Ecological Applications, 0(0), 2016, pp. 1–14

Conservation implications of a lack of relationship between baseline glucocorticoids and fitness in a wild passerine

Christine L. Madliger[1,3] and Oliver P. Love[1,2]

[1]*Department of Biological Sciences, University of Windsor, 401 Sunset Avenue, Windsor, Ontario N9B 3P4 Canada*
[2]*Great Lakes Institute for Environmental Research, University of Windsor, 401 Sunset Avenue, Windsor, Ontario N9B 3P4 Canada*

Abstract. The application of physiological measures to conservation monitoring has been gaining momentum and, while a suite of physiological traits are available to ascertain disturbance and condition in wildlife populations, glucocorticoids (i.e., GCs; cortisol and corticosterone) are the most heavily employed. The interpretation of GC levels as sensitive indicators of population change necessitates that GCs and metrics of population persistence are linked. However, the relationship between GCs and fitness may be highly context-dependent, changing direction, or significance, depending on the GC measure, fitness metric, life history stage, or other intrinsic and extrinsic contexts considered. We examined the relationship between baseline plasma corticosterone (CORT) levels measured at two periods of the breeding season and three metrics of fitness (offspring quality, reproductive output, and adult survival) in female Tree Swallows (*Tachycineta bicolor*). Specifically, we investigated whether (1) a relationship between baseline CORT metrics and fitness exists in our population, (2) whether the inclusion of energetic contexts, such as food availability, reproductive investment, or body mass, could alter or improve the strength of the relationship between CORT and fitness, and (3) whether energetic contexts could better predict fitness compared to CORT metrics. Importantly, we investigated these relationships in both natural conditions and under an experimental manipulation of foraging profitability (feather clipping) to determine the influence of an environmental constraint on GC–fitness relationships. We found a lack of relationship between baseline CORT and both short- and long-term metrics of fitness in control and clipped birds. In contrast, loss in body mass over reproduction positively predicted reproductive output (number of chicks leaving the nest) in control birds; however, the relationship was characterized by a low R^2 (5%), limiting the predictive capacity, and therefore the application potential, of such a measure in a conservation setting. Our results stress the importance of ground-truthing GC–fitness relationships and indicate that baseline GCs will likely not be easily employed as conservation biomarkers across some species and life history stages. Given the accumulating evidence of temporally dynamic, inconsistent, and context-dependent GC–fitness relationships, placing effort towards directly measuring fitness traits, rather than plasma GC levels, will likely be more worthwhile for many conservation endeavours.

Key words: baseline; biomarker; conservation physiology; CORT-fitness; corticosterone; fitness; flexibility; glucocorticoid; reproductive success; survival; Tachycineta bicolor.

Introduction

Conservation biologists and managers are often tasked with urgently and accurately determining how wildlife populations respond to changing environmental conditions (Angelier and Wingfield 2013, Wingfield 2013). Traditionally, demographics have been employed to monitor changes in population persistence over time; however, these approaches provide little guidance on possible mitigation strategies and require large longitudinal datasets to ensure that changes in population numbers represent significant trends (Carey 2005, Wikelski and Cooke 2006, Ellis et al. 2011). Consequently, the measurement of physiological metrics, such as hormones, metabolites, immune factors, etc., forms the foundation of the growing field of conservation physiology, which seeks to provide proactive insight into population health and condition (Wikelski and Cooke 2006, Cooke et al. 2013). One of the essential requirements of employing any physiological metric as this type of biomarker is the presence of a predictable relationship with fitness (Busch and Hayward 2009, Cooke and O'Connor 2010, Madliger and Love 2015). Without such a relationship, variation in physiology over time cannot be interpreted as a sensitive indicator of future population change; predictive capacity necessitates a linkage between physiology and the metrics that drive population persistence (Fig. 1).

Although the physiological measures available to conservation managers (see Cooke et al. 2013 for an overview) and used successfully for conservation goals (Madliger et al. 2016) are diverse, the field is currently dominated by measures of stress physiology (i.e., glucocorticoids, GCs;

Manuscript received 8 January 2016; revised 9 June 2016; accepted 13 July 2016. Corresponding Editor: D. Brunton.

[3]E-mail: madlige@uwindsor.ca

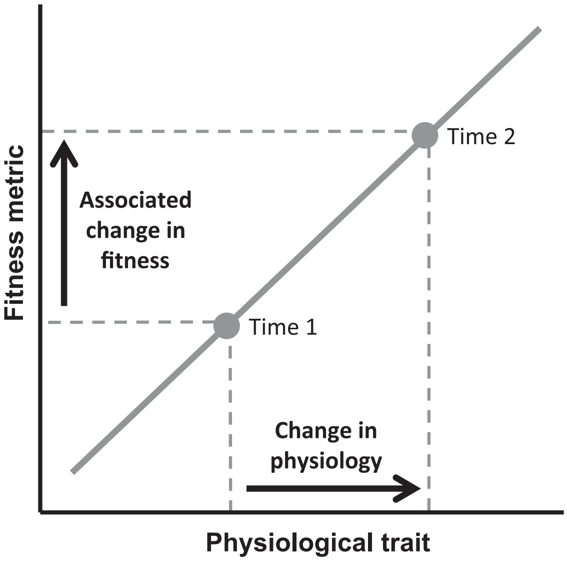

Fig. 1. Schematic displaying a hypothetical relationship between a physiological variable and fitness. Such a relationship is necessary to interpreting changes in physiology as predictive indicators of population health or persistence. It should be noted that negative, positive, and non-linear relationships would all be interpretable.

cortisol and corticosterone, CORT; Lennox and Cooke 2014). At baseline levels, GCs regulate feeding, promote regular activity patterns, and maintain energy metabolism (Landys et al. 2006). Unexpected acute perturbations result in a transient increase in GCs that mobilize energy reserves, temporarily suspend reproduction and other non-essential activities, promote subsequent foraging, and regulate immune function (Wingfield et al. 1998, Wingfield and Kitaysky 2002, Wingfield and Sapolsky 2003, Landys et al. 2006). If an organism is unable to overcome a stressful event (i.e., restore homeostasis), GCs can remain elevated for a prolonged period leading to allostatic overload with negative consequences for reproductive activities, foraging, immunity, vigilence, and/or survival (Korte et al. 2005, McEwen and Wingfield 2010). Overall, GC levels are expected to parallel energetic demand (i.e., allostatic load; Wingfield 2005), with recent evidence suggesting a causal link between baseline GCs and reproductive investment, likely mediated through behavior or metabolism (Cook et al. 2011, Ouyang et al. 2011, 2013*a*, Love et al. 2014).

Much of the application of GCs in the context of conservation has been based on the assumption of a negative relationship between baseline levels and fitness (Bonier et al. 2009*a*, Busch and Hayward 2009), with higher baseline levels correlating with lower condition, reproductive success, and/or survival probability (Bonier et al. 2009*a*). Although this relationship, formally known as the "CORT–fitness hypothesis" (Bonier et al. 2009*a*), has been investigated relatively extensively by ecological and evolutionary physiologists, results have been extremely mixed (Bonier et al. 2009*a*, Angelier et al. 2010, Breuner 2011, Crespi et al. 2013, Escribano-Avila et al. 2013, Sopinka et al. 2015). For example, Bonier et al. (2009*b*) found that the relationship between baseline GCs and reproductive success in female Tree Swallows (*Tachycineta bicolor*) switched from negative during the early breeding (incubation) stage to positive during the offspring provisioning stage. This discrepancy is likely a reflection of the role of GCs in promoting foraging; when the fitness value of a reproductive attempt is high (e.g., late in a reproductive attempt), increases in GCs will promote allocation of resources towards the breeding attempt leading to a positive relationship between GCs and fitness (Bonier et al. 2009*a*, *b*, 2011, Breuner 2011). In contrast, when the fitness value of a reproductive attempt is low (e.g., early in reproduction), increases in GCs are more likely to interfere with successful reproduction (Bonier et al. 2009*a*, *b*, 2011, Breuner 2011). While some support for this idea exists (Love et al. 2004, Bókony et al. 2009, Bonier et al. 2011, Escribano-Avila et al. 2013), it is not yet clear how well it explains discrepancies across the entire range of GC–fitness investigations.

It is also possible that context-dependent GC–fitness relationships may be causing much of the ambiguity in results (Breuner 2011, Crespi et al. 2013, Jaatinen et al. 2013, Madliger and Love 2014). For example, Ebensperger et al. (2013) could not detect a relationship between fecal GCs and survival in degus (*Octodon degus*), but levels did predict whether females would produce a second litter, illustrating the importance of investigating multiple fitness metrics simultaneously. Other contexts, such as sex and environmental quality, have also proven to be important. For instance, Angelier et al. (2010) found a negative relationship between baseline GCs and 5-yr reproductive output in male, but not female, Black-browed Albatross (*Thalassarche melanophris*). Similarly, D'Alba et al. (2011) found that a negative GC–fitness relationship was only evident in Common Eiders (*Somateria mollissima*) nesting in poorer, exposed sites as compared to sheltered sites. Finally, the modulation of GC levels across demanding time periods may better predict fitness outcomes than static (single-point) measures (Love and Williams 2008, Williams 2008, Ouyang et al. 2011, Love et al. 2014). For example, a manipulation of brood size in Tree Swallows indicated that the change in baseline plasma GCs over reproduction, but not static measures, could positively predict maternal foraging rates and subsequent fledging success (Bonier et al. 2011). In other words, initial physiological state may influence the subsequent response to environmental constraint or reproductive demand (Ouyang et al. 2011, 2013*b*, Arlettaz et al. 2014, Love et al. 2014). This possibility has important consequences for management applications as it would necessitate the collection of two or more measurements of GCs across time to infer population health or persistence. Overall, context-dependent GC–fitness relationships have important implications for employing GC levels in conservation, as levels will need to be interpreted differently depending on intrinsic and extrinsic factors, and in relation to different fitness metrics (e.g., current reproductive success vs. survival probability; Madliger and Love 2014).

