

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

<sup>3</sup> 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>  
<sup>4</sup> McCune KB<sup>3\*</sup>

<sup>5</sup> 2021-06-30

<sup>6</sup> **Affiliations:**

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

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

<sup>11</sup> **Cite as:** Berens JM, Logan CJ, Folsom M, Sevchik A, Bergeron L, McCune KB. Submitted to PCI Ecology  
<sup>12</sup> Nov 2020. Validating morphological condition indices and their relationship with reproductive success in  
<sup>13</sup> 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**

<sup>20</sup> Morphological variation among individuals has the potential to influence multiple life history characteristics  
<sup>21</sup> such as dispersal, migration, reproductive success, and survival. Individuals that are in better “condition”  
<sup>22</sup> can disperse or migrate further or more successfully, have greater reproductive success, and survive longer,  
<sup>23</sup> particularly in years where environmental conditions are harsh. energetic condition is defined in various ways,  
<sup>24</sup> but is most often measured using an individual’s energetic or immune state. These traits are difficult to  
<sup>25</sup> measure directly, therefore a variety of morphological proxies to quantify energetic condition are used instead,  
<sup>26</sup> including fat score, weight, ratio of weight to tarsus length, a scaled mass index, as well as hematological  
<sup>27</sup> indices for immune system function. However, there is mixed support regarding whether these energetic  
<sup>28</sup> condition indices relate to life history characteristics, and whether the relationship is linear. Additionally,  
<sup>29</sup> although some investigations use multiple morphological proxies for energetic condition, rarely have there  
<sup>30</sup> been direct comparisons among proxies to validate that they measure the same trait. In this investigation,  
<sup>31</sup> we define condition as an energetic state and we attempt to measure it by comparing two indices (fat score  
<sup>32</sup> and the scaled mass index) to validate whether they measure the same trait and whether they correlate with  
<sup>33</sup> measures of reproductive success in our study system, the great-tailed grackle (*Quiscalus mexicanus*). We  
<sup>34</sup> found that the morphological proxies did not correlate with each other, indicating that they do not measure  
<sup>35</sup> the same trait. Further, neither proxy correlated with reproductive success in males, measured as whether  
<sup>36</sup> a male held a territory containing nests or not. We found that females with a high scaled mass index had

37 a significantly lower probability that their nest would survive on any given day. However, there was no  
38 relationship between female fat score and nest survival. These results indicate that measures of energetic  
39 condition should be validated before relying on their use as a condition proxy in grackles and birds in general.  
40 Future research should further investigate our unexpected result that higher scaled mass index correlated  
41 with lower nest survival to better understand the importance of energetic condition for reproductive success  
42 - a necessary component for selection to act.

## 43 INTRODUCTION

44 Morphological variation among individuals has the potential to influence multiple life history characteristics  
45 such as dispersal, migration, reproductive fitness, and survival (Wilder et al. 2016). One morphological  
46 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 and 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 weight to tarsus length, ratio of weight to wing chord length; Labocha et al. (2014)]. However, there  
61 is mixed support regarding whether and how these proxies relate to life history characteristics (Labocha  
62 et al. 2014; Wilder et al. 2016). A review conducted by Barnett (2015) shows that, while mass or body  
63 size measures of energetic condition are often assumed to have a positive linear relationship with fitness,  
64 this is not always the case, and the relationship should first be empirically validated before being used as a  
65 proxy (Barnett et al. 2015). In some instances, the proxy might relate to life history characteristics, but in  
66 a non-linear way. For example, theoretical simulations of small birds show that survival does not increase  
67 linearly with energy (i.e., fat) reserves (McNamara et al. 2005). If the reserves are too low, the individual  
68 is at risk of starvation. However, once the reserves get too high, the individual is at an increased risk of  
69 predation (McNamara et al. 2005). Thus, fat reserves can relate to a life history variable (survival), but in  
70 a U-shaped relationship rather than a linear one.

71 Although some studies use multiple morphological proxies for energetic condition (e.g., Warnock and 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 three  
75 estimates of energetic condition based on fat content or on the relationship between body mass and body  
76 length (scaled mass index or ordinary least squares regression) did not correlate with each other (Kelly et  
77 al. 2014), thus indicating that they do not measure the same trait. This is an example of the jingle fallacy  
78 (Block 1995; Carter et al. 2013), where a single trait label (“energetic condition”) actually encompasses  
79 more than one distinct trait. In this case, two investigations using different proxies can be conducted on the  
80 same research question, using the same species, but may end up with different results. This is problematic  
81 because inconsistency in results among researchers can result in potentially misleading interpretations of the  
82 impact of variation in morphology in relation to life history and population variables (Stevenson and Woods  
83 Jr 2006).

84 Here we compare two indices (fat score and the scaled mass index) of an individual’s energetic state to  
85 validate whether they correlate with each other, which would indicate that they both measure energetic  
86 condition. Fat score, as described by Kaiser (1993), is a numerical estimate of the amount of fat visible

under the skin (Fig. 1). The score ranges from 0 to 8 depending on the size and appearance of the fat located in the individual's abdomen and interclavicular depression, with 0 indicating no visible fat and 8 indicating extensive fat covering the ventral surface such that no muscle tissue is visible. For example, a score of 1 corresponds to sparse traces of fat visible in the interclavicular depression and abdomen. This measure is frequently used in birds (Merilä and Svensson 1997; Erciyas et al. 2010; Cornelius Ruhs et al. 2019), and is a straightforward, non-invasive method for estimating energetic condition. However, previous research found that it does not always positively relate with life history variables. For example, Haas (1998) found no difference between fat scores in individuals that had successful or failed nests in American robins and brown thrashers, indicating that fat score may not explain much of the variation in nest success in some species. Further research is needed to understand the relationship between fat score measures and life history characteristics.

