



Original Article

Achromatic plumage brightness predicts stress resilience and social interactions in tree swallows (*Tachycineta bicolor*)

Conor C. Taff,^{a,b,●} Cedric Zimmer,^a and Maren N. Vitousek^{a,b}

^aCorson Hall, Cornell University, Ithaca, NY 14853 ^bLab of Ornithology, Cornell University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA

Received 18 December 2017; revised 5 December 2018; editorial decision 4 January 2019; accepted 17 January 2019.

Theory suggests that signal honesty may be maintained by differential costs for high and low quality individuals. For signals that mediate social interactions, costs can arise from the way that a signal changes the subsequent social environment via receiver responses. These receiver-dependent costs may be linked with individual quality through variation in resilience to environmental and social stress. Here, we imposed stressful conditions on female tree swallows (*Tachycineta bicolor*) by attaching groups of feathers during incubation to decrease flight efficiency and maneuverability. We simultaneously monitored social interactions using an RFID network that allowed us to track the identity of every individual that visited each nest for the entire season. Before treatments, plumage coloration was correlated with baseline and stress-induced corticosterone. Relative to controls, experimentally challenged females were more likely to abandon their nest during incubation. Overall, females with brighter white breasts were less likely to abandon, but this pattern was only significant under stressful conditions. In addition to being more resilient, brighter females received more unique visitors at their nest-box and tended to make more visits to other active nests. In contrast, dorsal coloration did not reliably predict abandonment or social interactions. Taken together, our results suggest that females differ in their resilience to stress and that these differences are signaled by plumage brightness, which is in turn correlated with the frequency of social interactions. While we do not document direct costs of social interaction, our results are consistent with models of signal honesty based on receiver-dependent costs.

Key words: glucocorticoids, plumage signals, social feedback.

INTRODUCTION

Models of signal evolution often focus on the costs of signal production as the key to understanding how signal honesty and between-individual variation in signals are maintained (e.g., Zahavi 1975; Hamilton and Zuk 1982; Folstad and Karter 1992). These receiver-independent production costs are undoubtedly important in many cases, but some costs—and all benefits—of signals are receiver-dependent and derived from their use in a social environment (Safran et al. 2008; Cornwallis and Uller 2010; Vitousek et al. 2014; Webster et al. 2017). Thus, the social environment that an individual inhabits plays a critical role in determining the fitness value of a given signaling phenotype. However, signals themselves can alter the social environment that an animal experiences by changing the frequency and nature of social interactions, which may be costly (Møller 1987; Dey et al. 2014; Webster et al. 2017). Furthermore, social experiences can result in physiological changes

that ultimately influence the production and maintenance of signals (Dey et al. 2014); thus, there is the potential for dynamic feedback between physiology, signals, and the social environment (Vitousek et al. 2013; Levin et al. 2016; Tibbetts et al. 2016; Crocker-Buta and Leary 2018; Levin et al. 2018). Under this scenario, mechanisms that link physiology, signals, and the ability to withstand challenges have the potential to generate the integrated suites of traits associated with honest signaling.

One mechanism that might link individual quality and the ability to tolerate receiver-dependent costs is the glucocorticoid stress response. Glucocorticoids can both organize behavioral phenotypes on various timescales (Wingfield et al. 1998; Spencer and Verhulst 2007; Romero et al. 2009) and can be influenced by the social environment (Creel et al. 2013; Dantzer et al. 2017). For example, experimental dosing with corticosterone can alter investment in breeding and parental behavior (Spée et al. 2011; Ouyang et al. 2015; Perez et al. 2016; Taff et al. 2018; Vitousek et al. 2018b). At the same time, repeated social stressors can alter the glucocorticoid response to subsequent stressors (Brunton and Russell 2010). Despite the costs associated with repeated or long-term activation

Address correspondence to C.C. Taff. E-mail: cct63@cornell.edu

of the stress response system, an appropriate response to challenging conditions is adaptive and is associated with the ability to survive and reproduce in challenging environmental or social conditions (Wingfield 2003). Thus, resilience to social stress and environmental stress may share common mechanistic underpinnings and both are likely to be influenced by the social environment that arises as a product of signal expression (Husak and Moore 2008; Killen et al. 2013). In the case of repeated social interactions mediated by signals, differential costs could arise through a combination of a difference in absolute costs per encounter or through differences in the frequency of encounters. For example, some individuals might be better at coping with repeated brief increases in glucocorticoids without suffering long-term costs (Taff et al. 2018).