We used an explicit, context-dependent approach to investigate the relationship between baseline (plasma) GC levels and fitness in female Tree Swallows over three years. We measured baseline CORT at two different stages of reproduction (late incubation and mid-offspring provisioning) and assessed the relationship of each measure, as well as the change in CORT over the season, with three distinct metrics of fitness: (1) reproductive output, (2) offspring quality, and (3) survival probability. Moreover, we determined whether the inclusion of additional contexts such as food availability, reproductive investment, or body mass could alter the CORT–fitness relationship, or whether any of these contexts were able to better predict fitness than CORT measures. Finally, we included an experimental manipulation of foraging profitability (feather clipping) to examine whether the CORT–fitness relationship changes when individuals are pushed outside of expected environmental conditions.

We predicted that the CORT–fitness relationship would change from negative during incubation to positive during nestling provisioning in the control group (i.e., under natural conditions), in line with the increase in brood value over the reproductive period and the role of GCs in promoting allocation of resources to reproduction (Bonier et al. 2011). In contrast, we predicted that the CORT–fitness relationship would remain negative at the nestling provisioning stage in the clipped group as we expected the individuals with the highest CORT levels would be those most negatively impacted by the decline in foraging profitability and also the least able to successfully raise offspring or survive to the following year. We also predicted that the change in baseline CORT over the reproductive season would be positively related to fitness in control birds. By examining contexts such as reproductive stage, GC metric, fitness metric, energetic contexts, and environmental quality, we can better assess when, and how, GC levels may predict population-level demographic consequences. Such an approach is essential to determining the applicability of GC levels as conservation biomarkers and is currently considered one of the highest priority questions in the field of conservation physiology (Cooke 2014).

METHODS

Study species

Tree Swallows are small, migratory passerines that readily breed in artificial boxes. The species represents a model organism that has been heavily studied in the context of life-history tradeoffs, ecological requirements, toxicology, and immune responses (Jones 2003). Tree Swallows are aerial insectivores, a guild of birds which feed on flying insects, that has shown precipitous declines in North America (Nebel et al. 2010). Although they are an abundant, widespread member of this group, they have nonetheless declined by 2.5% per annum over the past 20 years (McCracken 2008), with some populations declining more rapidly (Paquette et al. 2014). Our study population is located in an area of southern Ontario, Canada, that is characterized by agricultural expansion and loss of wetland habitat similar to the alterations that this species would be exposed to across a large proportion of its range. As such, it is an ideal species to investigate relationships between physiology and fitness and to draw conclusions about how changes in environmental quality on the breeding grounds may lead to demographic consequences for aerial insectivores.

Nest monitoring and sampling protocol

We monitored breeding attempts of pairs of Tree Swallows in a nest box colony located in southern Ontario, Canada, from 2010 to 2014. In total, our study area consists of 175 nest boxes located across two conservation sites 4 km apart: Taquanyah Conservation Area (42°57′ N, 79°54′ W) and Ruthven Park National Historic Site (42°58′ N, 79°52′ W). Boxes are located adjacent to and within a variety of landscapes including active agricultural fields, roadways, fallow fields, cattle pastures, and riparian areas along the Grand River. We checked boxes once daily and recorded date of the first egg laid, clutch size, egg masses, hatching success, and the number of offspring that successfully left the nest (fledging success). We also recorded the mass of each nestling at 6 and 12 d of age. The summed nestling mass measurements taken at 12 d of age were used as our estimate of brood mass. At this point in time, nestlings can have masses equal or greater to adults (Quinney et al. 1986, McCarty 2001). Overall, nestling growth can influence post-fledging survival, with chicks with delayed or interrupted growth showing diminished long-term survival (McCarty 2001).

We focused on females in our study because they are the sole incubators and the most accessible sex throughout reproduction. Females were captured at their nest box at two time periods: (1) late incubation (10 d after clutch completion) and (2) mid-nestling provisioning (12 d post-hatch). Females were blood-sampled through puncture of the brachial (wing) vein to obtain <150 μL of blood (i.e., <10% of total blood volume). Blood samples were obtained between 08:00 and 12:00 to control for diel changes in baseline CORT values and within 2 min of trapping a bird in the nest box to ensure acquisition of baseline samples (Romero and Reed 2005). We also recorded mass, wing length, and age, and visually scored fat and muscle condition. Female Tree Swallows in their first year are characteristically brown in color, while females aged 2 yr or older are iridescent blue-green (Hussell 1983). We also gave unbanded birds a federal numbered band (Canadian Wildlife Service Permit: 10,808). Finally, we determined return rates for each female to serve as a proxy for survival probability. Tree Swallows live an average of 2.7 yr and to a maximum of 12 yr (Winkler et al. 2011). Ninety-five percent of birds that fledge at least one offspring will return to the same breeding site in the subsequent year (many to the exact same nest box), and even after complete nest failure, females are still 72% likely to return to the same breeding site (Winkler et al. 2004). As a result, we considered return rates to be valid proxies for female survival in this species due to their extremely high level of philopatry (Winkler et al. 2004).

We also monitored food availability throughout the reproductive period using four-sided, commercially available malaise traps. In total, we used five traps across our study sites. Our sampling protocol has been described and supported for use in Tree Swallows in detail elsewhere (Madliger et al. 2015). Briefly, we calculated the average insect (dry) biomass (as per Hussell and Quinney 1987) that was available over the most demanding stage of the nestling provisioning period (day 5–10; McCarty 2001) for each individual as our measure of food availability for subsequent analyses.

Manipulation of foraging profitability

In 2011, we used a feather clipping manipulation to decrease foraging profitability (Winkler and Allen 1995) and therefore simulate an unexpected decline in habitat quality in the context of food availability (Madliger et al. 2015). We clipped four flight feathers at the base of the feather on each wing (i.e., every other primary flight feather) of 33 females (Winkler and Allen 1995, Ardia and Clotfelter 2007; control, $n = 38$). This level of feather clipping causes a handicap in this species by increasing the energetic cost of flight (Winkler and Allen 1995, Ardia and Clotfelter 2007), leading to decreased foraging rate (Winkler and Allen 1995, Patterson et al. 2011, Madliger et al. 2015) and lower body condition compared to control birds (Ardia and Clotfelter 2007, Patterson et al. 2011). In addition, this manipulation increased baseline GCs over reproduction and led to the clipped group having higher average levels of GCs compared to control birds at the nestling provisioning stage in our population (Madliger et al. 2015). We performed this manipulation at the late incubation stage, immediately following acquisition of the first blood sample. All other birds were handled in the same way, but no flight feathers were clipped. Only females aged 2 yr or older were included in the clipping manipulation due to a considerably smaller number of 1-yr-old females across our sites. Females faced this handicap for two weeks before the nestling provisioning blood sample was obtained and remained feather clipped until the natural moult that follows breeding (prior to migration to wintering grounds; Stutchbury and Rohwer 1990). We recorded no instances of nest abandonment as a result of the manipulation. All manipulation and monitoring protocols were approved by the University of Windsor's Animal Care Committee (AUPP #10-10) and the Canadian Wildlife Service (Permit CA 0266).

Corticosterone assay

We stored blood samples on ice for up to 5 h and then centrifuged to separate plasma. Plasma was stored at −80°C until assay. We measured non-extracted levels of baseline CORT (the primary GC in avian species) in plasma using a Corticosterone Enzyme-linked Immunoabsorbent Assay (EIA; Assay Designs, Ann Arbor, Michigan, USA, catalog #901-097) that has been validated in our lab (Love and Williams 2008) and used extensively in passerines (e.g., Breuner et al. 2006, Wada et al. 2007, Zhang et al. 2011, Rivers et al. 2012). Samples were run in triplicate at a total volume of 100 μL with 1:40 dilution and 1.5% steroid displacement buffer (SDB). We calculated the detection limit of the assay as 0.74 ng/mL (as per the manufacturer's method). Of 442 total plasma samples analyzed, 24 fell below that value and were assigned the value of the detection limit. Intra-assay variation was 7.7%, 8.0%, and 10.3% in 2010, 2011, and 2012, respectively. Inter-assay variation was 6.7%, 13.3%, and 6.0% in 2010, 2011, and 2012, respectively (calculated from control samples of spiked laying hen plasma; one control run in triplicate per assay plate). Raw CORT values were: incubation, 0.74–9.47 ng/mL; nestling provisioning, 0.74–12.47 ng/mL.

Statistical analyses

We analyzed whether measures of baseline CORT could predict fitness metrics in clipped and control birds separately to allow for conclusions about how the relationship may be different when individuals are faced with an unexpected change in environmental quality (i.e., feather clipping). Specifically, we used three metrics representing different components of fitness as dependent variables, each analyzed separately. First, we calculated offspring quality as the residual of brood mass (calculated as the total mass, g, of all chicks in the nest at 12 d of age) on number of chicks. Residual brood mass provides a measure of offspring quality that is uncorrelated with the number of offspring (output), where individuals with heavier than average offspring for a given brood size can be discerned from those with smaller than average offspring. Second, we used the total number of offspring that successfully fledged from a nest as a measure reproductive output. Third, adult survival probability was recorded as 0 for birds that were not subsequently recorded as returning to the breeding site and as 1 for birds that returned to the breeding site in a subsequent year.

