

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

³ Berens JM¹ Logan CJ² Folsom M² Sevchik A¹ Bergeron L³
⁴ McCune KB^{3*}

2020-11-12

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

¹⁰ *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.

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

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

- ¹⁸ • Reviewers: Javier Seoane and Isabel López-Rull

19 ABSTRACT

Morphological variation among individuals has the potential to influence multiple life history characteristics such as dispersal, migration, reproductive success, and survival (Wilder et al. 2016). Research has shown that individuals that are in better “condition” can disperse or migrate further or more successfully, have greater reproductive success, and survive for longer (Heidinger et al. 2010; Liao et al. 2011; Wilder et al. 2016), particularly in years where environmental conditions are harsh (Milenkaya et al. 2015). An individual’s body condition can be defined in various ways, but is most often considered an individual’s energetic or immune state (Milenkaya et al. 2015). Since these traits are hard to measure directly, researchers have instead used a variety of morphological proxies to quantify condition such as fat score (Kaiser 1993), weight, ratio of weight to tarsus length (Labocha et al. 2014), a scaled mass index (Peig and Green 2009), as well as hematological indices for immune system function (Fleskes et al. 2017; Kraft et al. 2019). However, there is mixed support regarding whether these condition indices relate to life history characteristics (Labocha et al. 2014; Wilder et al. 2016), and whether the relationship is linear (McNamara et al. 2005; Milenkaya et al. 2015). Additionally, although some researchers use multiple morphological proxies for condition (e.g. Warnock and Bishop 1998), 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 indices (fat score and the scaled mass index) to validate whether they measure the same trait in our study system, the great-tailed grackle (*Quiscalus mexicanus*). We found that the morphological proxy

37 variables did not correlate with each other, indicating that they do not measure the same trait. Further,
38 neither proxy correlated with reproductive success in males, measured as whether a male held a territory
39 containing nests or not. We found that females with a high scaled mass index had a significantly lower
40 probability that their nest would survive on any given day. However, there was no relationship between
41 female fat score and nest survival. These results improve our understanding of measures of condition in
42 grackles and birds in general. Future research should further investigate our unexpected result that higher
43 scaled mass index led to lower nest survival to better understand the importance of energetic condition for
44 reproductive success - a necessary component for selection to act.

45 INTRODUCTION

46 Morphological variation among individuals has the potential to influence multiple life history characteristics
47 such as dispersal, migration, reproductive fitness, and survival (Wilder et al. 2016). One morphological
48 trait that might be particularly likely to influence these life history characteristics is energetic condition.
49 Research has shown that individuals that are in better “condition” can disperse or migrate further or more
50 successfully, have greater reproductive success, and survive for longer (Heidinger et al. 2010; Liao et al.
51 2011; Wilder et al. 2016), particularly in years where environmental conditions are harsh (Milenkaya et al.
52 2015). For example, a study conducted on vipers showed that while the level of fat reserves in males had
53 no effect on their sexual activity, females with low fat reserves engaged in sexual interactions less frequently
54 than those with higher fat reserves (Aubret et al. 2002). In contrast, mantids showed conflicting results
55 regarding the relationship between fat reserves and reproductive success. Female mantids were fed either a
56 high protein, low lipid diet, or a high lipid, low protein diet. The females that received the high lipid diet
57 had higher lipid content in most parts of their body compared to that of their high protein diet counterparts.
58 However, they were not able to produce even half as many eggs as the females fed the high protein, low
59 lipid diet. This led to lower male attraction, measured by the number of copulation events, thus negatively
60 impacting further reproductive success (Barry and Wilder 2013).

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

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

85 Here we compare two indices (fat score and the scaled mass index) of an individual’s energetic state to validate
86 whether they correlate with each other, which would indicate that they both measure body condition. Fat

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

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

111
$$Mass_i \left[\frac{AvgLength_p}{Length_i} \right]^{slope_p}$$
 where $Mass_i$ is each individual's weight in grams, $Length_i$ is the value of the
112 chosen measure of structural body length for each bird, $AvgLength_p$ is the average structural body length
113 in the population, and $slope_p$ is the value of the slope from a standard major axis regression of structural
114 body size on mass (Peig and Green 2009) which is used to compare variables that were both measured and
115 so have residual error. Therefore, individuals in better energetic condition (larger weight for their structural
116 body size) will have a higher SMI compared to individuals in poor condition. Studies across taxa found that
117 the SMI relates positively to reproductive success and survival. For example, mallards with a lower SMI
118 had lower rates of survival compared to their higher SMI counterparts (Champagnon et al. 2012), while in
119 crimson finches SMI was positively related to the number of young that survived to independence (Milenkaya
120 et al. 2015).

121 Our research will determine whether these two indices of energetic condition measure the same trait, and
122 whether this trait relates to an important life history characteristic- reproductive success. To quantify
123 reproductive success in birds, researchers must find and monitor nests. However, nests are usually built in
124 cryptic locations and parents behave secretly. Additionally, it is very difficult and time-consuming to track
125 the survival of offspring once they leave the nest. Therefore, the predominantly used measure of reproductive
126 success is whether a nest fledged offspring (Mayfield 1961).

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



137

138 Figure 1: A male grackle showing the yellow/orange tint of fat unnder the skin in the intraclavicular de-
139 pression, and a female grackle showing *no* fat under the skin of the intraclavicular region but significant fat
140 deposits under the skin of the abdomen.

141 **HYPOTHESES**

142 We measured two proxies of body condition and observed reproductive success in grackles to test two
143 hypotheses:

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

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

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

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

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

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

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

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

172 **ASSOCIATED PREREGISTRATION**

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

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

- 182 1) We realized that the sexual dimorphism of male and female body sizes necessitates separate analyses.
183 Therefore, we calculated SMI for males and females separately, ran separate models for each sex for
184 the repeatability analysis, P1 and P2.
- 185 2) Fat score data were distributed such that the majority of scores were 0, with some 1's and very
186 few higher numbers. This made it difficult to fit models using an ordinal regression. The function
187 "simulateResiduals", which we used to check our data, does not work with data in the ordinal family.
188 Consequently, we used logistic regression where the dependent variable FatScore represents no fat
189 (score = 0), or some fat (score = 1).