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

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

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

Our study system is a population of great-tailed grackles (*Quiscalus mexicanus*), hereafter "grackles", in Tempe, Arizona. This system is ideal for this investigation because grackles are native to the tropical climates of Central America (Johnson and Peer 2001), but have rapidly expanded their geographic range into new areas (Wehtje 2003). Because grackles are a water-associated species, the desert habitat of Tempe presents physiological challenges that could lead to an increased likelihood of a tradeoff between survival and reproductive attempts (Henderson et al. 2017). Deserts are characterized by a scarcity of water and extreme temperature fluctuations, which require behavioral and physiological adaptations (Costa 2012). Wide variation in energetic condition and reproductive success is possible if grackle physiology requires more water than is present in the environment, and some individuals may cope with physiological stress, or find hidden sources of water, better than others (Henderson et al. 2017).



139

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

143 **HYPOTHESES**

144 We measured two proxies of energetic condition and observed reproductive success in grackles to test two  
 145 hypotheses:

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

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

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

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

160 **H2 - Energetic condition (as measured by fat score and the scaled mass index) relates to  
 161 reproductive success (measured as a binary variable of whether a female had one or more  
 162 fledglings (1) or not (0), and whether a male defended a territory containing nests (1) or not  
 163 (0)).**

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

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

171 **Prediction 2 alternative 2:** Morphological indices of energetic condition (fat score and the scaled mass  
 172 index) do not correlate with reproductive success. This indicates that other, potentially non-morphological,

173 individual characteristics relate to reproductive success (i.e., cognition, nest site selection, breeding experi-  
174 ence, predator vigilance, etc.).

175 **ASSOCIATED PREREGISTRATION**

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

185 **After pre-study peer review: Deviations from the planned methods**

- 186 1) We realized that the sexual dimorphism of male and female body sizes necessitates separate analyses.  
187 Therefore, we calculated SMI for males and females separately, and ran separate models for each sex  
188 for the repeatability analysis (P1 and P2).
- 189 2) Fat score data were distributed such that the majority of scores were 0, with some 1's and very few  
190 higher numbers. Specifically, 15 out of 21 males and 26 out of 47 females had fat score values of 0.  
191 This lack of variance in the response variable led to problems when we ran the models: it was difficult  
192 to fit models using an ordinal regression. The function “simulateResiduals”, which we used to check  
193 our data, does not work with data in the ordinal family. Consequently, we modified the model to use  
194 a logistic regression where the dependent variable FatScore is categorized as individuals that showed  
195 no visible fat ( $y = 0$ ), or some fat was present ( $y = 1$ ) where we combined all individuals that had fat  
196 score values of 1 or greater. Subsequent data checking indicated that these data were not zero-inflated  
197 or overdispersed.

198 **P1: correlation between SMI and Fat score**

- 199 3) Warning messages occurred during the repeatability analysis using the “rptR” package in R (Stoffel et  
200 al. 2017) indicating that the fit was singular, likely because the variance for the Experimenter random  
201 effect in the model for both female and male wing length was 0.001. We thus conducted an unregistered  
202 analysis where we confirmed that our repeatability values from the repeatability models were valid,  
203 despite the warning, by hand calculating repeatability following Nakagawa and Schielzeth (2010). The  
204 hand-calculated repeatabilities were nearly identical (female  $R = 0.5$ , male  $R = 0.71$ ) to the output  
205 from the rpt function.
- 206 4) Despite the data checking which indicated our model was not overdispersed or zero inflated, we could  
207 not get the fixed effects or random effect to converge using the Bayesian package in R “MCMCglmm”.  
208 We found no improvement in model fit by tweaking the priors or iterations/burnin/thin options. There-  
209 fore, we fit these models using the function glmer, a frequentist framework.
- 210 5) The Season variable only includes 2 males in the breeding season category, thus we do not have a  
211 large enough sample to produce reliable estimates. We removed the Season variable from the model  
212 for males.

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

- 214 6) Only two females had reproductive success data from more than one year in our study (2019 and 2020).  
215 Consequently, there were very few repeated measures in this sample and our random effect of bird ID  
216 accounted for zero variance. This led to a warning that our model fit was singular. Therefore, we  
217 removed the data for these females for 2020 so we could remove ID as a random effect from the model,  
218 which resulted in the model running without warnings. We removed the 2020 data for these females  
219 because their energetic condition data was collected in 2019 and these measures were more likely to  
220 relate to their 2019 reproductive success data than to their reproductive success in 2020.
- 221 7) The fit of the model analyzing the relationship between energetic condition and male reproductive  
222 success (ability to hold a territory containing female nests) was singular. The Year random effect  
223 accounted for zero variance in the data, so we removed it. The fit was still singular, but we retained  
224 the ID random effect (although it also explained zero variance) to account for repeated measures in  
225 this sample.
- 226 8) The model fit was again singular in our logistic exposure model because the Year random effect ex-  
227 plained zero variance in the data. We removed this random effect from the analysis.

Table 1: Table 1. 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

## 228 RESULTS

### 229 Prediction 1: correlation between SMI and Fat Score

230 Table 1: Sample sizes for P1 and P2. The “Breeding” and “Non-breeding season” categories refer to the  
 231 number of individuals measured in each season. The “Reprod. success” category represents the total number  
 232 of individuals in each year observed engaging in breeding behaviors. Note that the 2019 and 2020 reproductive  
 233 success sample sizes include some of the same individuals that were observed in both years. Whereas, the  
 234 “Prop. successful” category represents the proportion of the total individuals observed engaging in breeding  
 235 behaviors in each year that held a territory containing nests (males) or fledged young (females). Category  
 236 | Males | Females | ————— | — | —— | Fat Score | 22 | 47 | Breeding season Fat | 2 | 12 |  
 237 Non-breed. season Fat | 20 | 35 | SMI | 24 | 62 | Breeding season SMI | 6 | 24 | Non-breed. season SMI | 18  
 238 | 38 | Aviaries | 16 | 9 | Reprod. success 2019 | 8 | 9 | Reprod. success 2020 | 17 | 13 | Prop. successful 2019  
 239 | 0.63 | 0.22 |  
 240 Prop. successful 2020 | 0.47 | 0.54 | —————

241 We were able to calculate SMI for 24 males and 62 females, and fat score values were available for 22 males  
 242 and 47 females.