We used three metrics of baseline CORT as independent variables: (1) late incubation CORT, (2) mid-nestling provisioning CORT, and (3) percentage of change in CORT over the reproductive season, calculated as the absolute difference in CORT levels divided by the incubation CORT level. We used the percent change in CORT, rather than the absolute difference, to better take into account the overall degree of change. We also included the percent change in body mass over the reproductive season as an energetic context. For within-season fitness metrics (reproductive output and offspring quality), we also included insect biomass over each female's peak offspring provisioning period as a measure of food availability. For the survival analyses, we included number of offspring fledged as an additional independent variable to represent previous reproductive investment. All analyses included site (random) and lay date (fixed) as covariates. Since Tree Swallow reproductive performance declines over the season (Stutchbury and Robertson 1988), including lay date allowed us to control for the potential influences of timing on fitness outcomes. Lay date was standardized to represent a relative lay date within each year (by subtracting the average) to make the timing of reproduction comparable across years. In addition, in control analyses (n = 122), which included three years of data, we included year and individual ID as random effects. In contrast, analyses in the clipped group included only one year of data (n = 33). Analyzing control birds within only a single year (i.e., investigating whether treatment specifically alters the CORT–fitness relationship) provided similar results to the multi-year analysis across all three CORT metrics. As such, we subsequently only provide results for the full analyses. We checked for collinearity of independent variables by calculating variance inflation factors (VIFs). Due to a high correlation between the change in CORT and the single time-point CORT measures ($R > 0.70$), we did not construct any models with both variables included simultaneously. All other VIFs were below 1.40.

We used AIC (Akaike's information criterion) to perform model selection and determine which physiological or other contexts were best able to predict short- and long-term fitness metrics. More specifically, we used AIC$_c$ (AIC corrected for small sample size) values to calculate ΔAIC$_c$, Akaike's weights (ω), and cumulative weights, allowing us to determine the best-supported models from our candidate sets. Each fitness metric was analyzed separately in each treatment group (i.e., we performed a total of six analyses). Each candidate set included 20 models; we used an all-subsets approach, except (as outlined previously) the percent change in CORT was never included in a model with either incubation CORT or nestling provisioning CORT due to high collinearity. We used cumulative weights to determine 95% confidence sets of models in each analysis. A confidence set represents a list of models in which we can be 95% certain the best model from our original candidate set is included (Symonds and Moussalli 2011). Where model uncertainty was evident (i.e., when no single model could be identified as the top model), we performed multi-model inference using the 95% confidence set (Burnham and Anderson 2002). This allowed us to obtain model-averaged parameter estimates (β), unconditional standard errors, and 95% confidence intervals (Burnham and Anderson 2002, Johnson and Omland 2004). Model-averaged β-values and unconditional standard errors were calculated by weighting them by the Akaike weights of the models included in the confidence set (Burnham and Anderson 2002).

Analyses with offspring quality as the dependent variable were completed using linear mixed effects models (LMMs); residuals were normally distributed in all cases. Analyses with reproductive output (i.e., count data) and survival probability (i.e., binary data) as dependent variables were completed using generalized linear mixed effects models (GLMMs). In the case of survival probability, we used a binomial distribution and a logit-link function. For reproductive output, we used a Poisson distribution and a log-link function. All continuous covariates were mean-centered (Bolker et al. 2009). Analyses were completed in R (R Development Team 2015) using the lme4 package with the lmer and glmer function (Bates et al. 2015). Marginal and conditional R^2 of LMMs and GLMMs were calculated with the package MuMIN with the function rsquared.glmm (Barton 2015).

RESULTS

In the analyses investigating which CORT metrics and other energetic contexts (food availability, reproductive investment, mass loss) predicted fitness in control birds, the null model was among the best supported models in both the offspring quality and the survival analyses (Table 1). There was considerable model uncertainty with 11 and 15 models included in the 95% confidence sets of the offspring quality and survival analyses, respectively. For all CORT and energetic variables, parameter estimates were very low and 95% confidence intervals crossed zero, indicating a lack of association of the variables investigated with both offspring quality and adult survival (Table 2). In contrast, the model composed of only percent change in body mass represented the best supported model in the analysis of reproductive output (Table 1). More specifically, greater losses of body mass were associated with a greater number of successfully fledged offspring (Fig. 2); however, the marginal R^2 of the relationship was only 0.05 indicating a large degree of unexplained variance. The 95% confidence set included 14 models, and all other variables investigated showed no association with reproductive output (Table 2).

In contrast to control birds, the top model for reproductive output in the clipped group was the null model (Table 3). The 95% confidence set included 14 models, and confidence intervals of all variables crossed zero, indicating poor precision of parameter estimation and a lack of association between energetic and CORT metrics and the number of offspring produced (Table 4). In terms of offspring quality in the clipped group, the models containing single or multiple CORT metrics represented the best supported models (Table 3). However, there was considerable model uncertainty and all variables included in the 95% confidence set had confidence intervals that crossed zero (Table 4). Similarly, in the survival analysis, the best supported model contained only the percent change in CORT and represented the only model in the

TABLE 1. Summary of confidence (95%) set of models predicting fitness metrics in female tree swallows in control birds (n = 122).

Model variables	AIC$_c$	ΔAIC$_c$	ω	Cum. ω	R^2
(a) Offspring fledged (output)					
massΔ	476.63	0.00	0.28	0.28	0.05
massΔ, CORT1	478.73	2.10	0.10	0.37	0.05
CORTΔ, massΔ	478.77	2.14	0.10	0.47	0.05
massΔ, CORT2	478.86	2.23	0.09	0.56	0.05
massΔ, food	478.88	2.25	0.09	0.65	0.05
Null (lay date, ID, site, year)	479.19	2.56	0.08	0.73	0.01
massΔ, CORT1, CORT2	480.96	4.33	0.03	0.76	0.05
massΔ, food, CORT1	481.02	4.39	0.03	0.79	0.05
CORTΔ, massΔ, food	481.05	4.43	0.03	0.82	0.05
massΔ, food, CORT2	481.15	4.52	0.03	0.85	0.05
CORTΔ	481.34	4.72	0.03	0.88	0.01
CORT2	481.36	4.74	0.03	0.90	0.01
food	481.39	4.76	0.03	0.93	0.01
CORT1	481.40	4.77	0.03	0.95	0.01
(b) Offspring quality					
Null (lay date, ID, site, year)	874.82	0.00	0.30	0.30	0.25
CORTΔ	876.13	1.31	0.16	0.46	0.25
CORT2	876.53	1.71	0.13	0.58	0.22
CORT1	876.89	2.07	0.11	0.69	0.24
massΔ	877.62	2.80	0.07	0.76	0.28
CORT1, CORT2	878.57	3.75	0.05	0.81	0.22
CORTΔ, massΔ	879.01	4.19	0.04	0.85	0.28
massΔ, CORT2	879.40	4.58	0.03	0.88	0.25
food	879.50	4.68	0.03	0.91	0.31
massΔ, CORT1	879.71	4.89	0.03	0.93	0.27
CORTΔ, food	880.93	6.11	0.01	0.95	0.31
(c) Survival					
Null (lay date, ID, site, year)	174.54	0.00	0.22	0.22	0.01
massΔ	175.83	1.29	0.12	0.34	0.02
fledged	176.53	1.99	0.08	0.42	0.01
CORT1	176.56	2.02	0.08	0.50	0.01
CORTΔ	176.62	2.08	0.08	0.58	0.01
CORT2	176.68	2.14	0.08	0.66	0.01
CORTΔ, massΔ	177.95	3.41	0.04	0.70	0.02
massΔ, CORT1	177.95	3.42	0.04	0.74	0.02
massΔ, fledged	177.98	3.44	0.04	0.78	0.02
massΔ, CORT2	178.00	3.46	0.04	0.82	0.02
fledged, CORT1	178.59	4.05	0.03	0.85	0.01
CORTΔ, fledged	178.64	4.10	0.03	0.88	0.01
fledged, CORT2	178.70	4.17	0.03	0.90	0.01
CORT1, CORT2	178.72	4.19	0.03	0.93	0.01
CORTΔ, massΔ, fledged	180.14	5.61	0.01	0.95	0.02

Notes: The summary provides AIC values corrected for small sample size (AIC$_c$), ΔAIC$_c$, Akaike weight (ω), cumulative Akaike weights (cum. ω), and conditional R^2 for each model. Abbreviations are as follows: massΔ is percent loss in body mass from incubation to nestling provisioning, CORT1 is baseline CORT at the incubation stage, CORT2 is baseline CORT at the nestling provisioning stage, CORTΔ is percent change in CORT from incubation to nestling provisioning, food is average insect biomass (food availability) during peak nestling provisioning, and fledged is the number of offspring fledged (see *Methods* for details). All models included lay date as a fixed effect, and female ID, site, and year as random effects.

candidate set ranking higher than the null (Table 3). Again, there was model uncertainty with 13 models constituting the 95% confidence set. Model-averaged parameter estimates and confidence intervals indicate that all variables investigated show a lack of association with survival probability (Table 4).

DISCUSSION

We used an integrative and environmentally relevant, context-dependent approach to examine the relationship between baseline GC levels and multiple fitness metrics in a declining aerial insectivore. Coupling this context-dependent approach with a multi-year dataset and an experimental manipulation of environmental quality, we were able to assess whether GC measures may be useful as broader conservation biomarkers. We found no relationships between baseline CORT metrics and any short- or longer-term fitness component that we measured in breeding female Tree Swallows. More specifically, baseline CORT at the incubation and at the nestling provisioning stage, and the change in CORT over the reproductive season, failed to predict key components of fitness, namely offspring quality, reproductive output, and adult survival probability. Importantly, this was the case under both natural conditions and when females were faced with an unexpected, experimentally induced decrease in foraging profitability during the nestling

TABLE 2. Model-averaged parameter estimates, unconditional standard errors, and 95% confidence intervals for models predicting fitness of control female Tree Swallows (n = 122).