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

- 191 3) Warning messages occurred during the repeatability analysis using the "rptR" package in R (Stoffel
192 et al. 2017) indicating that the fit was singular, likely because the variance for the Experimenter
193 random effect in the model for both female and male wing length was 0.001. We thus conducted an
194 unregistered analysis where we confirmed that our repeatability values from the repeatability models
195 were valid despite the warning by hand calculating repeatability following Nakagawa and Schielzeth
196 (2010). The hand-calculated repeatabilities were nearly identical (female R = 0.5, male R = 0.71) to
197 the output from the rpt function.
- 198 4) Despite the data checking which indicated our model was not overdispersed or zero inflated, we could
199 not get the fixed effects or random effect to converge using the Bayesian package in R "MCMCglmm".
200 We found no improvement in model fit by tweaking the priors or iterations/burnin/thin options. There-
201 fore, we fit these models using the function glmer, a frequentist framework.
- 202 5) The Season variable only includes 2 males in the breeding season category, thus we do not have a
203 large enough sample to produce reliable estimates. We removed the Season variable from the model
204 for males.

205 **P2: body condition and reproductive success**

- 206 6) Only two females had reproductive success data from more than one year in our study (2019 and 2020).
207 Consequently, there were very few repeated measures in this sample and our random effect of bird ID
208 accounted for zero variance. This led to a warning that our model fit was singular. Therefore, we
209 removed the data for these females for 2020 so we could remove ID as a random effect from the model,
210 which resulted in the model running without warnings. We removed the 2020 data for these females
211 because their condition data was collected in 2019 and these measures were more likely to relate to
212 their 2019 reproductive success data than to their reproductive success in 2020.
- 213 7) The fit of the model analyzing the effect of body condition on male reproductive success (ability to
214 hold a territory containing female nests) was singular. The Year random effect accounted for zero
215 variance in the data, so we removed it. The fit was still singular, but we retained the ID random effect
216 (although it also explained zero variance) to account for repeated measures in this sample.

Table 1: Results from the logistic mixed-effect regression for females and fixed-effect regression for 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

217 8) The model fit was again singular in our logistic exposure model because the Year random effect ex-
 218 plained zero variance in the data. We removed this random effect from the analysis.

219 RESULTS

220 Prediction 1: correlation between SMI and Fat Score

221 We calculated SMI for 24 males and 62 females. We had fat score values on 21 males and 47 females.

222 We found that wing length was more tightly correlated with body mass than tarsus length in both sexes,
 223 therefore we used wing length in our SMI calculations (female n = 62, r = 0.26, p = 0.03; male n = 24,
 224 r = 0.35, p = 0.08). This allows us to account for as much variation in body mass as possible that is
 225 associated with skeletal body size because leftover variation in body mass is more likely to relate to energetic
 226 condition. Consequently, we used wing length in our calculation of SMI as: $Mass_i \left[\frac{AvgWing_p}{AvgWing_i} \right]^{slope_p}$. $Mass_i$
 227 is each individual's weight in grams, $AvgWing_i$ is the average value of the measures of the left and right
 228 wing lengths of each bird, $AvgWing_p$ is the average wing length in the population, and $slope_p$ is the value
 229 of the slope from a standard major axis regression of wing length on mass (Peig and Green 2009).

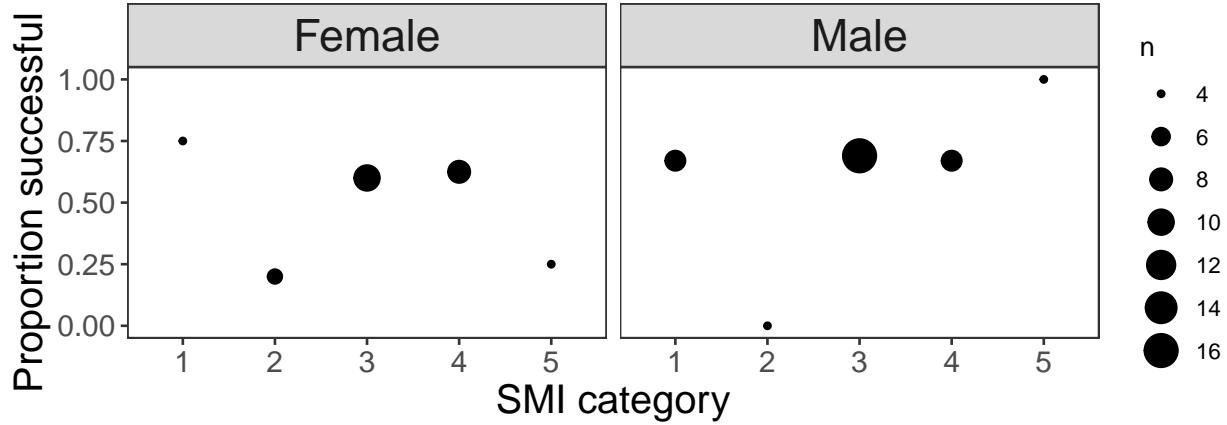
230 To validate that we were measuring structural body size consistently across experimenters, we analyzed the
 231 repeatability of wing length in the birds in our sample that were measured more than once. We found that
 232 average wing length was repeatable (n = 17 females, Repeatability \pm standard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

235 We found that fat score was not correlated with SMI, which indicates that they are not measuring the same
 236 trait (female p = 0.81; male p = 0.50; Table 1). There was also no effect of season (breeding or non-breeding)
 237 on female fat score (p = 0.71). Only 2 males were measured during the breeding season, therefore we omitted
 238 season as an independent variable in the male model.

