mating system of  
790 great-tailed grackles. *Behavioral Ecology*, 11(2), 132–141.
- 791 Johnson, K., & Peer, B. D. (2001). *Great-tailed grackle: Quiscalus mexicanus*. Birds of North America,  
792 Incorporated.
- 793 Kaiser, A. (1993). A new multi-category classification of subcutaneous fat deposits of songbirds (una nueva  
794 clasificación, con multi-categorías, para los depósitos de grasa en aves canoras). *Journal of Field  
795 Ornithology*, 246–255.
- 796 Kelly, C. D., Tawes, B. R., & Worthington, A. M. (2014). Evaluating indices of body condition in two  
797 cricket species. *Ecology and Evolution*, 4(23), 4476–4487.
- 798 Labocha, M. K., & Hayes, J. P. (2012). Morphometric indices of body condition in birds: A review.  
799 *Journal of Ornithology*, 153(1), 1–22.
- 800 Labocha, M. K., Schutz, H., & Hayes, J. P. (2014). Which body condition index is best? *Oikos*, 123(1),  
801 111–119.
- 802 Langston, N. E., Freeman, S., Rohwer, S., & Gori, D. (1990). The evolution of female body size in  
803 red-winged blackbirds: The effects of timing of breeding, social competition, and reproductive  
804 energetics. *Evolution*, 44(7), 1764–1779.
- 805 Liao, C.-Y., Rikke, B. A., Johnson, T. E., Gelfond, J. A., Diaz, V., & Nelson, J. F. (2011). Fat maintenance  
806 is a predictor of the murine lifespan response to dietary restriction. *Aging Cell*, 10(4), 629–639.
- 807 Maceda-Veiga, A., Green, A. J., & De Sostoa, A. (2014). Scaled body-mass index shows how habitat  
808 quality influences the condition of four fish taxa in north-eastern spain and provides a novel indicator of  
809 ecosystem health. *Freshwater Biology*, 59(6), 1145–1160.
- 810 Maspons, J., Molowny-Horas, R., & Sol, D. (2019). Behaviour, life history and persistence in novel  
811 environments. *Philosophical Transactions of the Royal Society B*, 374(1781), 20180056.
- 812 Mayfield, H. (1961). Nesting success calculated from exposure. *The Wilson Bulletin*, 255–261.
- 813 McGuire, L. P., Kelly, L. A., Baloun, D. E., Boyle, W. A., Cheng, T. L., Clerc, J., Fuller, N. W., Gerson,  
814 A. R., Jonasson, K. A., Rogers, E. J., & others. (2018). Common condition indices are no more  
815 effective than body mass for estimating fat stores in insectivorous bats. *Journal of Mammalogy*, 99(5),  
816 1065–1071.
- 817 McNamara, J. M., Barta, Z., Houston, A. I., & Race, P. (2005). A theoretical investigation of the effect of  
818 predators on foraging behaviour and energy reserves. *Proceedings of the Royal Society B: Biological  
819 Sciences*, 272(1566), 929–934.
- 820 Merilä, J., & Svensson, E. (1997). Are fat reserves in migratory birds affected by condition in early life?  
821 *Journal of Avian Biology*, 279–286.
- 822 Milenkaya, O., Catlin, D. H., Legge, S., & Walters, J. R. (2015). Body condition indices predict  
823 reproductive success but not survival in a sedentary, tropical bird. *PLoS One*, 10(8), e0136582.
- 824 Milenkaya, O., Weinstein, N., Legge, S., & Walters, J. R. (2013). Variation in body condition indices of  
825 crimson finches by sex, breeding stage, age, time of day, and year. *Conservation Physiology*, 1(1).

- 826 Montreuil-Spencer, C. (2017). *Relationships between winter energetic condition and reproductive*  
827 *investment in a wild bird* [PhD thesis].
- 828 Murphy, M. T. (1986). Body size and condition, timing of breeding, and aspects of egg production in  
829 eastern kingbirds. *The Auk*, 103(3), 465–476.
- 830 Musacchia, X. (1953). A study of the lipids in arctic migratory birds. *The Condor*, 55(6), 305–312.
- 831 Nakagawa, S., & Schielzeth, H. (2010). Repeatability for gaussian and non-gaussian data: A practical  
832 guide for biologists. *Biological Reviews*, 85(4), 935–956.
- 833 Nip, E. J., Frei, B., & Elliott, K. H. (2019). Seasonal and temporal variation in scaled mass index of  
834 black-capped chickadees (*poecile atricapillus*). *The Canadian Field-Naturalist*, 132(4), 368–377.
- 835 Peig, J., & Green, A. J. (2009). New perspectives for estimating body condition from mass/length data:  
836 The scaled mass index as an alternative method. *Oikos*, 118(12), 1883–1891.
- 837 Pendlebury, C., MacLeod, M., & Bryant, D. (2004). Variation in temperature increases the cost of living in  
838 birds. *Journal of Experimental Biology*, 207(12), 2065–2070.
- 839 Perrins, C. (1970). The timing of birds ‘breeding seasons. *Ibis*, 112(2), 242–255.
- 840 Pond, C. (1981). Storage. *Physiological Ecology*, 190–219.
- 841 R Core Team. (2017). *R: A language and environment for statistical computing*. R Foundation for  
842 Statistical Computing. <https://www.R-project.org>
- 843 Shaffer, T. L. (2004). A unified approach to analyzing nest success. *The Auk*, 121(2), 526–540.
- 844 Stevenson, R., & Woods Jr, W. A. (2006). Condition indices for conservation: New uses for evolving tools.  
845 *Integrative and Comparative Biology*, 46(6), 1169–1190.
- 846 Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). rptR: Repeatability estimation and variance  
847 decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 8(11),  
848 1639–1644.
- 849 Walsberg, G. E. (1988). Evaluation of a nondestructive method for determining fat stores in small birds  
850 and mammals. *Physiological Zoology*, 61(2), 153–159.
- 851 Warnock, N., & Bishop, M. A. (1998). Spring stopover ecology of migrant western sandpipers. *The  
852 Condor*, 100(3), 456–467.
- 853 Wehtje, W. (2003). The range expansion of the great-tailed grackle (*quiscalus mexicanus gmelin*) in north  
854 america since 1880. *Journal of Biogeography*, 30(10), 1593–1607.
- 855 Welbergen, J. A. (2011). Fit females and fat polygynous males: Seasonal body mass changes in the  
856 grey-headed flying fox. *Oecologia*, 165(3), 629–637.
- 857 Wilder, S. M., Raubenheimer, D., & Simpson, S. J. (2016). Moving beyond body condition indices as an  
858 estimate of fitness in ecological and evolutionary studies. *Functional Ecology*, 30(1), 108–115.
- 859 Zhang, Y., Eyster, K., Liu, J.-S., & Swanson, D. L. (2015). Cross-training in birds: Cold and exercise  
860 training produce similar changes in maximal metabolic output, muscle masses and myostatin expression  
861 in house sparrows (*passer domesticus*). *Journal of Experimental Biology*, 218(14), 2190–2200.