Variable	Estimate	Unconditional SE	95% confidence interval (lower, upper)
(a) Offspring fledged (output)			
CORT1	0.002	0.01	−0.04, 0.06
CORT2	0.0008	0.01	−0.05, 0.04
CORTΔ	0.002	0.02	−0.08, 0.06
food	0.00004	0.001	−0.004, 0.005
lay date	0.008	0.006	−0.004, 0.02
massΔ	0.02	0.01	0.002, 0.05
(b) Offspring quality			
CORT1	−0.003	0.2	−0.89, 0.86
CORT2	−0.08	0.22	−1.05, 0.40
CORTΔ	−0.04	0.29	−1.39, 1.00
food	−0.005	0.02	−0.14, 0.02
lay date	−0.02	0.1	−0.21, 0.18
massΔ	−0.04	0.11	−0.56, 0.02
(c) Survival			
CORT1	0.007	0.05	−0.18, 0.25
CORT2	−0.0009	0.04	−0.18, 0.17
CORTΔ	0.006	0.06	−0.33, 0.26
fledged	−0.01	0.07	−0.29, 0.21
lay date	−0.02	0.02	−0.07, 0.03
massΔ	−0.01	0.03	−0.13, 0.05

Notes: All values were calculated using models included in the 95% confidence sets, weighting by Akaike weights. Abbreviations are as follow: massΔ is percent loss in body mass from incubation to nestling provisioning, CORT1 is baseline CORT at the incubation stage, CORT2 is baseline CORT at the nestling provisioning stage, CORTΔ is percent change in CORT from incubation to nestling provisioning, food is average insect biomass (food availability) during peak nestling provisioning, and fledged is the number of offspring fledged (see *Methods* for details).

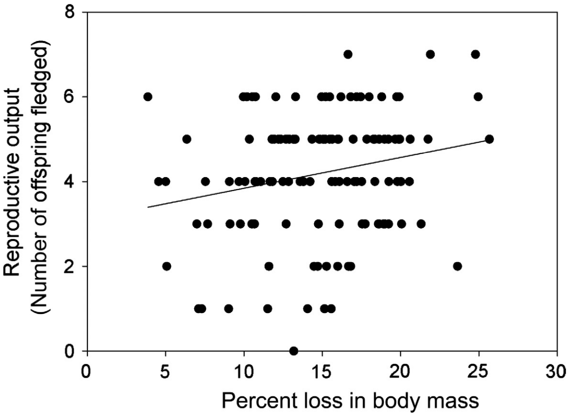

FIG. 2. Reproductive output (number of offspring fledged) in relation to percent loss in body mass over the reproductive season (from the late incubation to the mid-nestling provisioning stage) in female tree swallows (R^2 = 0.05).

provisioning stage. In contrast, control birds that lost a greater percentage of body mass over the reproductive season raised more offspring.

Lack of GC–fitness relationships

Variation in circulating GCs is often proposed as a useful conservation biomarker of exposure to anthropogenically induced stressors (Bonier et al. 2009*a*, Busch and Hayward 2009, Baker et al. 2013, Dantzer et al. 2014). However, the lack of GC–fitness relationships we found adds to the already variable findings previously reported across populations and species (Bonier et al. 2009*a*). It has been proposed that such a high level of variability could, at least in part, be due to the presence of underlying context-dependency in GC–fitness relationships (Bonier et al. 2009*a*, Madliger and Love 2014). In our investigation, the addition of energetic contexts such as food availability, reproductive investment, and somatic investment (i.e., decline in body mass) did not improve the capacity of baseline CORT measures to predict fitness outcomes. We were also rigorous in our work by (1) limiting analyses by sex and broad age class, (2) controlling for factors such as habitat (site) and reproductive timing, and (3) explicitly investigating the potential of varying CORT–fitness relationships across breeding stages and fitness metrics. Despite this attention to intrinsic and extrinsic context, both static measures and the change in baseline CORT over the reproductive season failed to be useful in predicting components of fitness. In theory, it is possible that additional contexts may be necessary to link baseline CORT levels and fitness. For example, as individuals are expected to alter their investment in their current brood based on their future reproductive potential, with baseline GCs potentially reflecting anticipated risks and demands during breeding, brood value may be particularly important in understanding the relationship between baseline GCs and fitness outcomes (Bókony et al. 2009). As would be the case in many conservation situations, we lacked knowledge on the specific age of each individual. We were therefore only able to partition birds into two age categories: (1) less than or equal to one year old, or (2) equal to/greater than two years, potentially encompassing ages of 2–12 years (Winkler et al. 2011) and as such, a range of potential brood values. While the possibility remains that age-related changes in baseline GCs (Angelier et al. 2006, Mateo 2006, Riechert et al. 2012, Hämäläinen et al. 2015) could influence our ability to detect a GC–fitness relationship, if fine-scale age metrics are necessary (i.e., if pooling age categories will mask a relationship between GCs and fitness), this will limit the application of the technique in many species of concern that are not easily aged in a field setting.

As is often the case regardless of whether a study is evolutionary or applied in nature, the measurement of lifetime reproductive success rather than fitness components is preferable (Newton 1989). This could therefore also be the case for properly assessing GC–fitness relationships. For example, Angelier et al. (2010) was able to predict five-year reproductive success with breeding baseline CORT levels in male Black-browed Albatross. In addition, Boonstra and Singleton (1993) found that

TABLE 3. Summary of confidence (95%) set of models predicting fitness metrics in female tree swallows in feather-clipped birds (n = 33).

Model variables	AIC$_c$	ΔAIC$_c$	ω	Cum. ω	R^2
(a) Offspring fledged (output)					
Null (lay date + site)	140.76	0.00	0.24	0.24	0.003
massΔ	141.52	0.76	0.16	0.40	0.06
CORT1	142.89	2.13	0.08	0.48	0.02
CORTΔ	143.24	2.49	0.07	0.55	0.01
CORT2	143.24	2.49	0.07	0.62	0.01
food	143.30	2.55	0.07	0.68	0.01
massΔ, CORT2	143.57	2.81	0.06	0.74	0.08
massΔ, CORT1	144.02	3.26	0.05	0.79	0.07
massΔ, food	144.15	3.39	0.04	0.83	0.07
CORTΔ, massΔ	144.27	3.52	0.04	0.87	0.06
CORT1, CORT2	145.66	4.90	0.02	0.89	0.02
food, CORT1	145.67	4.92	0.02	0.91	0.02
CORTΔ, food	146.00	5.24	0.02	0.93	0.01
food, CORT2	146.00	5.24	0.02	0.94	0.01
(b) Offspring quality					
CORT2	259.62	0.00	0.25	0.25	0.25
CORT1, CORT2	260.18	0.56	0.19	0.44	0.28
CORT1	261.21	1.59	0.11	0.55	0.29
CORTΔ	261.63	2.02	0.09	0.64	0.25
Null (lay date + site)	261.68	2.06	0.09	0.73	0.22
massΔ, CORT2	262.02	2.41	0.07	0.81	0.25
massΔ, CORT1, CORT2	262.75	3.13	0.05	0.86	0.27
massΔ, CORT1	263.32	3.70	0.04	0.90	0.29
massΔ	263.67	4.05	0.03	0.93	0.22
(c) Survival					
CORTΔ	47.34	0.00	0.31	0.31	0.19
Null (lay date + site)	48.83	1.49	0.15	0.45	0.0004
CORT1	49.80	2.46	0.09	0.54	0.06
CORTΔ, massΔ	49.91	2.57	0.08	0.62	0.21
CORTΔ, fledged	50.10	2.76	0.08	0.70	0.19
CORT2	51.06	3.71	0.05	0.75	0.02
massΔ	51.25	3.91	0.04	0.79	0.01
CORT1, CORT2	51.41	4.07	0.04	0.83	0.12
fledged	51.43	4.08	0.04	0.87	0.0006
massΔ, CORT1	52.49	5.15	0.02	0.90	0.07
fledged, CORT1	52.52	5.18	0.02	0.92	0.07
CORTΔ, massΔ, fledged	52.92	5.58	0.02	0.94	0.21
massΔ, CORT2	53.80	6.46	0.01	0.95	0.02

Notes: All models included lay date as a fixed effect and site as a random effect. The summary provides AIC values corrected for small sample size (AIC$_c$), ΔAIC$_c$, Akaike weight (ω), cumulative Akaike weights (cum. ω), and conditional R^2 for each model. Abbreviations are as follows: massΔ is percent loss in body mass from incubation to nestling provisioning, CORT1 is baseline CORT at the incubation stage, CORT2 is baseline CORT at the nestling provisioning stage, CORTΔ is percent change in CORT from incubation to nestling provisioning, food is average insect biomass (food availability) during peak nestling provisioning, and fledged is the number of offspring fledged (see *Methods* for details).

during years of population decline, showshoe hares (*Lepus americanus*) experience higher CORT levels compared to hares in good condition with ample food availability, illustrating the capacity for GCs to reflect longer-term population processes. Nonetheless, baseline GCs have been useful in predicting similar shorter-term fitness proxies to those that we measured across a variety of species (Bonier et al. 2009*b*). Instead, given that circulating plasma baseline GC levels are a more instantaneous measure of current energetic demand or environmental conditions, if CORT levels at the time periods we measured are driven by social or environmental conditions that do not have downstream consequences on fitness, this could inhibit the ability to relate baseline CORT to reproductive outcomes or survival. For example, one of the important known drivers of variation in baseline GCs is food availability (Astheimer et al. 1992, Kitaysky et al. 1999, 2010, Pravosudov et al. 2001, Corbel and Groscolas 2008, Jenni-Eiermann et al. 2008, Fokidis et al. 2012). As a result, short-lived nadirs in food availability (e.g., due to temporary decreases in temperature) could have large consequences for variability in baseline GCs (Astheimer et al. 1992) but may not be detrimental enough on body condition, reproductive behavior, or offspring growth to confer downstream fitness effects.