243 We found that wing length was more tightly correlated with body mass than tarsus length in both sexes,  
 244 therefore we used wing length in our SMI calculations (female n = 62, r = 0.26, p = 0.03; male n = 24, r =  
 245 0.35, p = 0.08). This allows us to account for as much variation in body mass as possible that is associated  
 246 with skeletal body size because leftover variation in body mass is more likely to relate to energetic condition.  
 247 Consequently, we used wing length in our calculation of SMI.

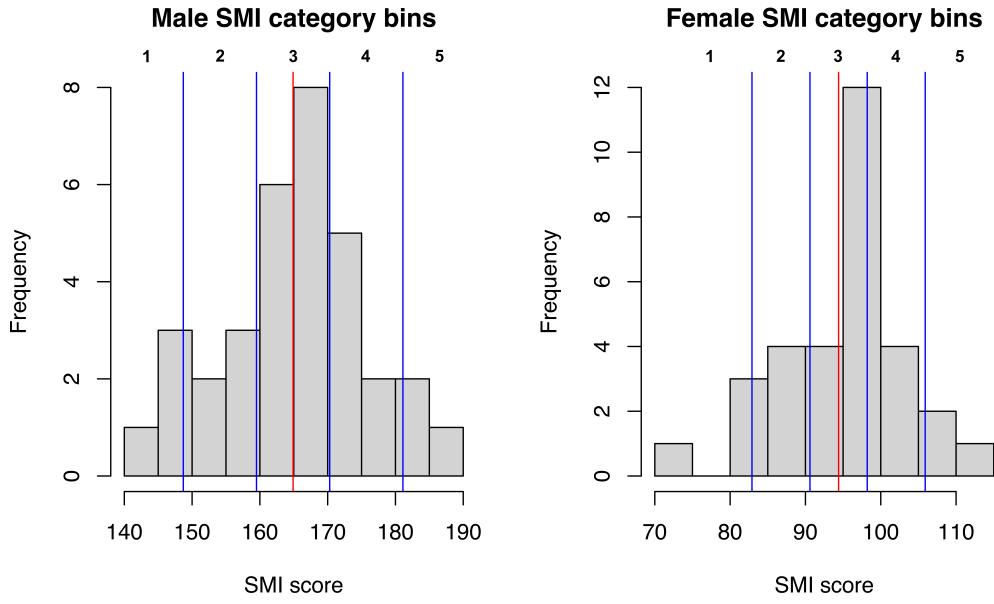
248 To validate that we were measuring structural body size consistently across experimenters, we analyzed the  
 249 repeatability of wing length in the birds in our sample that were measured more than once. We found that  
 250 average wing length was repeatable (n = 17 females, Repeatability  $\pm$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

253 We found that fat score was not correlated with SMI, which indicates that they are not measuring the same  
 254 trait (female  $p = 0.81$ ; male  $p = 0.50$ ; Table 1). There was also no relationship between season (breeding  
 255 or non-breeding) and female fat score ( $p = 0.71$ ). Only 2 males were measured during the breeding season,  
 256 therefore we omitted season as an independent variable in the male model.

### 257 P2: energetic condition and reproductive success

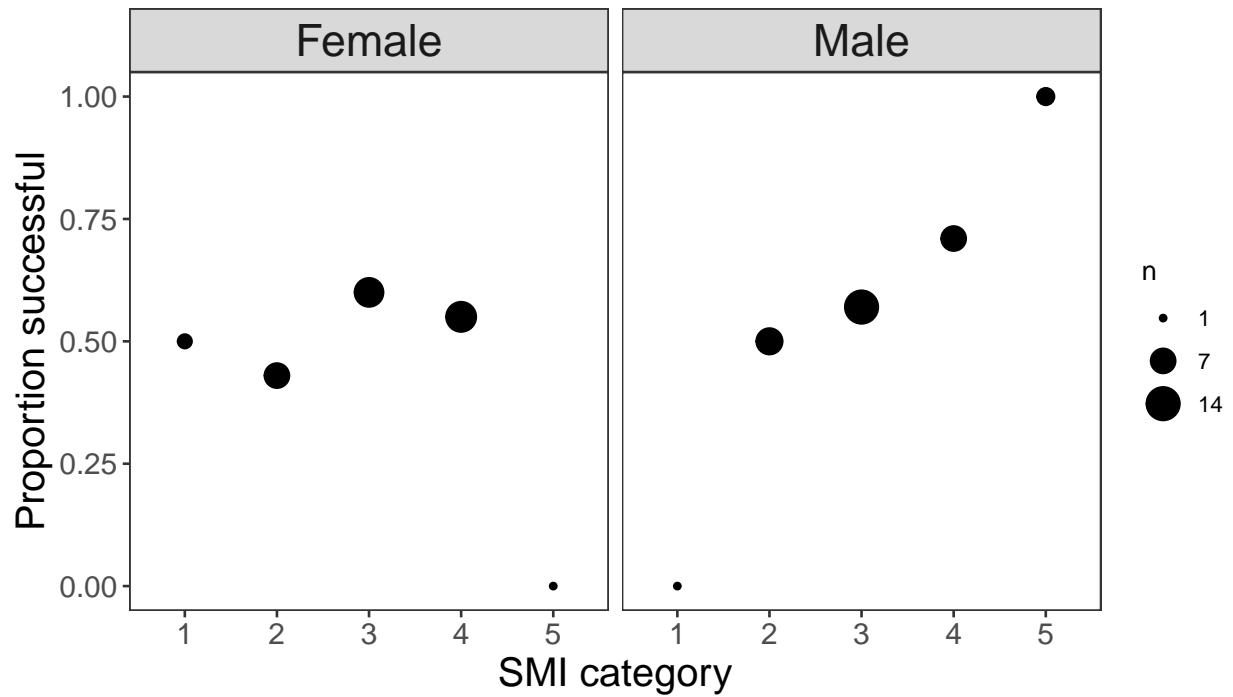
258 Our sample size for P2, where individuals had measures of reproductive success, SMI, and fat scores, was 20  
 259 for females and 20 for males.

260 To determine whether we should include any non-linear effects of SMI in our models [Milenkaya et al. (2015);  
 261 gosler1995predation], we visually evaluated whether individuals in any of 5 categories, ranging from low to  
 262 high SMI, were more likely to be reproductively successful (Fig. 2). We found no visual evidence for a  
 263 non-linear relationship between reproductive success and SMI for males or females (Fig. 3). Consequently,  
 264 we did not include non-linear terms in subsequent models.