Here, we studied plumage coloration, social activity, and resilience to experimental stressors in a free-living population of tree swallows (*Tachycineta bicolor*). Each of these aspects of tree swallow biology has been studied extensively in isolation, but they have not been linked previously. Both baseline and stress-induced corticosterone have been associated with reproductive success and survival (Bonier et al. 2011; Vitousek et al. 2018a), although these relationships are often inconsistent and context dependent (Bonier et al. 2009; Madliger and Love 2016). In this population, we previously found that females that exhibited low baseline corticosterone during periods of challenging conditions and a robust stress response had higher seasonal reproductive success (Vitousek et al. 2018a).

Male and female tree swallows have striking white ventral feathers and iridescent blue/green dorsal feathers. In previous studies at other locations, dorsal coloration has been identified as a sexually selected signal and variation in dorsal coloration is related to age (Bitton and Dawson 2008; Bentz and Siefferman 2013), extra-pair mating (Bitton et al. 2007; Van Wijk et al. 2016; Whittingham and Dunn 2016), coordination of male and female parental care (Dakin et al. 2016), nestling quality (Bitton et al. 2008; Berzins and Dawson 2016), and some immune measures (Beck et al. 2015). Much less is known about variation in the white ventral plumage (but see Beck et al. 2015). Finally, tree swallows are highly social throughout the year. Outside of the breeding season they form large foraging flocks and roost communally. In the breeding season, aggregations of individuals often nest at high densities and forage communally, such that conspecific social interactions are extensive year round. In populations with limited nest sites, aggression and nest-box takeovers are common and aggressive interactions can result in injury or death to adults and nestlings (Leffelaar and Robertson 1985; Robertson and Stutchbury 1988). Moreover, simulated conspecific box visitations prompt aggressive responses from females and more aggressive females are better at competing for limited boxes (Rosvall 2008). Importantly, several recent studies suggest that plumage characteristics mediate the intensity or frequency of these aggressive interactions; females with immature plumage receive less direct aggression (Coady and Dawson 2013) and experimentally enhancing female dorsal coloration alters nest site retention, reproductive success, and nestling quality (Berzins and Dawson 2016, 2018). While direct receiver-dependent costs have not yet been conclusively demonstrated, these patterns are most consistent with the idea that signals play a role in mediating social interactions that result in costs associated with the level of aggression or frequency of interaction.

Although tree swallow social interactions are common, they have previously been difficult to measure quantitatively because males and females look similar, and because the species breeds at high densities without exclusive territories and is not amenable to easy

individual identification with binoculars. These social interactions occur frequently on and around the nest-box in the breeding season, but due to the difficulty in tracking individuals, relatively little is known about these interaction patterns (but see Lombardo 1986; Lombardo 1987a; Lombardo 1987b). We used a novel approach to characterize social activity by installing a network of radio frequency identification (RFID) readers at each active nest-box in the population so that all visits by any previously captured bird to any active nest were recorded for the entire breeding season. We also measured feather coloration and the glucocorticoid response to standardized handling stress.

Finally, we exposed females to feather restraint treatments that simulated the effect of feather clipping, but that could be reversed after the treatment period. Feather clipping experiments have been used extensively in tree swallows and serve to increase flight and foraging costs for these obligate aerial insectivores (Winkler and Allen 1995). Previous wing-clipping experiments of various intensity and at different times of the breeding season have resulted in reduced condition, smaller clutches, lower feeding rates, reduced over winter survival, and reduced humoral immunocompetence (Winkler and Allen 1995; Ardia and Clotfelter 2006; Hasselquist et al. 2001b). Based on prior work suggesting that brighter coloration is associated with individual quality, we predicted that brighter females would be more resilient to experimentally imposed stressors and would exhibit a more robust glucocorticoid stress response. We did not make any a priori prediction about how coloration would relate to social activity because several patterns were possible; for example, brighter females might be predicted to receive more nest visits because they are more attractive or a better source of public information or they might be predicted to receive fewer visits because they are more aggressive and better at excluding potential visitors.