239 P2: body condition and reproductive success

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

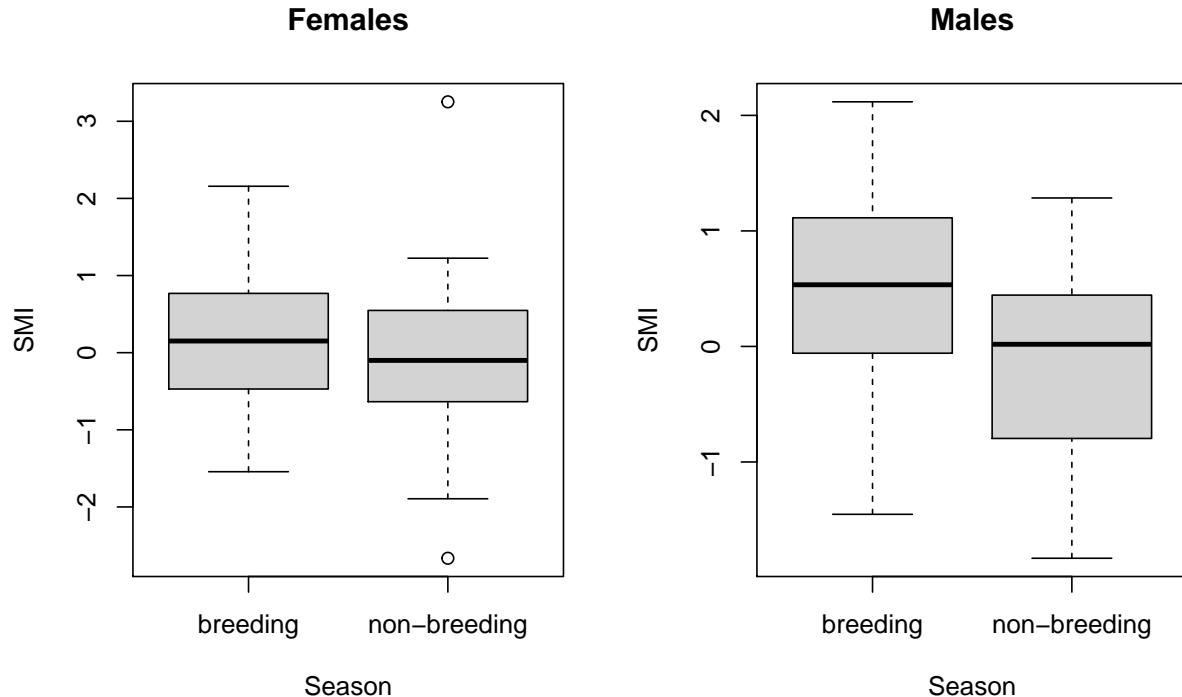
242 In some cases body condition shows a non-linear relationship with reproductive success (Milenkaya et al.
 243 2015). To test for this we calculated the SMI categories using 0.5 standard deviation increments around
 244 the mean to determine if individuals in some categories were more likely to be reproductively successful.
 245 However, we found no evidence for a non-linear relationship between reproductive success and SMI for males
 246 or females (Fig. 2).



247

248 Figure 2: The proportion of individuals that successfully fledged nests (females: left)
 249 or held a territory (males: right) in low (1), moderately low (2), moderate (3), moderately high (4) and high (5) SMI categories.
 250 Dots are sized according to the number (n) of individuals in that category. There is no evidence of a non-linear
 251 relationship.

252 Neither SMI nor fat score differed by season in males or females (Fig. 3) so we did not include season as an
 253 independent variable in the final models.



254

255 Figure 3: Scaled mass index (SMI) was not significantly different between the breeding and non-breeding
 256 seasons for either sex.

257 Because fat score and SMI did not correlate, we included both as independent variables in our models testing
 258 prediction 2. We found that neither SMI ($p = 0.13$), nor fat score ($p = 0.82$) was associated with a female's
 259 ability to produce fledglings (Table 2). There was also no evidence of an effect on the ability of a female
 260 to produce fledglings after having spent time in the aviaries ($p = 0.22$). For males, the ability to defend a
 261 territory was also unrelated to either SMI ($p = 0.13$) or fat score ($p = 0.76$). Additionally, we found that

Table 2: Results from the logistic regression for females and males to test whether condition relates to reproductive success. 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

Table 3: Results of the logistic exposure model showing the relationship between the probability of daily nest survival and scaled mass index (SMI), fat score, the amount of time spent in the aviaries, and the day of the year. Odds ratios (OR) are the exponentiated estimates to increase interpretability.