265

266 Figure 2: Frequency histogram of the SMI scores, illustrating the SMI categories, for the 33 males and 31  
 267 females for which we also had reproductive success data. The mean SMI value is indicated by a red vertical  
 268 line. We created SMI category bins (indicated with vertical blue lines) in 1 standard deviation increments,  
 269 centered on the mean. Category 3 indicates the SMI value is close to the population mean value. Categories  
 270 1 and 2 are individuals that have SMI scores that are low, and moderately low, respectively, compared to  
 271 the population mean value. Similarly, categories 4 and 5 contain individuals that have SMI scores that are  
 272 moderately high and high, respectively, compared to the population mean value.

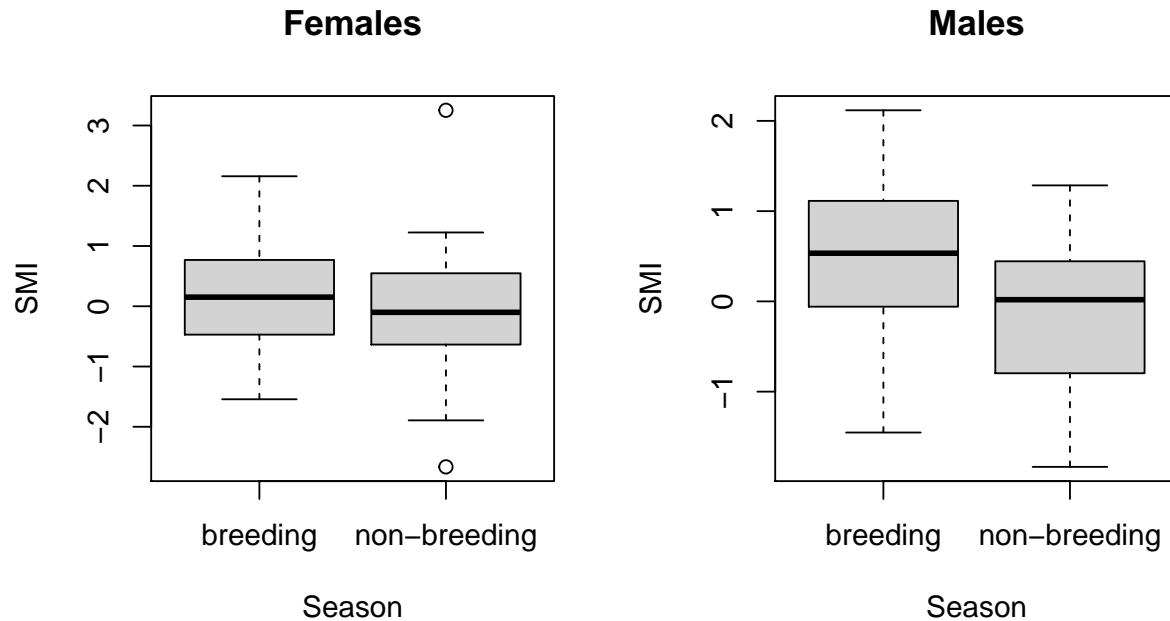


273

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

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

278 We used linear models to determine whether season would be important to include in our models testing P2.  
279 We found that neither SMI (female  $p = 0.26$ , male  $p = 0.15$ ) nor fat score (female  $p = 0.68$ , male  $p = 0.99$ )  
280 differed by season in females or males (Fig. 4). Although we note that, as stated above, we lack sufficient fat  
281 score data from males in the breeding season so results from that model should be interpreted with caution.  
282 Consequently, we did not include season as an independent variable in the P2 models.



283

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

286 Because fat score and SMI did not correlate, we included both as independent variables in our models testing  
287 prediction 2. We found that neither SMI ( $p = 0.13$ ), nor fat score ( $p = 0.82$ ) was associated with whether  
288 a female fledged offspring (Table 2). There was also no evidence of a relationship between the ability of a  
289 female to produce fledglings and having previously spent time in the aviaries ( $p = 0.22$ ). For males, the  
290 ability to defend a territory was also unrelated to either SMI ( $p = 0.13$ ) or fat score ( $p = 0.76$ ). Additionally,  
291 we found that those males who spent time in the aviaries were less likely to hold a territory compared with  
292 males who were never in the aviaries ( $p = 0.02$ ). However, we stress that our sample size was relatively  
293 small (20 males), and we did not have a balanced sample because there were no males that did not defend a  
294 territory and were never in the aviaries. Additionally, only five males had data from more than one breeding  
295 season, which resulted in our model fit being singular because the random effect for bird ID accounted for  
296 essentially zero variance. However, we kept ID in the model to account for the repeated samples.

Table 2: Table 2. 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*

297 **P2: energetic condition and probability of daily nest survival**

298 Logistic regression analyses to determine reproductive success from nests discovered in different stages will  
 299 be systematically biased (Shaffer 2004). Nests discovered at a more progressed stage (i.e., nestling stage  
 300 compared to building stage) are statistically more likely to succeed and nests with frequent and prolonged  
 301 adult visits (such as those that occur when nests survive longer) are more likely to be discovered. Therefore,  
 302 nests that fail early are less likely to be detected (Shaffer 2004). Consequently, we analyzed female repro-  
 303 ductive success using a logistic exposure model (Bolker 2014), which uses survival analysis to determine the  
 304 factors affecting the probability of daily nest survival, while accounting for incomplete nest observations. We  
 305 found that the probability of daily nest survival was significantly negatively related to SMI ( $p = 0.03$ ; Table  
 306 3), where, for every unit increase in SMI, the odds of daily nest survival decreased by half. This indicates  
 307 that a female with a larger SMI (more mass for her structural body size) was less likely to have her nest  
 308 survive each day (Fig. 5). There was no statistically significant relationship between the probability of daily  
 309 nest survival and fat score, day of the year, or time spent in the aviaries (Table 3). Although not statistically  
 310 significant, the effect size for the relationship between fat score and daily nest survival is large (Fig. 5) and  
 311 potentially biologically meaningful. The odds of a nest surviving on a given day are almost 2.5 times greater  
 312 for birds with some fat (a score of 1) compared to no fat (a score of 0).