In addition, baseline GCs are known to vary in response to internal changes in state such as body condition and mass both within (Schoech et al. 1997, Romero and Wikelski 2001, Love et al. 2005, Cabezas et al. 2007, Williams et al. 2008) and across species (Hau et al. 2010). When female passerine birds begin to drop body mass at the end of incubation (most likely as an adaptive mechanism to decrease wing loading for the subsequent

Table 4. Model-averaged parameter estimates, unconditional standard errors, and 95% confidence intervals for models predicting fitness of feather-clipped female Tree Swallows ($n = 33$).

Variable	Estimate	Unconditional SE	95% confidence interval (lower, upper)
(a) Offspring fledged (output)			
CORT1	−0.006	0.02	−0.13, 0.07
CORT2	−0.005	0.02	−0.12, 0.07
CORTΔ	0.001	0.02	−0.12, 0.14
food	0.0001	0.001	−0.005, 0.07
lay date	0.005	0.02	−0.03, 0.04
massΔ	0.01	0.02	−0.01, 0.07
(b) Offspring quality			
CORT1	−0.35	0.89	−3.39, 1.66
CORT2	−0.99	1.17	−3.89, 0.51
CORTΔ	−0.06	0.6	−3.78, 2.88
lay date	0.43	0.39	−0.4, 1.23
massΔ	−0.03	0.01	−1.27, 1.04
(c) Survival			
CORT1	0.06	0.15	−0.16, 0.72
CORT2	−0.02	0.1	−0.61, 0.28
CORTΔ	−0.31	0.41	−1.38, 0.09
fledged	0.007	0.1	−0.43, 0.5
lay date	0.01	0.07	−0.13, 0.16
massΔ	0.002	0.05	−0.22, 0.24

Notes: All values were calculated using models included in the 95% confidence sets, weighting by Akaike weights. Abbreviations are as follows: massΔ is percent loss in body mass from incubation to nestling provisioning, CORT1 is baseline CORT at the incubation stage, CORT2 is baseline CORT at the nestling provisioning stage, CORTΔ is percent change in CORT from incubation to nestling provisioning, food is average insect biomass (food availability) during peak nestling provisioning, and fledged is the number of offspring fledged (see *Methods* for details).

nestling provisioning period; Freed 1981, Norberg 1981, Neto and Gosler 2009, Boyle et al. 2012), they may be experiencing changes in metabolic costs or may be adjusting food intake in preparation for chick hatching (Portugal et al. 2007, Boyle et al. 2012). If baseline CORT acts as a mediator of this change (e.g., through mobilization of fat stores), small differences in the timing of when females initiate this adaptive reduction in mass could lead to highly variable CORT levels across individuals at this sampling time. This potential involvement of (or influence on) GCs highlights the importance of also considering the metabolic role of GCs across fine temporal scales and illustrates that even small differences in sampling times could lead to altered relationships between GCs and fitness (Crespi et al. 2013).

Our results may differ from other findings in the same species (e.g., Bonier et al. 2009*b*, 2011) for a number of methodological reasons. Although we chose sampling windows to coincide with expected demands within individual reproductive stages, our baseline CORT measures were obtained at a later stage in both incubation and nestling provisioning compared to previous work in this species (Bonier et al. 2009*b*, 2011). It is possible that the time periods we sampled are less sensitive to the environmental factors that may influence reproductive success or that females are highly committed to nesting attempts very late in incubation (i.e., they may be more able to buffer extrinsic environmental changes due to increased body reserves; Boyle et al. 2012). It is also possible that relationships may differ between years or sites; it has been proposed that differences in the contribution of ecological factors and breeding effort to allostatic load may alter the relationship between GCs and fitness (Ebensperger et al. 2013). More specifically, when characteristics of the ecological or social environment are the main drivers of GC levels, as opposed to reproductive effort, a lack of relationship between GCs and fitness is expected (Bonier et al. 2009*a*, Ebensperger et al. 2013). For example, Ebensperger et al. (2011) only found an association between fecal GCs and fitness (number and survival of offspring) in degus during years of high food availability. As a result, performing our feather-clipping manipulation in a different year may have led to alternative results given that our manipulation year (2011) was a reasonably stable one in terms of weather, as compared to others with harsher or unanticipated weather events, or above-average food resources (e.g., May snowfalls; Christine Madliger, *personal observation*). Indeed, weather conditions have been shown to alter responses to GC manipulations (O'Connor et al. 2010, Thierry et al. 2013, Ouyang et al. 2015). For example, Ouyang et al. (2015) found that CORT-implanted and control Tree Swallows had similar nest failure rates when weather conditions were harsh (cold and rainy) but that CORT supplementation caused quicker brood failure during favorable conditions. Additionally, it has been proposed that GC–fitness relationships may be more readily found when a portion of, but not all, individuals in a population are constrained by their environment (Angelier 2010). Overall, the females in our study may have been able to cope with the constraints of the manipulation (Patterson et al. 2011), potentially decreasing its total influence on CORT–fitness relationships. Overall, it remains unclear how environmental variability may influence GC–fitness relationships.

A final explanation for a general lack of CORT–fitness relationships, particularly at the nestling provisioning stage when chicks begin to plateau in body mass (Quinney et al. 1986, McCarty 2001), is the possibility that males could compensate for reduced foraging ability in females (Patterson et al. 2011), leading to unaltered nesting success despite alteration in female CORT levels. This would be a particularly important factor for manipulated females, whose foraging rates decrease (Madliger et al. 2015), but whose breeding success was comparable to controls (Madliger and Love 2016). Indeed, the total number of foraging trips to manipulated and control nests was equivalent (Madliger and Love 2016), indicating that males did compensate to ensure a certain overall foraging rate for their brood (Patterson et al. 2011). In particular, this highlights the possibility that GC levels fail to reflect fitness due to unmeasured variables (e.g., mate quality) and has implications for measuring GC levels during the breeding season in species with bi-parental care, particularly if environmental conditions affect the sexes differentially.

Loss of body mass as a predictor of reproductive success

Our results indicate that the loss in body mass over reproduction was a more sensitive predictor of within-season reproductive success than multiple measures of baseline CORT. It is possible that the change in body mass is directly indicative of energetic investment, where individuals that invest the most in offspring are accruing the greatest fitness benefits but are suffering from greater losses in somatic body condition (Drent and Daan 1980, Bryant 1988, Neto and Gosler 2009). Additionally, lower body masses are likely indicative of an adaptive change in body mass to increase flight efficiency during demanding stages of reproduction (Freed 1981, Norberg 1981, Neto and Gosler 2009), particularly in species that forage solely on the wing (Boyle et al. 2012). It may therefore be the ability of females to adjust their body mass to foraging demands that may represent the most reliable indicator of individual quality (Boyle et al. 2012).

Overall, the greater utility of a body mass metric compared to GCs is of interest from a management perspective, given the monetary cost of analyzing GC levels, considerations for storage in field settings (Sheriff et al. 2011), and invasiveness of blood sampling. However, while a change in body mass did predict reproductive output statistically, the fit (R^2) of the relationship was weak and was only evident in control birds. As a result, there is still a large amount of variation in reproductive output that is not well-captured with metrics of body mass, limiting the application of such a measure in conservation field settings. Nonetheless, our findings do reinforce the idea that, at some stages, loss in body condition can actually be an indication of a high quality individual investing heavily in reproduction, with concomitant fitness benefits (Hillstrom 1995, Gillooly and Baylis 1999, Golet and Irons 1999, Breuner 2011). This further draws attention to the importance of considering expected energetic demands of the organism of interest at the stage of sampling in conservation applications, and otherwise (Madliger and Love 2014).

Implications for baseline GCs as conservation biomarkers and recommendations for future study

We have illustrated that measures of baseline GCs may fail to provide reliable biomarkers of reproductive success or survival probability in some populations and at certain time points, despite a robust experimental design, the measurement of multiple within-individual metrics, and careful investigation of the intrinsic and extrinsic contexts that may influence GC–fitness relationships. Most importantly, these results indicate the importance of validation prior to application of GCs in conservation settings, since varying GC levels may not always be indicative of population-level persistence. Additionally, examination of our results in comparison to others in the same species within the same geographic region (e.g., Bonier et al. 2009*b*, 2011) provides additional evidence that GC–fitness relationships can change based on the time frame in which a GC measure is obtained (potentially in as little time as one week) or that the relationship could vary substantially by site or year. Finally, our results draw attention to the potential importance of measuring multiple time points per individual when investigating biomarkers of fitness. All of these considerations have important implications for the ease of use, costs, and time frames that may be necessary to monitor populations of conservation concern using physiological traits. Overall, our results demonstrate that GC measures, particularly those as labile as plasma baseline GCs, will need to be validated within specific populations prior to use as conservation monitoring tools. Indeed, mounting evidence is indicating that species- and context-specific studies are necessary before conservation managers can feel confident about the interpretation of changing GC levels in their systems (Sopinka et al. 2015).