METHODS

General field methods and experimental treatments

We studied tree swallows breeding in nest-boxes at 2 locations near Ithaca, NY (site one: 42.503°N, 76.437°W, 342 m elevation; site 2: 42.461°N, 76.368°W, 516 m elevation). Nest-boxes at these sites are identically constructed with a 12.7 by 12.7 cm floor and 25.4 cm height; boxes are all oriented so that the entrance hole faces east, and placed at a uniform spacing of 20 m apart. Habitat across the box locations was very similar within each site. As in previous studies at this site (Winkler et al. 2013; Vitousek et al. 2018a), boxes were monitored for nest construction every other day starting in the first week of May. This allowed us to record clutch initiation, start of incubation, and dates of nest abandonment or failure to within 1 day. We checked nests every day when hatching was expected, allowing us to identify the exact date of hatching. We installed radio frequency identification (RFID) units on each box on the fourth day of incubation (see below). We intended to also collect data on nestling mass, physiology, and fledging success, but a late bout of cold weather from 6 to 12 June caused almost complete nest failure across all treatment groups in this year (only 7 out of 40 focal nests successfully fledged any young). The cold weather coincided with peak nestling growth, when nestlings are most susceptible to inclement conditions. Most nestlings at our focal nests had not been sampled before the cold snap; therefore, our analyses focus on adult females only.

Females were initially captured at the nest-box on incubation day 6 or 7. At this capture, we collected feathers to characterize coloration (see below), took morphological measurements (head + bill length, flat wing length, and mass), and collected 3 small blood samples (~70 μL for 1st sample and ~30 μL for 2nd and 3rd samples) in heparinized hematocrit tubes by brachial venipuncture to characterize the corticosterone response to handling stress. The first blood sample was collected within 3 min of disturbance at the box. After handling, birds were held in cloth bags and a second blood sample was taken 30 min after the initial disturbance. As part of a separate study not reported here, we also measured the efficacy of negative feedback in each individual using a dexamethasone injection at 30 min, followed by a final blood sample taken 30 min after injection (Zimmer et al. 2019). All adults were captured between 7:00 AM and 10:00 AM to control for circadian variation in corticosterone secretion. Blood samples were stored on ice in the field for 1–3 h and then centrifuged at 3500 rpm for 5 min to separate plasma, which was stored at $-30\text{ }^{\circ}\text{C}$ until analysis. Each bird was banded with a unique USGS aluminum band and a celluloid color band with attached passive integrated transponder (PIT) tag encoding a 10 digit hexadecimal string (see below).

At the time of the first capture, each female was randomly assigned to the control treatment, or to 1 of 2 treatments designed to experimentally restrict foraging ability.

“Control” birds experienced only the procedures listed above. In the other 2 treatments, groups of 3 primary feathers were bound together on each wing using miniature plastic zip ties. Binding a group of feathers together prevented them from fully spreading and thus reduced the total surface area of the wing in a way similar to feather clipping procedures. The goal of these treatments was to reduce flight ability and thereby increase the cost of foraging. Tree swallows are obligate aerial insectivores and flight maneuverability is critical to effective foraging. This treatment was designed to mimic the effects of feather clipping, which has been widely used to impose flight costs in tree swallows and other birds (e.g., Winkler and Allen 1995; Hasselquist et al. 2001a), but in a way that allowed us to restore flight ability once the treatment period was concluded (see Senar et al. 2001 for a similar reversible treatment using feather binding in great tits, *Parus major*). Our treatment most likely impacted both flight efficiency (the energy required for a given distance of flight) and flight maneuverability (ability to make brief changes in direction needed to catch flying insects). In the “low feather restraint” (low FR) group we attached primary feathers 4-5-6 on each wing in a single bundle. In the “high feather restraint” (high FR) group, we attached primaries 2-3-4 and 6-7-8 on each wing in 2 separate bundles (Supplementary Figure S1 illustrates the effect of zip ties on wing area). The weight of each individual zip tie was 0.05 g. Thus, in addition to the decreased wing surface area, individuals in the low FR and high FR groups carried an added weight of 0.1 and 0.2 g, respectively.