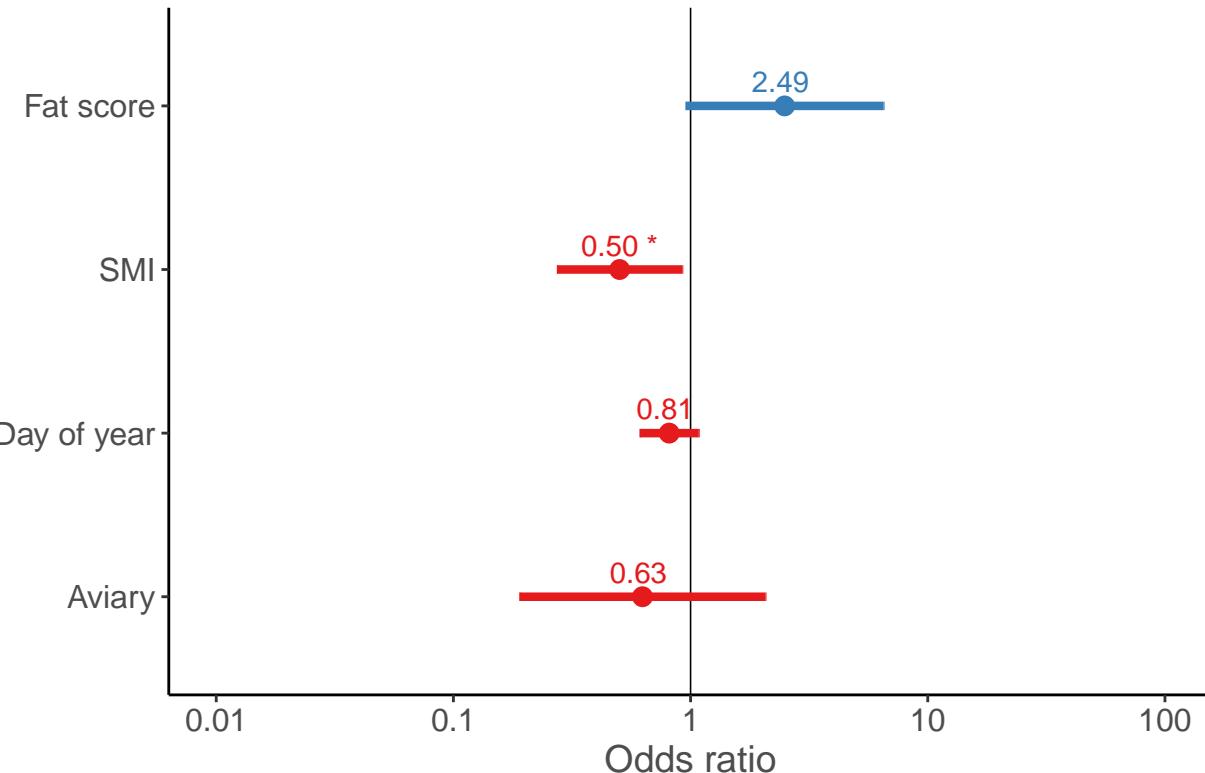
313 \begin{table}

314 \caption{Table 3. Results of the logistic exposure model showing the relationship between the probability  
 315 of daily nest survival and scaled mass index (SMI), fat score, the amount of time spent in the aviaries, and  
 316 the day of the year. Parameter estimates are presented with the standard error in parentheses. Odds ratios  
 317 (OR) represent the exponentiated estimates, and are presented to increase interpretability with 95%  
 318 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

320 \end{table}

## Probability of daily nest survival



321  
322 Figure 5: Odds ratios for independent variables affecting the probability of a nest surviving a given day.  
323 The dots and corresponding values represent the odds ratio values, and lines represent the confidence  
324 intervals around the odds ratio value. The vertical line at  $x = 1$  delineates the odds ratio value for no  
325 relationship between the estimates and the probability of daily nest survival. The asterisk indicates an  
326 odds ratio value that is statistically significant.

## 327 DISCUSSION

328 Although it is often implicitly assumed that most energetic condition proxies measure the same trait, we  
329 found that two proxies of energetic condition, fat score and SMI, did not correlate with each other in the  
330 great-tailed grackle, regardless of whether it was the breeding or non-breeding season. Further, we found  
331 that neither fat score nor SMI correlated with a female's ability to produce fledglings or a male's ability to  
332 hold a territory containing nests. However, we did find that the probability a female's nest will survive a  
333 given day is negatively related to SMI. These results have implications for the interpretation of results that  
334 are based on such proxies and for the use of these proxies in future research.

335 There are several potential reasons why grackle fat score and SMI did not correlate. First, it is possible  
336 that we were unable to accurately measure the amount of fat the birds actually stored. In addition to  
337 storing fat under their skin, birds may also store fat intraperitoneally (Musacchia 1953), which would not  
338 have been detected with our fat score measure. Second, variation in mass among grackles might have  
339 resulted from not only variation in fat content, but also from variation in muscle content (Labocha and  
340 Hayes 2012). However, measuring muscle content requires destructive methods [i.e. sacrificing the birds;  
341 Zhang et al. (2015)], which was beyond the scope of the current research program. Third, it is possible  
342 that fat score and SMI did not correlate due to experimenter error in collecting these measurements. We  
343 were unable to quantify the repeatability of our measures within and between experimenters because we  
344 did not collect repeated measurements on the same grackles when they were in hand (to reduce the amount

345 of processing time a bird experiences). Finally, our sample size might have been too small to detect an  
346 effect. However, the effect size for the relationship between fat score and SMI was essentially zero (0.001),  
347 therefore it is unlikely that a larger sample size would find a biologically informative relationship between  
348 these two proxies.