It should be noted that the functioning of the HPA axis in the acute stress response may be a more reliable indicator of chronic stressors than baseline plasma GCs. For example, a meta-analysis by Dickens and Romero (2013) indicated that the lack of ability to terminate a stress response (in investigations employing dexamethasone) was relatively consistently related to chronic stress (Dantzer et al. 2014). However, measurements related to stress responsiveness and attenuation necessitate capture, restraint, repeated blood sampling, injections of adrenocorticotropic hormone or dexamethasone and, in some cases anesthesia (Sheriff et al. 2011), which will mean that this approach will be extremely limited and case-specific within the realm of conservation. As a result, it is important to acknowledge that a diversity of media exist that require substantially less invasive sampling than plasma metrics. Feces, feathers, hair, and other integuments (baleen, claws, skin, etc.) provide alternatives, some of which do not even require direct capture of the organism of interest (Millspaugh and Washburn 2004, Sheriff et al. 2011, Berkvens et al. 2013, Baxter-Gilbert et al. 2014, Hunt et al. 2014). Because these samples integrate GC activity over longer time periods, they may also provide a more realistic snapshot of how an individual is responding to prevailing environmental conditions (Dantzer et al. 2014). Furthermore, in the case of integumentary GCs like hair and feathers, they may allow acquisition of information regarding distant time periods (e.g., overwintering, migration) when animals are inaccessible. A recent meta-analysis indicated that fecal GCs show a consistent increase in response to anthropogenic disturbance, while baseline plasma GCs do not; however, the fitness consequences of these increases have been markedly under-investigated (Dantzer et al. 2014). Despite a less extensive literature base (particularly for integumentary GCs) there is certainly still context-dependency and the presence of variable results, as in the case of plasma GCs. We suggest that non-invasive measures warrant special attention, given their compatibility with many conservation goals of minimizing disturbance, and we urge researchers to continue to carefully validate these techniques in terms of repeatability (Madliger and Love 2015, Killen et al. 2016), links to fitness metrics (Bonier et al. 2009*a*), time periods of integration (Millspaugh and Washburn 2004, Bortolotti et al. 2008, Meyer and Novak 2012), stability of GC levels (Bortolotti et al. 2009), hormone extraction (e.g., small sample artifacts; Tempel and Gutiérrez 2004, Lattin et al. 2011), and relationship to environmental and/or chronic stressors (Baker et al. 2013, Dickens and Romero 2013).

Future studies should also place emphasis on whether static measures of GCs are sufficient to understand variation in disturbance, condition, or fitness. In addition, it is becoming increasingly clear that a GC–fitness relationship observed using one metric of fitness should not be expected to apply to other metrics (e.g., reproductive output vs. survival probability). As a result, investigating the GC-fitness metric that is most important to the population of interest will be of paramount importance to properly interpreting changing GC levels over time. This also has implications for seasonal changes in GC–fitness relationships, indicating that it will be important to determine when in the life cycle a predictive relationship exists. Finally, manipulative studies that compare how changes within normal baseline GC levels may change behavior and fitness will further elucidate when GC–fitness relationships may exist, and, more broadly, how GC levels may mediate life history decisions (Crespi et al. 2013, Sopinka et al. 2015). In particular, habitat quality, disturbance, or GC manipulations that influence both sexes would likely be advantageous and relevant for determining how GC–fitness relationships may directly change in response to environmental alteration. Overall, there is still a great deal of validation necessary before GC levels can be reliably utilized as conservation biomarkers in many species and, in some cases, the effort required for monitoring and sampling the contexts necessary for successful interpretation of baseline GC levels may be better placed towards directly measuring fitness metrics.

Acknowledgments

We would like to thank Ruthven Park National Historic Site, the Grand River Conservation Authority, and Habitat Haldimand for access to study sites. We would also like to thank Rick Ludkin, Nancy Furber, and Peter Marier for assistance with data collection. We appreciate the constructive comments from two anonymous reviewers that improved the manuscript. Funding for this work was provided by an Ontario Graduate Scholarship and an NSERC PGS-D (427552) grant to C. L. Madliger, and NSERC Discovery (2015-478021), Canada Research Chair (30342), and Canada Foundation for Innovation (29401) awards to O. P. Love.

Literature Cited

Angelier, F., S. A. Shaffer, H. Weimerskirch, and O. Chastel. 2006. Effect of age, breeding experience and senescence on corticosterone and prolactin levels in a long-lived seabird: the wandering albatross. General and Comparative Endocrinology 149:1–9.

Angelier, F., and J. C. Wingfield. 2013. Importance of the glucocorticoid stress response in a changing world: theory, hypotheses and perspectives. General and Comparative Endocrinology 190:118–128.

Angelier, F., J. C. Wingfield, H. Weimerskirch, and O. Chastel. 2010. Hormonal correlates of individual quality in a long-lived bird: a test of the "corticosterone-fitness hypothesis." Biology Letters 6:846–849.

Ardia, D. R., and E. D. Clotfelter. 2007. Individual quality and age affect responses to an energetic constraint in a cavity-nesting bird. Behavioral Ecology 18:259–266.

Arlettaz, R., S. Nusslé, M. Baltic, P. Vogel, R. Palme, S. Jenni-Eiermann, P. Patthey, and M. Genoud. 2014. Disturbance of wildlife by outdoor winter recreation: allostatic stress response and altered activity–energy budgets. Ecological Applications 25:1197–1212.

Astheimer, L. B., W. A. Buttemer, and J. C. Wingfield. 1992. Interactions of corticosterone with feeding, activity and metabolism in passerine birds. Ornis Scandinavica 23:355–365.

Baker, M. R., K. S. Gobush, and C. H. Vynne. 2013. Review of factors influencing stress hormones in fish and wildlife. Journal for Nature Conservation 21:309–318.

Barton, K. 2015. MuMIn: multi-model inference. R package, version 1.15.1.

Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Baxter-Gilbert, J. H., J. L. Riley, G. F. Mastromonaco, J. D. Litzgus, and D. Lesbarrères. 2014. A novel technique to measure chronic levels of corticosterone in turtles living around a major roadway. Conservation Physiology 2:cou036. http://dx.doi.org/10.1093/conphys/cou036

Berkvens, C. N., C. Hyatt, C. Gilman, D. L. Pearl, I. K. Barker, and G. F. Mastromonaco. 2013. Validation of a shed skin corticosterone enzyme immunoassay in the African House Snake (*Lamprophis fuliginosus*) and its evaluation in the Eastern Massasauga Rattlesnake (*Sistrurus catenatus catenatus*). General and Comparative Endocrinology 194:1–9.

Bókony, V., Á. Z. Lendvai, A. Liker, F. Angelier, J. C. Wingfield, and O. Chastel. 2009. Stress response and the value of reproduction: Are birds prudent parents? American Naturalist 173:589–598.

Bolker, B. M., M. E. Brooks, C. J. Clark, S. W. Geange, J. R. Poulsen, M. H. H. Stevens, and J.-S. S. White. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology & Evolution 24:127–135.

Bonier, F., P. R. Martin, I. T. Moore, and J. C. Wingfield. 2009*a*. Do baseline glucocorticoids predict fitness? Trends in Ecology & Evolution 24:634–642.

Bonier, F., I. T. Moore, P. R. Martin, and R. J. Robertson. 2009*b*. The relationship between fitness and baseline glucocorticoids in a passerine bird. General and Comparative Endocrinology 163:208–213.

Bonier, F., I. T. Moore, and R. J. Robertson. 2011. The stress of parenthood? Increased glucocorticoids in birds with experimentally enlarged broods. Biology Letters 7:944–946.

Boonstra, R., and G. R. Singleton. 1993. Population declines in the snowshoe hare and the role of stress. General and Comparative Endocrinology 91:126–143.

Bortolotti, G. R., T. Marchant, J. Blas, and S. Cabezas. 2009. Tracking stress: localisation, deposition and stability of corticosterone in feathers. Journal of Experimental Biology 212:1477–1482.

Bortolotti, G. R., T. A. Marchant, J. Blas, and T. German. 2008. Corticosterone in feathers is a long-term, integrated measure of avian stress physiology. Functional Ecology 22: 494–500.

Boyle, W. A., D. W. Winkler, and C. G. Guglielmo. 2012. Rapid loss of fat but not lean mass prior to chick provisioning supports the flight efficiency hypothesis in tree swallows. Functional Ecology 26:895–903.

Breuner, C. W. 2011. Stress and reproduction in birds. Pages 129–151 *in* D. O. Norris and K. H. Lopez, editors. Hormones and reproduction of vertebrates: birds. Academic Press, London, UK.

Breuner, C. W., S. Lynn, G. Julian, J. Cornelius, B. Heidinger, O. Love, R. Sprague, H. Wada, and B. Whitman. 2006. Plasma-binding globulins and acute stress response. Hormone and Metabolic Research 38:260–268.

Bryant, D. M. 1988. Energy expenditure and body mass changes as measures of reproductive costs in birds. Functional Ecology 2:23–34.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and multimodel inference. Second edition. Springer, New York, USA.

Busch, D. S., and L. S. Hayward. 2009. Stress in a conservation context: a discussion of glucocorticoid actions and how levels change with conservation-relevant variables. Biological Conservation 142:2844–2853.

Cabezas, S., J. Blas, T. A. Marchant, and S. Moreno. 2007. Physiological stress levels predict survival probabilities in wild rabbits. Hormones and Behavior 51:313–320.

Carey, C. 2005. How physiological methods and concepts can be useful in conservation biology. Integrative and Comparative Biology 45:4–11.