We next recaptured females on day 12 or 13 of incubation. At this capture, we took a single blood sample to measure baseline corticosterone and weighed each bird before release. For the low FR and high FR treatments, zip ties were removed at this capture, so that the treatment period lasted for 5–7 days. Among birds that were recaptured after feather restraint, 3 had lost one of their zip-ties (1 low FR and 2 high FR); since these individuals still retained 1 (low FR) or 3 (high FR) zip ties and presumably had the full treatment for some time before losing one zip tie, we retained them in our analyses. Any females that were not recaptured would have lost their zip-ties during molt shortly after the

breeding season. Finally, we recaptured females 6–8 days after eggs had hatched. At this final capture, we again took a series of 3 blood samples as described above. We also attempted to capture males at each box when females were captured for the third time (males cannot be reliably caught until nestlings have hatched and provisioning begins). Males were treated in the same way as females at the initial capture (i.e., we collected feathers, morphological measurements, a stress series of blood samples, and attached aluminum bands plus a PIT tag). No treatments were assigned to males. To minimize disturbance, we only attempted to capture males during the hour while females were held, thus only 20 out of 40 males at the focal nests were equipped with PIT tags.

Coloration

We measured the color of feathers plucked from the center of the breast and the center of the back from each bird. Four feathers from each plumage patch were stacked on index cards and taped in place. Coloration was measured using an Ocean Optics USB 2000 spectrophotometer with PX-2 pulsed Xenon light source and WS-1 white standard in OceanView version 1.5.2 (Ocean Optics, Dunedin, FL). Each saved spectrum represented an average of 10 scans, with a box-car width of 20 nm and integration time of 60 ms; the fiber optic UV/VIS probe was positioned in a holder that blocked external light while maintaining a distance of 5 mm from the probe to measurement surface and was held at 90° to the feathers.

For each stack of feathers, 4 separate reflectance spectra were collected with the probe removed from the feathers between each reading and replaced haphazardly within the color patch. All feathers were measured by CCT. At the time of measurement, only identification numbers were associated with each feather sample and CCT was blind to treatment group and nest fate. Breast feathers were measured near the center of the white patch, which covered most of each feather. Back feathers have blue/green tips that vary in both size and color between individuals (in first year females in our population, these tips are absent and feathers are uniformly drab brown). We made measurements at the center of the blue/green patch or as close as possible to that location on feathers for which this patch was small or absent. The blue/green feathers of tree swallows also have an iridescent component to their color that has been related to breeding performance in other populations (Van Wijk et al. 2015, 2016); our spectrophotometer set up did not allow us to characterize variation in angular iridescence in this study.

Raw reflectance spectra were processed in R version 3.3.3 (R Core Development Team 2016) using the package “pavo” (Maia et al. 2013). For each plumage patch, we calculated mean total brightness (“B2” in the pavo package) as the average reflectance from 300 to 700 nm. We also calculated color measurements derived from avian visual models in tetrahedral color space to represent bird-perceived colors. For these measurements, we used the blue tit visual model provided by “pavo” because cone sensitivities for tree swallows have not been directly measured (as in Dakin et al. 2016). In tetrahedral color space, a color is described by the angles theta (θ) and phi (Φ), which together define the hue of a color; along with the vector r , which defines the saturation (Stoddard and Prum 2008). Because the maximum length of the vector is constrained by the hue, we also calculated r achieved (r_a), as the degree of saturation given the maximum possible saturation of the vector (Stoddard and Prum 2008). Several of these metrics were highly correlated with each other (see Supplementary Tables S1 and S2); in our main analysis, we focused on the back measurements θ and

r_{as} , which were recently identified as mediators of parental care in another population of tree swallows (Dakin et al. 2016) and on the mean brightness of the white breast, which was highly correlated with breast θ and Φ .

Social Activity

Social activity was quantified using a network of RFID units that were maintained at each active nest-box. Circular antennae were fastened around each entrance hole so that birds had to pass directly through an antenna to enter or exit the box. We programmed our RFID units to sample for PIT tags every second between 5 AM and 8 PM each day (as in Vitousek et al. 2018b). This time period was chosen to conserve battery power and because tree swallows are much less active during the night. At the first capture, each bird was fitted with a passive integrated transponder (PIT) tag attached to a color band; many returning birds already carried PIT tags from previous years allowing for behavioral data collection even before the date of first capture. Each PIT tag encoded a unique 10 digit hexadecimal string that allowed subsequent identification of that individual when it perched at a box entrance anywhere in our site. These systems are only powerful enough to detect PIT tags that are within a few centimeters of the antenna, so only individuals that were perched on the nest-box entrance or passing through the opening would have been recorded. Because tree swallows nearly always perch at the entrance hole before entering a box, we expect that our system captured almost every visit made by individuals at our site, though it is possible that some small percentage of visits were missed if individuals passed very quickly into and out of the box (Vitousek et al. 2018b). We also could not detect visits in which birds were chased off before landing at the entrance hole or when they perched on top of or hovered around the box; thus, our social interaction data represent a minimum estimate for the frequency of total interactions that occur around the nest-box. When nest-boxes were in use, we downloaded data and changed the battery at each box approximately every 5 days. In some cases, RFID units or batteries temporarily failed so that complete RFID records were not available for every box on every day of the study.