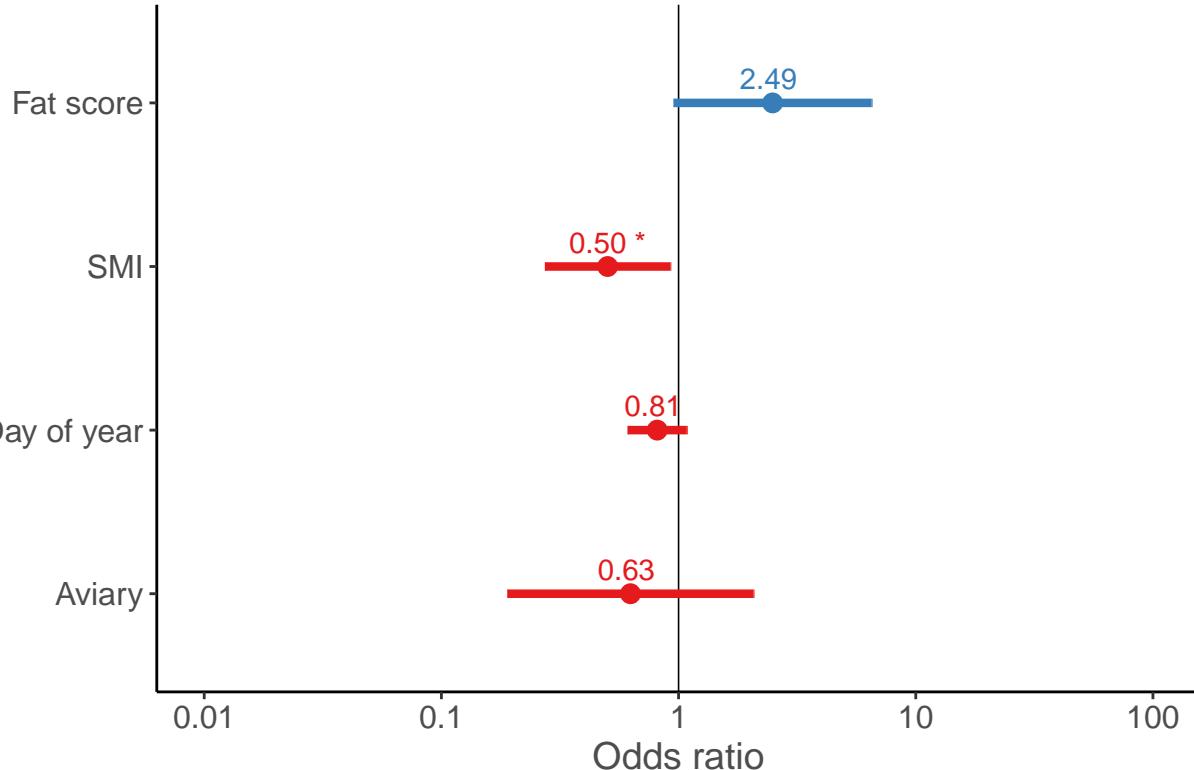
Parameter	Estimate (SE)	OR (CI)	p-value
Intercept	1.99 (0.40)	7.32 (3.3-16.0)	<0.001
Fat score	0.91 (0.49)	0.50 (0.27-0.92)	0.06
SMI	-0.69 (0.31)	2.48 (0.95-6.49)	0.03*
Day of year	-0.21 (0.15)	0.63 (0.19-2.10)	0.16
Aviary	-0.47 (0.61)	0.81 (0.60-1.10)	0.44

262 those males who spent more time in the aviaries were less likely to hold a territory compared with males
 263 who were never in the aviaries or who spent less time in the aviaries ($p = 0.02$). However, we stress that
 264 our sample size was relatively small (20 males), and we did not have a balanced sample because there were
 265 no males that failed to defend a territory and were never in the aviaries. Additionally, only five males had
 266 data from more than one breeding season, which resulted in our model fit being singular because the random
 267 effect for bird ID accounted for essentially zero variance. However, we kept ID in the model to account for
 268 the repeated samples.

269 P2: body condition and probability of daily nest survival

270 Logistic regression analyses to determine reproductive success from nests discovered in different stages will
 271 be systematically biased (Shaffer 2004). Nests discovered at a more progressed stage (i.e. nestling stage
 272 compared to building stage) are statistically more likely to succeed and nests with frequent and prolonged
 273 adult visits (such as those that occur when nests survive longer) are more likely to be discovered. Therefore,
 274 nests that fail early are less likely to be detected (Shaffer 2004). Consequently, we also analyzed female
 275 reproductive success using a logistic exposure model (Bolker 2014), which uses survival analysis to determine
 276 the factors affecting the probability of daily nest survival, while accounting for incomplete nest observations.
 277 We found that the probability of daily nest survival was significantly negatively related to SMI ($p = 0.03$;
 278 Table 3), where for every unit increase in SMI, the odds of daily nest survival decreased by half. This
 279 indicates that a female with a larger SMI (more mass for her structural body size) was less likely to have her
 280 nest survive each day (Fig. 4). There was no statistically significant relationship between the probability
 281 of daily nest survival and fat score, day of the year, or time spent in the aviaries (Table 3). Although not
 282 statistically significant, the effect size for the relationship between fat score and daily nest survival is large
 283 (Fig. 4) and potentially biologically meaningful. The odds of a nest surviving on a given day are almost 2.5
 284 times greater for birds with some fat (a score of 1) compared to no fat (a score of 0).

Probability of daily nest survival



285
286 Figure 4: Odds ratios for independent variables affecting the probability of a nest surviving a given day.
287 The dots and corresponding values represent the odds ratio values, lines represent the confidence intervals
288 around the odds ratio value. The vertical line at $x = 1$ delineates the odds ratio value for no relationship
289 between the estimates and the probability of daily nest survival. The asterisks indicates an odds ratio value
290 that is statistically significant.

291 DISCUSSION

292 Although researchers implicitly assume that most condition proxies measure the same trait, we found that
293 the two proxies of energetic condition, fat score and SMI, did not correlate with each other in the great-tailed
294 grackle regardless of whether it was breeding or non-breeding season. Further, we found that neither fat
295 score nor SMI correlated with a female's ability to produce fledglings or a male's ability to hold a territory
296 containing nests. However, we did find that the probability a female's nest will survive a given day is
297 influenced by SMI. These results have implications for the interpretation of results that are based on such
298 proxies and for the use of such proxies in future research.

299 There are several potential reasons why grackle fat score and SMI did not correlate. First, it is possible that
300 we were unable to accurately measure the amount of fat the birds actually stored. In addition to storing
301 fat under their skin, birds may also store fat intraperitoneally (Musacchia 1953), which would not have
302 been detected with our fat score measure. Second, variation in mass among grackles might have resulted
303 from not only variation in fat content, but also from variation in muscle content (Labocha and Hayes 2012).
304 However, measuring muscle content requires destructive methods (i.e. sacrificing the birds; Zhang et al.
305 2015), which was beyond the scope of the current research program. Third, it is possible that fat score
306 and SMI did not correlate due to experimenter error in collecting these measurements. We were unable
307 to quantify the repeatability of our measures within and between experimenters because we did not collect
308 repeated measurements on the same grackles when they were in hand (to reduce the amount of processing
309 time a bird experiences). Finally, our sample size might have been too small to detect an effect. However, the

310 effect size for the relationship between fat score and SMI was essentially zero (0.001), therefore it is unlikely
311 that a larger sample size would find a biologically informative relationship between these two proxies.