Cook, K. V., S. H. McConnachie, K. M. Gilmour, S. G. Hinch, and S. J. Cooke. 2011. Fitness and behavioral correlates of pre-stress and stress-induced plasma cortisol titers in pink salmon (*Oncorhynchus gorbuscha*) upon arrival at spawning grounds. Hormones and Behavior 60:489–497.

Cooke, S. J. 2014. Conservation physiology today and tomorrow. Conservation Physiology 2:cot033. http://dx.doi.org/10.1093/conphys/cot033

Cooke, S. J., and C. M. O'Connor. 2010. Making conservation physiology relevant to policy makers and conservation practitioners. Conservation Letters 3:159–166.

Cooke, S. J., L. Sack, C. E. Franklin, A. P. Farrell, J. Beardall, M. Wikelski, and S. L. Chown. 2013. What is conservation physiology? Perspectives on an increasingly integrated and essential science. Conservation Physiology 1:cot001. http://dx.doi.org/10.1093/conphys/cot001

Corbel, H., and R. Groscolas. 2008. A role for corticosterone and food restriction in the fledging of nestling white storks. Hormones and Behavior 53:557–566.

Crespi, E. J., T. D. Williams, T. S. Jessop, and B. Delehanty. 2013. Life history and the ecology of stress: How do glucocorticoid hormones influence life-history variation in animals? Functional Ecology 27:93–106.

D'Alba, L., K. A. Spencer, R. G. Nager, and P. Monaghan. 2011. State dependent effects of elevated hormone: nest site quality, corticosterone levels and reproductive performance in the common eider. General and Comparative Endocrinology 172:218–224.

Dantzer, B., Q. E. Fletcher, R. Boonstra, and M. J. Sheriff. 2014. Measures of physiological stress: A transparent or opaque window into the status, management and conservation of species? Conservation Physiology 2:cou023. http://dx.doi.org/10.1093/conphys/cou023

Dickens, M. J., and L. M. Romero. 2013. A consensus endocrine profile for chronically stressed wild animals does not exist. General and Comparative Endocrinology 191: 177–189.

Drent, R., and S. Daan. 1980. The prudent parent: energetic adjustments in avian breeding. Ardea 68:225–252.

Ebensperger, L. A., J. Ramírez-Estrada, C. León, R. A. Castro, L. O. Tolhuysen, R. Sobrero, V. Quirici, J. R. Burger, M. Soto-Gamboa, and L. D. Hayes. 2011. Sociality, glucocorticoids and direct fitness in the communally rearing rodent, *Octodon degus*. Hormones and Behavior 60:346–352.

Ebensperger, L. A., D. Tapia, J. Ramírez-Estrada, C. León, M. Soto-Gamboa, and L. D. Hayes. 2013. Fecal cortisol levels predict breeding but not survival of females in the short-lived rodent, *Octodon degus*. General and Comparative Endocrinology 186:164–171.

Ellis, R. D., T. J. McWhorter, and M. Maron. 2011. Integrating landscape ecology and conservation physiology. Landscape Ecology 27:1–12.

Escribano-Avila, G., N. Pettorelli, E. Virgós, C. Lara-Romero, J. Lozano, I. Barja, F. S. Cuadra, and M. Puerta. 2013. Testing Cort–fitness and Cort–adaptation hypotheses in a habitat suitability gradient for roe deer. Acta Oecologica 53:38–48.

Fokidis, H. B., M. B. des Roziers, R. Sparr, C. Rogowski, K. Sweazea, and P. Deviche. 2012. Unpredictable food availability induces metabolic and hormonal changes independent of food intake in a sedentary songbird. Journal of Experimental Biology 215:2920–2930.

Freed, L. A. 1981. Loss of mass in breeding wrens: Stress or adaptation? Ecology 62:1179–1186.

Gillooly, J. F., and J. R. Baylis. 1999. Reproductive success and the energetic cost of parental care in male smallmouth bass. Journal of Fish Biology 54:573–584.

Golet, G. H., and D. B. Irons. 1999. Raising young reduces body condition and fat stores in black-legged kittiwakes. Oecologia 120:530–538.

Hämäläinen, A., M. Heistermann, and C. Kraus. 2015. The stress of growing old: sex- and season-specific effects of age on allostatic load in wild grey mouse lemurs. Oecologia 178:1063–1075.

Hau, M., R. E. Ricklefs, M. Wikelski, K. A. Lee, and J. D. Brawn. 2010. Corticosterone, testosterone and life-history strategies of birds. Proceedings of the Royal Society of London B 277:3203–3212.

Hillstrom, L. 1995. Body mass reduction during reproduction in the pied flycatcher *Ficedula hypoleuca*: Physiological stress or adaptation for lowered costs of locomotion? Functional Ecology 9:807–817.

Hunt, K. E., R. Stimmelmayr, C. George, C. Hanns, R. Suydam, H. Brower, and R. M. Rolland. 2014. Baleen hormones: a novel tool for retrospective assessment of stress and reproduction in bowhead whales (*Balaena mysticetus*). Conservation Physiology 2:cou030. http://dx.doi.org/10.1093/conphys/cou030

Hussell, D. J. 1983. Age and plumage color in female tree swallows. Journal of Field Ornithology 54:312–318.

Hussell, D. J., and T. E. Quinney. 1987. Food abundance and clutch size of tree swallows *Tachycineta bicolor*. Ibis 129: 243–258.

Jaatinen, K., M. W. Seltmann, T. Hollmén, S. Atkinson, K. Mashburn, and M. Öst. 2013. Context dependency of baseline glucocorticoids as indicators of individual quality in a capital breeder. General and Comparative Endocrinology 191:231–238.

Jenni-Eiermann, S., E. Glaus, M. Grüebler, H. Schwabl, and L. Jenni. 2008. Glucocorticoid response to food availability in breeding barn swallows (*Hirundo rustica*). General and Comparative Endocrinology 155:558–565.

Johnson, J. B., and K. S. Omland. 2004. Model selection in ecology and evolution. Trends in Ecology & Evolution 19:101–108.

Jones, J. 2003. Tree swallows (*Tachycineta bicolor*): A new model organism? Auk 120:591–599.

Killen, S. S., B. Adriaenssens, S. Marras, G. Claireaux, and S. J. Cooke. 2016. Context dependency of trait repeatability and its relevance for management and conservation of fish populations. Conservation Physiology 4:cow007. http://dx.doi.org/10.1093/conphys/cow007

Kitaysky, A. S., J. F. Piatt, S. A. Hatch, E. V. Kitaiskaia, Z. M. Benowitz-Fredericks, M. T. Shultz, and J. C. Wingfield. 2010. Food availability and population processes: severity of nutritional stress during reproduction predicts survival of long-lived seabirds. Functional Ecology 24:625–637.

Kitaysky, A. S., J. C. Wingfield, and J. F. Piatt. 1999. Dynamics of food availability, body condition and physiological stress response in breeding black-legged kittiwakes. Functional Ecology 13:577–584.

Korte, S. M., J. M. Koolhaas, J. C. Wingfield, and B. S. McEwen. 2005. The Darwinian concept of stress: benefits of allostasis and costs of allostatic load and the trade-offs in health and disease. Neuroscience and Biobehavioral Reviews 29:3–38.

Landys, M. M., M. Ramenofsky, and J. C. Wingfield. 2006. Actions of glucocorticoids at a seasonal baseline as compared to stress-related levels in the regulation of periodic life processes. General and Comparative Endocrinology 148:132–149.

Lattin, C. R., J. M. Reed, D. W. DesRochers, and L. M. Romero. 2011. Elevated corticosterone in feathers correlates with corticosterone-induced decreased feather quality: a validation study. Journal of Avian Biology 42:247–252.

Lennox, R., and S. J. Cooke. 2014. State of the interface between conservation and physiology: a bibliometric analysis. Conservation Physiology 2:cou003. http://dx.doi.org/10.1093/conphys/cou003

Love, O. P., C. W. Breuner, F. Vézina, and T. D. Williams. 2004. Mediation of a corticosterone-induced reproductive conflict. Hormones and Behavior 46:59–65.

Love, O. P., E. H. Chin, K. E. Wynne-Edwards, and T. D. Williams. 2005. Stress hormones: a link between maternal condition and sex-biased reproductive investment. American Naturalist 166:751–766.

Love, O. P., C. L. Madliger, S. Bourgeon, C. A. D. Semeniuk, and T. D. Williams. 2014. Evidence for baseline glucocorticoids as mediators of reproductive investment in a wild bird. General and Comparative Endocrinology 199:65–69.

Love, O., and T. Williams. 2008. Plasticity in the adrenocortical response of a free-living vertebrate: the role of pre- and postnatal developmental stress. Hormones and Behavior 54:496–505.

Madliger, C. L., et al. 2016. Success stories and emerging themes in conservation physiology. Conservation Physiology 4: cov057. http://dx.doi.org/10.1093/conphys/cov057

Madliger, C. L., and O. P. Love. 2014. The need for a predictive, context-dependent approach to the application of stress hormones in conservation. Conservation Biology 28:283–287.

Madliger, C. L., and O. P. Love. 2015. The power of physiology in changing landscapes: considerations for the continued integration of conservation and physiology. Integrative and Comparative Biology 55:545–553.

Madliger, C. L., and O. P. Love. 2016. Do baseline glucocorticoids simultaneously represent fitness and environmental quality in a declining aerial insectivore? Oikos. http://dx.doi.org/10.1111/oik.03354

Madliger, C. L., C. A. Semeniuk, C. M. Harris, and O. P. Love. 2015. Assessing baseline stress physiology as an integrator of environmental quality in a wild avian population: implications for use as a conservation biomarker. Biological Conservation 192:409–417.