349 Although our first analysis of reproductive success, measured as the ability to produce fledglings (females)  
350 or to hold a territory containing nests (males), found no correlation with fat score or SMI, when we used  
351 logistic exposure models to determine whether female energetic condition related to the probability of daily  
352 nest survival, we found a negative relationship between SMI and the likelihood of daily nest survival. This  
353 result was surprising, but could be due to larger females actually carrying proportionally smaller energetic  
354 reserves than their smaller female counterparts, as seen in red-winged blackbirds (Langston et al. 1990). In  
355 some species, females with smaller body sizes are able to initiate breeding earlier because they can allocate  
356 more resources to reproduction compared to larger individuals that have higher bodily energy demands and  
357 therefore fewer excess energetic resources (Murphy 1986; Langston et al. 1990; Barbraud et al. 2000). This  
358 indirectly affects reproductive success because nesting earlier increases the probability of nesting success  
359 and multiple nesting attempts (Perrins 1970; Johnson and Peer 2001). Yet, we found no relationship  
360 between the probability of daily nest survival and day of the year, therefore this is unlikely to explain the  
361 negative relationship between SMI and nest survival. Alternatively, it is possible that larger females are  
362 unable to build a more concealed nest in the most dense vegetation, or that larger females are more likely to  
363 disrupt nest stability. The grackle nests were very high (often >10m above ground) and usually fairly well  
364 concealed, so we could not determine the causes of nest failure. Further investigations would be required to  
365 determine how energetic condition relates to specific threats to nesting success. In addition, the parameter  
366 estimate for the relationship between fat score and the daily probability of nest survival indicates that  
367 females with some visible fat are more than twice as likely to have a nest survive a given day. Because the  
368 direction of this effect is opposite to the relationship between SMI and nest survival, this is further  
369 evidence that these two proxies represent different traits, and that SMI is likely influenced by muscle mass.

370 Measurements of energetic condition are important for understanding variation in life history  
371 characteristics in studies across the animal kingdom. However, the results of this study highlight the need  
372 to better understand proxy measures of energetic condition, not only in grackles, but for birds in general.  
373 Most studies on avian energetic condition only use one proxy variable, but because energetic condition is  
374 difficult to measure directly, it is important to compare multiple proxy variables to ensure each proxy is  
375 measuring the intended trait [the jingle-jangle fallacy; Block (1995); Carter et al. (2013)]. Future research  
376 could add to this work by incorporating additional methods to measure energetic condition, for example,  
377 blood hematocrit levels (Dawson and Bortolotti 1997), protein storage (Houston et al. 1995), or by  
378 studying additional traits that could relate to variation in energy stores, such as dispersal (Ellers et al.  
379 1998) or survival (Liao et al. 2011). Furthermore, future research would benefit from using logistic  
380 exposure models to examine the relationship between energetic condition and reproductive success because  
381 these models control for the bias that arises when early nest failures are not detected, which is not possible  
382 in logistic regression models, and it is more sensitive to changes in a bird's nest status (Shaffer 2004).

## METHODS

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

386 **Planned Sample** Great-tailed grackles are caught in the wild in Tempe, Arizona using a variety of  
 387 methods (e.g., walk-in trap, bownet, mist net). After capture we immediately process birds by attaching  
 388 colored leg bands in unique combinations for individual identification, conducting morphological  
 389 measurements of weight, tarsus length, flattened wing length, tail length, skull length, bill length and fat  
 390 score (the amount of visible fat under the skin in the clavicle and abdomen as in Kaiser 1993). Most  
 391 grackles are released after completion of color band marking, measurements, and acquiring a blood sample.  
 392 A subset of grackles are held in aviaries for up to 6 months for behavioral testing, and then released back  
 393 to the wild at their location of capture.

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

398 Individuals included in this sample will be those for which we have measures of energetic condition when  
 399 they were adults. We will not include individuals whose data were collected as juveniles. As of 30 July  
 400 2019, we have fledgling data for 14 females that exhibited breeding behavior (5 had 1+ fledgling, 9 had no  
 401 fledglings) and breeding territory status for 10 males (7 territory holders, 3 non-territory holders, 2 not  
 402 observed so not part of this sample). Therefore, the minimum sample size for H2 will be 24. The minimum  
 403 sample size for H1 will be 72, because that is how many marked individuals we have biometric data for so  
 404 far. However, we expect to be able to add to the sample size for both H1 and H2 before the end of this  
 405 investigation in Tempe, Arizona. *UPDATE Oct 2020: In the second breeding season we had 20 females and*  
 406 *20 males with reproductive success and energetic condition data.*

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

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

415 **Open materials** Biometric measurement protocol:  
 416 <https://gitlab.com/corinalogan/the-grackle-project/blob/master/protocolBiometrics.pdf>  
 417 Nest check protocol:  
 418 <https://gitlab.com/corinalogan/the-grackle-project/blob/master/protocolNestCheck.pdf>

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

422 **Randomization and counterbalancing** There is no randomization or counterbalancing in this  
 423 investigation.

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

425           **Dependent Variables**

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

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

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

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

434           **Independent Variables**

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

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

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

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

457

## ANALYSIS PLAN

458

*UPDATE Oct 2020:*

459

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

462

*2) Fat score data were distributed such that the majority of scores were 0, with some 1's and very few higher numbers. This made it difficult to fit models using an ordinal regression. The function simulateResiduals, which we used to check our data, does not work with data in the ordinal family. Consequently, we used logistic regression where the dependent variable FatScore represents no fat (score = 0), or some fat (score = 1)*

467

*3) Despite the data checking which indicated our model was not overdispersed or zero inflated, we could not get the fixed effects or random effect to converge using the Bayesian MCMCglmm. We found no improvement in model fit by tweaking the priors or iterations/burnin/thin options. Therefore, we fit these models using the function glmer, a frequentist framework.*

471

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

473

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

482

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

492

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

494

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

495

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

497

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

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

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

512                          *Input:*  
513                          Effect size  $f^2 = 0.15$   
514                          err prob = 0.05  
515                          Power (1- err prob) = 0.86  
516                          Number of predictors = 3  
517                          *Output:*  
518                          Noncentrality parameter = 13.3500000  
519                          Critical F = 2.7119214  
520                          Numerator df = 3  
521                          Denominator df = 85  
522                          Total sample size = 89  
523                          Actual power = 0.8635760