312 Although our first analysis of reproductive success, measured as the ability to produce fledglings (females)
313 or to hold a territory containing nests (males), found no correlation with fat score or SMI, when we used
314 logistic exposure models to determine the effect of female body condition on the probability of daily nest
315 survival, we found a negative relationship between SMI and the likelihood of daily nest survival. This result
316 was very surprising, but could be due to larger females actually carrying proportionally smaller energetic
317 reserves than their smaller female counterparts, as seen in red-winged blackbirds (Langston et al. 1990). In
318 some species, females with smaller body sizes are able to initiate breeding earlier because they can allocate
319 more resources to reproduction compared to larger individuals that have higher bodily energy demands and
320 therefore fewer excess energetic resources (Murphy 1986; Langston et al. 1990; Barbraud et al. 2000). This
321 indirectly affects reproductive success because nesting earlier increases the probability of nesting success and
322 multiple nesting attempts (Perrins 1970; Johnson and Peer 2001). Yet, we found no relationship between
323 probability of daily nest survival and day of the year, so this is unlikely to explain the negative relationship
324 between SMI and nest survival. Alternatively, it is possible that larger females are unable to build ideally
325 concealed nests in the most dense vegetation, or that larger females are more likely to disrupt nest stability.
326 The grackle nests are very high off the ground and usually fairly well concealed, so we could not determine
327 the causes of the nest failure and more research is needed that relates body condition to specific threats to
328 nesting success. In addition, the parameter estimate for the relationship between fat score and the daily
329 probability of nest survival indicates that females with some visible fat are more than twice as likely to have
330 a nest survive a given day. As the direction of this effect is opposite to the relationship between SMI and nest
331 survival, this further supports that these two proxies represent different traits and SMI is likely influenced
332 by muscle mass.

333 Measurements of energetic condition are important for understanding variation in life history characteristics
334 in studies from across the animal kingdom. However, the results of this study highlight the need to better
335 understand proxy measures of condition, not only in grackles but for birds in general. Most studies on avian
336 energetic condition only use one proxy for condition, but because energetic condition is difficult to measure
337 directly, it is important to compare multiple proxy variables to ensure each proxy is measuring the intended
338 trait (the jingle-jangle fallacy; Block 1995; Carter et al. 2013). Future research could add to this work
339 by incorporating additional methods to measure energetic condition, for example, blood hematocrit levels
340 (Dawson and Bortolotti 1997), or protein storage (Houston et al. 1995), or by studying additional traits
341 that could be affected by variation in energy stores, like dispersal (Ellers et al. 1998) or survival (Liao et
342 al. 2011). Furthermore, future research would benefit from using logistic exposure models to examine the
343 relationship between body condition and reproductive success, as it controls for the bias that arises when
344 early nest failures are not detected, which is not possible in logistic regression models, and it is more sensitive
345 to changes in the nest statuses of birds (Shaffer 2004).

346 METHODS

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

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

357 From March - August, we monitor the behavior of all color-marked grackles to determine their nesting status.
358 We follow females carrying nesting materials to find their nest. We determine whether the male territory

359 owner is color-marked as well. Then we check each nest approximately every day to determine the status
360 based on the female's behavior (building, incubation, feeding nestlings, feeding fledglings, failed).

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

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

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

377 **Open materials** Biometric measurement protocol: [https://gitlab.com/corinalogan/the-grackle-project/
378 blob/master/protocolBiometrics.pdf](https://gitlab.com/corinalogan/the-grackle-project/blob/master/protocolBiometrics.pdf)

379 Nest check protocol: [https://gitlab.com/corinalogan/the-grackle-project/blob/master/protocolNestCheck.
380 pdf](https://gitlab.com/corinalogan/the-grackle-project/blob/master/protocolNestCheck.pdf)

381 **Open data** All data necessary for the analyses are available at [https://knb.ecoinformatics.org/view/doi:
382 10.5063/F1NZ862D](https://knb.ecoinformatics.org/view/doi:10.5063/F1NZ862D) and at [github](#) (the provided code will load these files directly from [github](#)).

383 **Randomization and counterbalancing** There is no randomization or counterbalancing in this investi-
384 gation.

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

386 **Dependent Variables**

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

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

392 **P2: condition and reproductive success**

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

395 **Independent Variables**

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

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

405 **P2: condition and reproductive success**

406 1) Fat score

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

418 **ANALYSIS PLAN**

419 *UPDATE Oct 2020:*

- 420 1) *We realized that the sexual dimorphism of male and female body sizes necessitates separate analyses.*
421 *Therefore, we calculated SMI for males and females separately, ran separate models for each sex for the*
422 *repeatability analysis, P1 and P2.*
- 423 2) *Fat score data were distributed such that the majority of scores were 0, with some 1's and very few higher*
424 *numbers. This made it difficult to fit models using an ordinal regression. The function simulateResiduals,*
425 *which we used to check our data, does not work with data in the ordinal family. Consequently, we used*
426 *logistic regression where the dependent variable FatScore represents no fat (score = 0), or some fat (score =*
427 *1)*
- 428 3) *Despite the data checking which indicated our model was not overdispersed or zero inflated, we could not*
429 *get the fixed effects or random effect to converge using the Bayesian MCMCglmm. We found no improvement*
430 *in model fit by tweaking the priors or iterations/burnin/thin options. Therefore, we fit these models using*
431 *the function glmer, a frequentist framework.*
- 432 4) *The Season variable only includes 2 males in the breeding season category, thus we do not have a large*
433 *enough sample to produce reliable estimates. We removed the Season variable from the model for males.*
- 434 We will **exclude** data that was collected from the grackles when they were released from the aviaries to
435 avoid any confounds due to their time in the aviary (e.g., perhaps unlimited nutritious food in the aviaries

436 decreased their fat score). However, to validate that our measures of structural body size (tarsus length
437 or wing length) are precise and accurate, we will measure twice a subset of grackles brought into aviaries -
438 once when they are initially caught, and again up to 6 months later when we release them. We will then
439 calculate the repeatability of these multiple measures. All other data included in this study will come only
440 from wild-caught grackles (including the birds that were brought into the aviaries on their first capture).
441 When **missing data** occur, the existing data for that individual will be included in the analyses for which
442 their data exist. Analyses will be conducted in R (current version 4.0.2; R Core Team 2017).