Mateo, J. M. 2006. Developmental and geographic variation in stress hormones in wild Belding's ground squirrels (*Spermophilus beldingi*). Hormones and Behavior 50:718–725.

McCarty, J. P. 2001. Variation in growth of nestling tree swallows across multiple temporal and spatial scales. Auk 118:176–190.

McCracken, J. D. 2008. Are aerial insectivores being "bugged out"? Birdwatch Canada 42:4–7.

McEwen, B. S., and J. C. Wingfield. 2010. What is in a name? Integrating homeostasis, allostasis and stress. Hormones and Behavior 57:105–111.

Meyer, J. S., and M. A. Novak. 2012. Minireview: Hair cortisol: a novel biomarker of hypothalamic-pituitary-adrenocortical activity. Endocrinology 153:4120–4127.

Millspaugh, J. J., and B. E. Washburn. 2004. Use of fecal glucocorticoid metabolite measures in conservation biology research: considerations for application and interpretation. General and Comparative Endocrinology 138:189–199.

Nebel, S., A. Mills, J. D. McCracken, and P. D. Taylor. 2010. Declines of aerial insectivores in North America follow a geographic gradient. Avian Conservation and Ecology 5:1.

Neto, J. M., and A. G. Gosler. 2009. Variation in body condition of breeding Savi's Warblers *Locustella luscinioides*: the reproductive stress and flight adaptation hypothesis revisited. Journal of Ornithology 151:201–210.

Newton, I. 1989. Lifetime reproduction in birds. Academic Press, London, UK.

Norberg, R. A. 1981. Temporary weight decrease in breeding birds may result in more fledged young. American Naturalist 118:838–850.

O'Connor, C. M., K. M. Gilmour, R. Arlinghaus, C. T. Hasler, D. P. Philipp, and S. J. Cooke. 2010. Seasonal carryover effects following the administration of cortisol to a wild teleost fish. Physiological and Biochemical Zoology: Ecological and Evolutionary Approaches 83:950–957.

Ouyang, J. Q., Á. Lendvai, R. Dakin, A. Domalik, V. Fasanello, B. Vassallo, M. Haussmann, I. Moore, and F. Bonier. 2015. Weathering the storm: parental effort and experimental manipulation of stress hormones predict brood survival. BMC Evolutionary Biology 15:219.

Ouyang, J. Q., M. Muturi, M. Quetting, and M. Hau. 2013*a*. Small increases in corticosterone before the breeding season increase parental investment but not fitness in a wild passerine bird. Hormones and Behavior 63:776–781.

Ouyang, J. Q., P. J. Sharp, A. Dawson, M. Quetting, and M. Hau. 2011. Hormone levels predict individual differences in reproductive success in a passerine bird. Proceedings of the Royal Society of London B 278:2537–2545.

Ouyang, J. Q., P. Sharp, M. Quetting, and M. Hau. 2013*b*. Endocrine phenotype, reproductive success and survival in the great tit, *Parus major*. Journal of Evolutionary Biology 26:1988–1998.

Paquette, S. R., F. Pelletier, D. Garant, and M. Bélisle. 2014. Severe recent decrease of adult body mass in a declining insectivorous bird population. Proceedings of the Royal Society of London B 281:20140649.

Patterson, S. H., D. W. Winkler, and C. W. Breuner. 2011. Glucocorticoids, individual quality and reproductive investment in a passerine bird. Animal Behaviour 81:1239–1247.

Portugal, S. J., J. A. Green, and P. J. Butler. 2007. Annual changes in body mass and resting metabolism in captive barnacle geese (*Branta leucopsis*): the importance of wing moult. Journal of Experimental Biology 210:1391–1397.

Pravosudov, V. V., A. S. Kitaysky, J. C. Wingfield, and N. S. Clayton. 2001. Long-term unpredictable foraging conditions and physiological stress response in mountain chickadees (*Poecile gambeli*). General and Comparative Endocrinology 123:324–331.

Quinney, T. E., D. J. Hussell, and C. D. Ankney. 1986. Sources of variation in growth of tree swallows. Auk 103:389–400.

R Development Team. 2015. R: a language and environment for statistical computing. Version 3.2.1. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org/

Riechert, J., O. Chastel, and P. H. Becker. 2012. Why do experienced birds reproduce better? Possible endocrine mechanisms in a long-lived seabird, the common tern. General and Comparative Endocrinology 178:391–399.

Rivers, J. W., A. L. Liebl, J. C. Owen, L. B. Martin, and M. G. Betts. 2012. Baseline corticosterone is positively related to juvenile survival in a migrant passerine bird. Functional Ecology 26:1127–1134.

Romero, L. M., and J. M. Reed. 2005. Collecting baseline corticosterone samples in the field: Is under 3 min good enough? Comparative Biochemistry and Physiology Part A 140: 73–79.

Romero, L. M., and M. Wikelski. 2001. Corticosterone levels predict survival probabilities of Galapagos marine iguanas during El Nino events. Proceedings of the National Academy of Sciences USA 98:7366–7370.

Schoech, S. J., R. L. Mumme, and J. C. Wingfield. 1997. Corticosterone, reproductive status, and body mass in a cooperative breeder, the Florida scrub-jay (*Aphelocoma coerulescens*). Physiological Zoology 70:68–73.

Sheriff, M. J., B. Dantzer, B. Delehanty, R. Palme, and R. Boonstra. 2011. Measuring stress in wildlife: techniques for quantifying glucocorticoids. Oecologia 166:869–887.

Sopinka, N. M., L. D. Patterson, J. C. Redfern, N. K. Pleizier, C. B. Belanger, J. D. Midwood, G. T. Crossin, and S. J. Cooke. 2015. Manipulating glucocorticoids in wild animals: basic and applied perspectives. Conservation Physiology 3:cov031. http://dx.doi.org/10.1093/conphys/cov031

Stutchbury, B. J., and R. J. Robertson. 1988. Within-season and age-related patterns of reproductive performance in female tree swallows (*Tachycineta bicolor*). Canadian Journal of Zoology 66:827–834.

Stutchbury, B. J., and S. Rohwer. 1990. Molt patterns in the tree swallow (*Tachycineta bicolor*). Canadian Journal of Zoology 68:1468–1472.

Symonds, M. R. E., and A. Moussalli. 2011. A brief guide to model selection, multimodel inference and model averaging in behavioural ecology using Akaike's information criterion. Behavioral Ecology and Sociobiology 65:13–21.

Tempel, D. J., and R. J. Gutiérrez. 2004. Factors related to fecal corticosterone levels in california spotted owls: implications for assessing chronic stress. Conservation Biology 18:538–547.

Thierry, A.-M., S. Massemin, Y. Handrich, and T. Raclot. 2013. Elevated corticosterone levels and severe weather conditions decrease parental investment of incubating Adélie penguins. Hormones and Behavior 63:475–483.

Wada, H., T. P. Hahn, and C. W. Breuner. 2007. Development of stress reactivity in white-crowned sparrow nestlings: total corticosterone response increases with age, while free corticosterone response remains low. General and Comparative Endocrinology 150:405–413.

Wikelski, M., and S. J. Cooke. 2006. Conservation physiology. Trends in Ecology & Evolution 21:38–46.

Williams, T. D. 2008. Individual variation in endocrine systems: moving beyond the "tyranny of the Golden Mean." Philosophical Transactions of the Royal Society B 363: 1687–1698.

Williams, C. T., A. S. Kitaysky, A. B. Kettle, and C. L. Buck. 2008. Corticosterone levels of tufted puffins vary with breeding stage, body condition index, and reproductive performance. General and Comparative Endocrinology 158:29–35.

Wingfield, J. C. 2005. The concept of allostasis: coping with a capricious environment. Journal of Mammalogy 86: 248–254.

Wingfield, J. C. 2013. Ecological processes and the ecology of stress: the impacts of abiotic environmental factors. Functional Ecology 27:37–44.

Wingfield, J. C., and A. S. Kitaysky. 2002. Endocrine responses to unpredictable environmental events: Stress or anti-stress hormones? Integrative and Comparative Biology 42: 600–609.

Wingfield, J. C., D. L. Maney, C. W. Breuner, J. D. Jacobs, S. Lynn, M. Ramenofsky, and R. D. Richardson. 1998. Ecological bases of hormone-behavior interactions: the "emergency life history stage." American Zoologist 38:191–206.

Wingfield, J. C., and R. M. Sapolsky. 2003. Reproduction and resistance to stress: When and how. Journal of Neuroendocrinology 15:711–724.

Winkler, D. W., K. K. Hallinger, D. R. Ardia, R. J. Robertson, B. J. Stutchbury, and R. R. Cohen. 2011. Tree swallow (*Tachycineta bicolor*). The Birds of North America. P. G. Rodewald, editor. Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America: https://birdsna.org/Species-Account/bna/species/treswa DOI: 10.2173/bna.11

Winkler, D. W., and P. E. Allen. 1995. Effects of handicapping on female condition and reproduction in tree swallows (*Tachycineta bicolor*). Auk 112:737–747.

Winkler, D. W., P. H. Wrege, P. E. Allen, T. L. Kast, P. Senesac, M. F. Wasson, P. E. Llambías, V. Ferretti, and P. J. Sullivan. 2004. Breeding dispersal and philopatry in the tree swallow. Condor 106:768–776.

Zhang, S., F. Lei, S. Liu, D. Li, C. Chen, and P. Wang. 2011. Variation in baseline corticosterone levels of tree sparrow (*Passer montanus*) populations along an urban gradient in Beijing, China. Journal of Ornithology 152:801–806.

DATA AVAILABILITY

Data associated with this manuscript have been deposited in Dryad: http://dx.doi.org/10.5061/dryad.mg529