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

526                          *code shown in .rmd*

527                          *P2 analysis: energetic condition and reproductive success*

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

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

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

547    *Input:*  
548    Effect size  $f^2 = 0.15$   
549    err prob = 0.05  
550    Power (1- err prob) = 0.90  
551    Number of predictors = 2  
552    *Output:*  
553    Noncentrality parameter = 13.2000000  
554    Critical F = 3.1038387  
555    Numerator df = 2  
556    Denominator df = 85  
557    Total sample size = 88  
558    Actual power = 0.9020264

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

561    *code shown in .rmd*

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

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

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

568    *code shown in .rmd*

## 569    ETHICS

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

- 571    1) US Fish and Wildlife Service (scientific collecting permit number MB76700A-0,1,2)
- 572    2) US Geological Survey Bird Banding Laboratory (federal bird banding permit number 23872)
- 573    3) Arizona Game and Fish Department (scientific collecting license number SP594338 [2017], SP606267  
574    [2018], and SP639866 [2019])
- 575    4) Institutional Animal Care and Use Committee at Arizona State University (protocol number 17-1594R)

## 576    AUTHOR CONTRIBUTIONS

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

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

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

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

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

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

583

## FUNDING

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

586

## ACKNOWLEDGEMENTS

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

592

## REFERENCES

- 593 Aubret F, Bonnet X, Shine R, Lourdais O. 2002. Fat is sexy for females but not males: The influence of  
 594 body reserves on reproduction in snakes (*vipera aspis*). Hormones and Behavior. 42(2):135–147.
- 595 Barbraud C, Lormée H, LeNevé A. 2000. Body size and determinants of laying date variation in the snow  
 596 petrel *pagodroma nivea*. Journal of Avian Biology. 31(3):295–302.
- 597 Barnett CA, Suzuki TN, Sakaluk SK, Thompson CF. 2015. Mass-based condition measures and their  
 598 relationship with fitness: In what condition is condition? Journal of Zoology. 296(1):1–5.
- 599 Barry KL, Wilder SM. 2013. Macronutrient intake affects reproduction of a predatory insect. Oikos.  
 600 122(7):1058–1064.
- 601 Berens J, Logan C, Folsom M, Sevchik A, Bergeron L, McCune K. 2020. Validating morphological  
 602 condition indices and their relationship with reproductive success in great-tailed grackles. Knowledge  
 603 Network for Biocomplexity. Data package. doi:10.5063/7P8WSM.
- 604 Bleeker M, Kingma SA, Szentirmai I, Székely T, Komdeur J. 2005. Body condition and clutch desertion in  
 605 penduline tit *remiz pendulinus*. Behaviour. 142:1465–1478.
- 606 Block J. 1995. A contrarian view of the five-factor approach to personality description. Psychological  
 607 bulletin. 117(2):187.
- 608 Bolker B. 2014. Logistic regression, accounting for differences in exposure. Version 0930 2014 RPubs.
- 609 Carter AJ, Feeney WE, Marshall HH, Cowlishaw G, Heinsohn R. 2013. Animal personality: What are  
 610 behavioural ecologists measuring? Biological Reviews. 88(2):465–475.
- 611 Champagnon J, Guillemain M, Elmberg J, Massez G, Cavallo F, Gauthier-Clerc M. 2012. Low survival  
 612 after release into the wild: Assessing ‘the burden of captivity’ on mallard physiology and behaviour.  
 613 European Journal of Wildlife Research. 58(1):255–267.
- 614 Cohen J. 1988. Statistical power analysis for the behavioral sciences 2nd edn.
- 615 Cornelius Ruhs E, Vézina F, Karasov WH. 2019. Physiological and immune responses of free-living  
 616 temperate birds provided a gradient of food supplementation. Physiological and Biochemical Zoology.  
 617 92(1):106–114.
- 618 Costa G. 2012. Behavioural adaptations of desert animals. Springer Science & Business Media.
- 619 Dawson RD, Bortolotti GR. 1997. Are avian hematocrits indicative of condition? American kestrels as a  
 620 model. The Journal of wildlife management.:1297–1306.
- 621 Delciellos AC, Barros C dos S de, Prevedello JA, Ferreira MS, Cerqueira R, Vieira MV. 2018. Habitat  
 622 fragmentation affects individual condition: Evidence from small mammals of the brazilian atlantic forest.  
 623 Journal of Mammalogy. 99(4):936–945.