443 **Ability to detect actual effects** To begin to understand what kinds of effect sizes we will be able to
444 detect given our sample size limitations, we used G*Power (v.3.1, Faul et al. 2007: @faul2009statistical) to
445 conduct power analyses based on confidence intervals. G*Power uses pre-set drop down menus and we chose
446 the options that were as close to our analysis methods as possible (listed in each analysis below). Note that
447 there were no explicit options for GLMMs, thus the power analyses are only an approximation of the kinds
448 of effect sizes we can detect. We realize that these power analyses are not fully aligned with our study design
449 and that these kinds of analyses are not appropriate for Bayesian statistics (e.g., our MCMCglmm below),
450 however we are unaware of better options at this time. Additionally, it is difficult to run power analyses
451 because it is unclear what kinds of effect sizes we should expect due to the lack of data on this species for
452 these particular research questions.

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

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

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

458 We use a Generalized Linear Mixed Model (GLMM; MCMCglmm function, MCMCglmm package; (Hadfield
459 2010)) with an ordinal distribution (for categorical variables in MCMCglmm) and probit link using 130,000
460 iterations with a thinning interval of 10, a burnin of 30,000, and minimal priors (V=1, nu=0) (Hadfield
461 2014). We will ensure the GLMM shows acceptable convergence (lag time autocorrelation values <0.01;
462 Hadfield 2010), and adjust parameters if necessary to meet this criterion. We will determine whether an
463 independent variable had an effect or not using the Estimate in the full model.

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

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

473 *Input:*

474 Effect size $f^2 = 0.15$

475 err prob = 0.05

476 Power (1- err prob) = 0.86

477 Number of predictors = 3

478 *Output:*

479 Noncentrality parameter = 13.3500000

480 Critical F = 2.7119214

481 Numerator df = 3
 482 Denominator df = 85
 483 Total sample size = 89
 484 Actual power = 0.8635760
 485 This means that, with a sample size of 89, we would have an 86% chance of detecting a medium effect
 486 (approximated at $f^2=0.15$ by Cohen 1988).
 487 *code not shown in pdf*
 488 *P2 analysis: condition and reproductive success*
 489 To model the effect of body condition on reproductive success, we will use two types of logistic mixed-
 490 effect models. Both types are supported in the literature, but are slightly different in the way in which
 491 the link function is specified. First, we will model reproductive success using a generalized linear mixed
 492 model framework with a logit link function (i.e. Milenkaya et al. 2015). We will also use a logistic exposure
 493 model that has a link function which accounts for the time interval between nest checks when estimating
 494 the probability of daily nest survival (Shaffer 2004; Bolker 2014). If fat score and the scaled mass index are
 495 positively correlated in P1, then we will use only fat score as the independent variable in this GLMM. If
 496 they are not positively correlated, we will include both as independent variables.
 497 Previous research found a non-linear relationship between reproductive success and body condition vari-
 498 ables (Milenkaya et al. 2015). To check whether this is occurring in our data, we will first plot our raw
 499 data to determine if we need to include a non-linear body condition independent variable into our model
 500 (i.e. FatScore²). Our dependent variable is binary, so to more clearly see the trends in the data, on the
 501 x-axis we will bin our condition scores into 5 categories based on standard deviations (sd) around the mean
 502 (low = < 2 sd, moderately low = -2 sd to -1 sd, moderate = -1 sd to +1 sd, moderately high = +1 sd to
 503 +2 sd, high = > 2 sd). Then on the y-axis we will use the proportion of individuals in each category that
 504 had successful nests. *UPDATE Oct 2020: Because most individuals fell within the medium category when*
 505 *we grouped data using 1 standard deviation around the mean, we switched to using half standard deviation*
 506 *increments around the mean.*
 507 A power analysis was conducted as above for P1 and the protocol reported here:
 508 *Input:*
 509 Effect size $f^2 = 0.15$
 510 err prob = 0.05
 511 Power (1- err prob) = 0.90
 512 Number of predictors = 2
 513 *Output:*
 514 Noncentrality parameter = 13.2000000
 515 Critical F = 3.1038387
 516 Numerator df = 2
 517 Denominator df = 85
 518 Total sample size = 88
 519 Actual power = 0.9020264
 520 This means that, with a sample size of 88, we would have a 90% chance of detecting a medium effect
 521 (approximated at $f^2=0.15$ by Cohen 1988).
 522 *code not shown in pdf*

523 **Do body condition variables vary by season?** *code not shown in pdf*

524 **Does body condition relate to reproductive success?** *code not shown in pdf*

525 **Does female body condition relate to the probability of daily nest survival?** Our measure of fe-
526 male nest success could be systematically biased against nests that failed early (Shaffer 2004). Consequently,
527 we also analyzed female reproductive success using a logistic exposure model. This type of model determines
528 the factors affecting daily nest survival probability.

529 *code not shown in pdf*

530 **ETHICS**

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

- 532 1) US Fish and Wildlife Service (scientific collecting permit number MB76700A-0,1,2)
533 2) US Geological Survey Bird Banding Laboratory (federal bird banding permit number 23872)
534 3) Arizona Game and Fish Department (scientific collecting license number SP594338 [2017], SP606267
535 [2018], and SP639866 [2019])
536 4) Institutional Animal Care and Use Committee at Arizona State University (protocol number 17-1594R)

537 **AUTHOR CONTRIBUTIONS**

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

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

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

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

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

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

544 **FUNDING**

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

547 **ACKNOWLEDGEMENTS**

548 We thank our research assistants for help with trapping the grackles and collecting the biometric and
549 nest/territory data: Aelin Mayer, Nancy Rodriguez, Brianna Thomas, Aldora Messinger, Elysia Mamola,
550 Michael Guillen, Rita Barakat, Adriana Bodenash, Olateju Ojekunle, August Sevchik, Justin Huynh, Amanda
551 Overholt, and Michael Pickett.

552 REFERENCES

- 553 Aubret F, Bonnet X, Shine R, Lourdais O. 2002. Fat is sexy for females but not males: The influence of
554 body reserves on reproduction in snakes (*vipera aspis*). *Hormones and Behavior*. 42(2):135–147.
- 555 Barbraud C, Lormée H, LeNevé A. 2000. Body size and determinants of laying date variation in the snow
556 petrel *pagodroma nivea*. *Journal of Avian Biology*. 31(3):295–302.
- 557 Barnett CA, Suzuki TN, Sakaluk SK, Thompson CF. 2015. Mass-based condition measures and their
558 relationship with fitness: In what condition is condition? *Journal of Zoology*. 296(1):1–5.
- 559 Barry KL, Wilder SM. 2013. Macronutrient intake affects reproduction of a predatory insect. *Oikos*.
560 122(7):1058–1064.
- 561 Bleeker M, Kingma SA, Szentirmai I, Székely T, Komdeur J. 2005. Body condition and clutch desertion in
562 penduline tit *remiz pendulinus*. *Behaviour*. 142:1465–1478.
- 563 Block J. 1995. A contrarian view of the five-factor approach to personality description. *Psychological
564 bulletin*. 117(2):187.
- 565 Bolker B. 2014. Logistic regression, accounting for differences in exposure. Version 0930 2014 RPubs.
- 566 Carter AJ, Feeney WE, Marshall HH, Cowlishaw G, Heinsohn R. 2013. Animal personality: What are
567 behavioural ecologists measuring? *Biological Reviews*. 88(2):465–475.
- 568 Champagnon J, Guillemain M, Elmberg J, Massez G, Cavallo F, Gauthier-Clerc M. 2012. Low survival after
569 release into the wild: Assessing ‘the burden of captivity’ on mallard physiology and behaviour. *European
570 Journal of Wildlife Research*. 58(1):255–267.
- 571 Cohen J. 1988. Statistical power analysis for the behavioral sciences 2nd edn.
- 572 Cornelius Ruhs E, Vézina F, Karasov WH. 2019. Physiological and immune responses of free-living temperate
573 birds provided a gradient of food supplementation. *Physiological and Biochemical Zoology*. 92(1):106–114.
- 574 Costa G. 2012. Behavioural adaptations of desert animals. Springer Science & Business Media.
- 575 Dawson RD, Bortolotti GR. 1997. Are avian hematocrits indicative of condition? American kestrels as a
576 model. *The Journal of wildlife management*.:1297–1306.
- 577 Delciellos AC, Barros C dos S de, Prevedello JA, Ferreira MS, Cerqueira R, Vieira MV. 2018. Habitat
578 fragmentation affects individual condition: Evidence from small mammals of the brazilian atlantic forest.
579 *Journal of Mammalogy*. 99(4):936–945.
- 580 Ellers J, Van Alphen JJ, Sevenster JG. 1998. A field study of size–fitness relationships in the parasitoid
581 *asobara tabida*. *Journal of Animal Ecology*. 67(2):318–324.
- 582 English MD, Robertson GJ, Peck LE, Pirie-Hay D, Roul S, Mallory ML. 2018. Body condition of american
583 black ducks (*anas rubripes*) wintering in atlantic canada using carcass composition and a scaled mass index.
584 *Canadian Journal of Zoology*. 96(10):1137–1144.
- 585 Erciyas K, Gürsoy A, Özsemir A, Barış Y. 2010. Body mass and fat score changes in recaptured birds during
586 the autumn migration at the cernek ringing station in turkey. *The Ring*. 32(1-2):3–15.
- 587 Faul F, Erdfelder E, Buchner A, Lang A-G. 2009. Statistical power analyses using g* power 3.1: Tests for
588 correlation and regression analyses. *Behavior research methods*. 41(4):1149–1160.
- 589 Faul F, Erdfelder E, Lang A-G, Buchner A. 2007. G* power 3: A flexible statistical power analysis program
590 for the social, behavioral, and biomedical sciences. *Behavior research methods*. 39(2):175–191.
- 591 Fleskes JP, Ramey AM, Reeves AB, Yee JL. 2017. Body mass, wing length, and condition of wintering ducks
592 relative to hematozoa infection. *Journal of Fish and Wildlife Management*. 8(1):89–100.
- 593 Haas CA. 1998. Effects of prior nesting success on site fidelity and breeding dispersal: An experimental
594 approach. *The Auk*. 115(4):929–936.