- 624 Ellers J, Van Alphen JJ, Sevenster JG. 1998. A field study of size-fitness relationships in the parasitoid  
 625 asobara tabida. *Journal of Animal Ecology*. 67(2):318–324.
- 626 English MD, Robertson GJ, Peck LE, Pirie-Hay D, Roul S, Mallory ML. 2018. Body condition of american  
 627 black ducks (*anas rubripes*) wintering in atlantic canada using carcass composition and a scaled mass  
 628 index. *Canadian Journal of Zoology*. 96(10):1137–1144.
- 629 Erciyas K, Gürsoy A, Özsemir A, Barış Y. 2010. Body mass and fat score changes in recaptured birds  
 630 during the autumn migration at the cernek ringing station in turkey. *The Ring*. 32(1-2):3–15.
- 631 Faul F, Erdfelder E, Buchner A, Lang A-G. 2009. Statistical power analyses using g\* power 3.1: Tests for  
 632 correlation and regression analyses. *Behavior research methods*. 41(4):1149–1160.
- 633 Faul F, Erdfelder E, Lang A-G, Buchner A. 2007. G\* power 3: A flexible statistical power analysis  
 634 program for the social, behavioral, and biomedical sciences. *Behavior research methods*. 39(2):175–191.
- 635 Gill F. 1995. *Ornithology*. Freeman; Company.
- 636 Haas CA. 1998. Effects of prior nesting success on site fidelity and breeding dispersal: An experimental  
 637 approach. *The Auk*. 115(4):929–936.
- 638 Hadfield J. 2010. MCMC methods for multi-response generalized linear mixed models: The MCMCglmm r  
 639 package. *Journal of Statistical Software*. 33(2):1–22.
- 640 Hadfield J. 2014. MCMCglmm course notes.  
 641 <http://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf>.
- 642 Hartig F. 2019. DHARMA: Residual diagnostics for hierarchical (multi-level / mixed) regression models.  
 643 <http://florianhartig.github.io/DHARMA/>.
- 644 Heidinger IMM, Hein S, Bonte D. 2010. Patch connectivity and sand dynamics affect dispersal-related  
 645 morphology of the blue-winged grasshopper *oedipoda caerulescens* in coastal grey dunes. *Insect  
 646 Conservation and Diversity*. 3(3):205–212.
- 647 Henderson L, Evans N, Heidinger B, Herborn K, Arnold K. 2017. Do glucocorticoids predict fitness?  
 648 Linking environmental conditions, corticosterone and reproductive success in the blue tit, *cyanistes  
 649 caeruleus*. *Royal Society open science*. 4(10):170875.
- 650 Houston DC, Donnan D, Jones P, Hamilton I, Osborne D. 1995. Changes in the muscle condition of female  
 651 zebra finches *poephila guttata* during egg laying and the role of protein storage in bird skeletal muscle.  
 652 *Ibis*. 137(3):322–328.
- 653 Huxley J. 1932. *Problems of relative growth*. Dover Publications.
- 654 Johnson K, Peer BD. 2001. Great-tailed grackle: *Quiscalus mexicanus*. *Birds of North America*,  
 655 Incorporated.
- 656 Kaiser A. 1993. A new multi-category classification of subcutaneous fat deposits of songbirds (una nueva  
 657 clasificación, con multi-categorías, para los depósitos de grasa en aves canoras). *Journal of Field  
 658 Ornithology*.:246–255.
- 659 Kelly CD, Tawes BR, Worthington AM. 2014. Evaluating indices of body condition in two cricket species.  
 660 *Ecology and Evolution*. 4(23):4476–4487.
- 661 Labocha MK, Hayes JP. 2012. Morphometric indices of body condition in birds: A review. *Journal of  
 662 Ornithology*. 153(1):1–22.
- 663 Labocha MK, Schutz H, Hayes JP. 2014. Which body condition index is best? *Oikos*. 123(1):111–119.
- 664 Langston NE, Freeman S, Rohwer S, Gori D. 1990. The evolution of female body size in red-winged  
 665 blackbirds: The effects of timing of breeding, social competition, and reproductive energetics. *Evolution*.  
 666 44(7):1764–1779.
- 667 Liao C-Y, Rikke BA, Johnson TE, Gelfond JA, Diaz V, Nelson JF. 2011. Fat maintenance is a predictor of  
 668 the murine lifespan response to dietary restriction. *Aging cell*. 10(4):629–639.

- 669 Maceda-Veiga A, Green AJ, De Sostoa A. 2014. Scaled body-mass index shows how habitat quality  
 670 influences the condition of four fish taxa in north-eastern spain and provides a novel indicator of ecosystem  
 671 health. *Freshwater biology*. 59(6):1145–1160.
- 672 Mayfield H. 1961. Nesting success calculated from exposure. *The Wilson Bulletin*.:255–261.
- 673 McNamara JM, Barta Z, Houston AI, Race P. 2005. A theoretical investigation of the effect of predators  
 674 on foraging behaviour and energy reserves. *Proceedings of the Royal Society B: Biological Sciences*.  
 675 272(1566):929–934.
- 676 Merilä J, Svensson E. 1997. Are fat reserves in migratory birds affected by condition in early life? *Journal*  
 677 *of Avian Biology*.:279–286.
- 678 Milenkaya O, Catlin DH, Legge S, Walters JR. 2015. Body condition indices predict reproductive success  
 679 but not survival in a sedentary, tropical bird. *PLoS One*. 10(8):e0136582.
- 680 Murphy MT. 1986. Body size and condition, timing of breeding, and aspects of egg production in eastern  
 681 kingbirds. *The Auk*. 103(3):465–476.
- 682 Musacchia X. 1953. A study of the lipids in arctic migratory birds. *The Condor*. 55(6):305–312.
- 683 Nakagawa S, Schielzeth H. 2010. Repeatability for gaussian and non-gaussian data: A practical guide for  
 684 biologists. *Biological Reviews*. 85(4):935–956.
- 685 Peig J, Green AJ. 2009. New perspectives for estimating body condition from mass/length data: The  
 686 scaled mass index as an alternative method. *Oikos*. 118(12):1883–1891.
- 687 Perrins C. 1970. The timing of birds ‘breeding seasons. *Ibis*. 112(2):242–255.
- 688 R Core Team. 2017. R: A language and environment for statistical computing. Vienna, Austria: R  
 689 Foundation for Statistical Computing. <https://www.R-project.org>.
- 690 Shaffer TL. 2004. A unified approach to analyzing nest success. *The Auk*. 121(2):526–540.
- 691 Stevenson R, Woods Jr WA. 2006. Condition indices for conservation: New uses for evolving tools.  
 692 *Integrative and comparative biology*. 46(6):1169–1190.
- 693 Stoffel MA, Nakagawa S, Schielzeth H. 2017. rptR: Repeatability estimation and variance decomposition  
 694 by generalized linear mixed-effects models. *Methods in Ecology and Evolution*. 8(11):1639–1644.
- 695 Warnock N, Bishop MA. 1998. Spring stopover ecology of migrant western sandpipers. *The Condor*.  
 696 100(3):456–467.
- 697 Wehtje W. 2003. The range expansion of the great-tailed grackle (*quiscalus mexicanus gmelin*) in north  
 698 america since 1880. *Journal of Biogeography*. 30(10):1593–1607.
- 699 Wilder SM, Raubenheimer D, Simpson SJ. 2016. Moving beyond body condition indices as an estimate of  
 700 fitness in ecological and evolutionary studies. *Functional Ecology*. 30(1):108–115.
- 701 Zhang Y, Eyster K, Liu J-S, Swanson DL. 2015. Cross-training in birds: Cold and exercise training  
 702 produce similar changes in maximal metabolic output, muscle masses and myostatin expression in house  
 703 sparrows (*passer domesticus*). *Journal of Experimental Biology*. 218(14):2190–2200.