- 595 Hadfield J. 2010. MCMC methods for multi-response generalized linear mixed models: The mcmcglmm r
596 package. *Journal of Statistical Software*. 33(2):1–22.
- 597 Hadfield J. 2014. MCMCglmm course notes. <http://cran.r-project.org/web/packages/MCMCglmm/vignettes/CourseNotes.pdf>.
- 599 Hartig F. 2019. DHARMA: Residual diagnostics for hierarchical (multi-level / mixed) regression models.
600 <http://florianhartig.github.io/DHARMA/>.
- 601 Heidinger IMM, Hein S, Bonte D. 2010. Patch connectivity and sand dynamics affect dispersal-related
602 morphology of the blue-winged grasshopper *oedipoda caerulescens* in coastal grey dunes. *Insect Conservation
603 and Diversity*. 3(3):205–212.
- 604 Henderson L, Evans N, Heidinger B, Herborn K, Arnold K. 2017. Do glucocorticoids predict fitness? Linking
605 environmental conditions, corticosterone and reproductive success in the blue tit, *cyanistes caeruleus*. *Royal
606 Society open science*. 4(10):170875.
- 607 Houston DC, Donnan D, Jones P, Hamilton I, Osborne D. 1995. Changes in the muscle condition of female
608 zebra finches *poephila guttata* during egg laying and the role of protein storage in bird skeletal muscle. *Ibis*.
609 137(3):322–328.
- 610 Johnson K, Peer BD. 2001. Great-tailed grackle: *Quiscalus mexicanus*. *Birds of North America, Incorporated*.
- 611
- 612 Kaiser A. 1993. A new multi-category classification of subcutaneous fat deposits of songbirds (una nueva clasificación, con multi-categorías, para los depósitos de grasa en aves canoras). *Journal of Field Ornithology*.:246–
613 255.
- 614
- 615 Kelly CD, Tawes BR, Worthington AM. 2014. Evaluating indices of body condition in two cricket species.
616 *Ecology and Evolution*. 4(23):4476–4487.
- 617 Kraft F-LO, Driscoll SC, Buchanan KL, Crino OL. 2019. Developmental stress reduces body condition
618 across avian life-history stages: A comparison of quantitative magnetic resonance data and condition indices.
619 *General and comparative endocrinology*. 272:33–41.
- 620 Labocha MK, Hayes JP. 2012. Morphometric indices of body condition in birds: A review. *Journal of
621 Ornithology*. 153(1):1–22.
- 622 Labocha MK, Schutz H, Hayes JP. 2014. Which body condition index is best? *Oikos*. 123(1):111–119.
- 623 Langston NE, Freeman S, Rohwer S, Gori D. 1990. The evolution of female body size in red-winged
624 blackbirds: The effects of timing of breeding, social competition, and reproductive energetics. *Evolution*.
625 44(7):1764–1779.
- 626 Liao C-Y, Rikke BA, Johnson TE, Gelfond JA, Diaz V, Nelson JF. 2011. Fat maintenance is a predictor of
627 the murine lifespan response to dietary restriction. *Aging cell*. 10(4):629–639.
- 628 Maceda-Veiga A, Green AJ, De Sostoa A. 2014. Scaled body-mass index shows how habitat quality influences
629 the condition of four fish taxa in north-eastern spain and provides a novel indicator of ecosystem health.
630 *Freshwater biology*. 59(6):1145–1160.
- 631 Mayfield H. 1961. Nesting success calculated from exposure. *The Wilson Bulletin*.:255–261.
- 632 McNamara JM, Barta Z, Houston AI, Race P. 2005. A theoretical investigation of the effect of predators
633 on foraging behaviour and energy reserves. *Proceedings of the Royal Society B: Biological Sciences*.
634 272(1566):929–934.
- 635 Merilä J, Svensson E. 1997. Are fat reserves in migratory birds affected by condition in early life? *Journal
636 of Avian Biology*.:279–286.
- 637 Milenkaya O, Catlin DH, Legge S, Walters JR. 2015. Body condition indices predict reproductive success
638 but not survival in a sedentary, tropical bird. *PLoS One*. 10(8):e0136582.

- 639 Murphy MT. 1986. Body size and condition, timing of breeding, and aspects of egg production in eastern
640 kingbirds. *The Auk.* 103(3):465–476.
- 641 Musacchia X. 1953. A study of the lipids in arctic migratory birds. *The Condor.* 55(6):305–312.
- 642 Nakagawa S, Schielzeth H. 2010. Repeatability for gaussian and non-gaussian data: A practical guide for
643 biologists. *Biological Reviews.* 85(4):935–956.
- 644 Peig J, Green AJ. 2009. New perspectives for estimating body condition from mass/length data: The scaled
645 mass index as an alternative method. *Oikos.* 118(12):1883–1891.
- 646 Perrins C. 1970. The timing of birds ‘breeding seasons. *Ibis.* 112(2):242–255.
- 647 R Core Team. 2017. R: A language and environment for statistical computing. Vienna, Austria: R
648 Foundation for Statistical Computing. <https://www.R-project.org>.
- 649 Shaffer TL. 2004. A unified approach to analyzing nest success. *The Auk.* 121(2):526–540.
- 650 Stevenson R, Woods Jr WA. 2006. Condition indices for conservation: New uses for evolving tools. *Integrative
651 and comparative biology.* 46(6):1169–1190.
- 652 Stoffel MA, Nakagawa S, Schielzeth H. 2017. RptR: Repeatability estimation and variance decomposition
653 by generalized linear mixed-effects models. *Methods in Ecology and Evolution.* 8(11):1639–1644.
- 654 Warnock N, Bishop MA. 1998. Spring stopover ecology of migrant western sandpipers. *The Condor.*
655 100(3):456–467.
- 656 Wehtje W. 2003. The range expansion of the great-tailed grackle (*Quiscalus mexicanus gmelin*) in north
657 America since 1880. *Journal of Biogeography.* 30(10):1593–1607.
- 658 Wilder SM, Raubenheimer D, Simpson SJ. 2016. Moving beyond body condition indices as an estimate of
659 fitness in ecological and evolutionary studies. *Functional Ecology.* 30(1):108–115.
- 660 Zhang Y, Eyster K, Liu J-S, Swanson DL. 2015. Cross-training in birds: Cold and exercise training produce
661 similar changes in maximal metabolic output, muscle masses and myostatin expression in house sparrows
662 (*Passer domesticus*). *Journal of Experimental Biology.* 218(14):2190–2200.