Unprocessed RFID records were stored as text files containing a list of PIT encounters, which each consisted of a unique PIT tag string and a timestamp associated with the box where the RFID unit was deployed. These raw files were combined and processed using a custom script in R. The full script and raw RFID files are available by request from CCT. Briefly, our script first performed data cleaning operations to filter out reads that did not match any known PIT tags and to correct reads that differed by only 1 or 2 digits from the known male or female breeders at a box (these represent cases where the hexadecimal digits are misread by the RFID unit). These criteria were conservative and it is possible they resulted in the removal of some real visits by nonbox owners; however, we expect that these cases were very rare. When considering all 138 unique PIT tags deployed in this study, there were a total of 9453 possible dyadic pairs and only 0.9% of those pairings had PIT tag strings that differed by just 1 or 2 digits. Because only a small subset of possible dyadic pairs actually visited each other in the study, we expect that the number of pairings that would have been affected by our criteria was substantially smaller. Remaining PIT tag records were merged with information about each banded bird in the population using a record of previously deployed PIT tag numbers. Records that did not match any known PIT tag strings were removed.

Next, we filtered out reads of the male and female pair breeding at the nest-box, which represented the majority of total reads. The remaining records represented cases where a known bird visited the box of another breeding pair in the population. Because our RFID units were programmed to poll the antenna every second, many of these visits generated a string of reads from a single visiting individual. The final step in our script filtered these strings of readings into unique visits by applying a time threshold, such that a sequence of reads with gaps of <120 s between each subsequent read was considered to be a single “visit” with a time duration calculated as the difference between the earliest and latest reading. This time threshold was somewhat arbitrary, but sensitivity analysis with thresholds ranging from 30 to 600 s yielded overall similar results. Further, most subsequent analyses were based on the number of unique visitors—rather than total visits; for metrics that describe patterns of unique visitors, the choice of a time threshold was irrelevant.

Using the final list of box visitations, we calculated several metrics for each focal female in the study. First, we calculated the total number of visits made to her nest-box by other tagged individuals in the population. Second, we calculated the number of visits made by unique individuals (i.e., each unique PIT tagged visitor counted only once). Each of these measures was calculated separately for male and female visitors along with the total for both sexes combined; we separated these visits by sex both because visits by males and females might have different social effects and because nearly all females were equipped with PIT tags while only a subset of males were. Finally, we also calculated the number of total trips made to other nest-boxes and the number of unique nest-boxes visited by each female.

One potential issue with the metrics described above is that the total amount of sampling effort varied considerably across dates and between boxes. The number of active RFID readers and total number of detected PIT tags differed throughout the season as breeding peaked and then waned and as more adults were given PIT tags (Supplementary Figure S2). Further, nests that failed early in the season had fewer days to accrue visits than did successful nests and visitation rates were higher overall later in the nesting cycle. As predicted, nests that had an RFID installed for more days had more total unique visitors over the course of the season (Supplementary Figure S3). To account for these sampling differences, we calculated the metrics described above (total visits, unique visits, total trips, and unique trips) separately for each day that an RFID unit was running on an active box and used these daily observations as the basis for our main analysis. Organizing the data in this way allowed us to explicitly model the relationship between social patterns and the day in the nesting cycle.

This study focused on 40 breeding females that received control, low FR, or high FR treatments; however, the study sites that we used also contained additional breeding pairs that were monitored in simultaneous, unrelated studies (51 additional breeding attempts occurred across the 2 sites). Each of those active boxes was also equipped with an RFID reader and adults were given PIT tags following procedures identical to those described here. A total of 138 unique PIT tags from known birds were identified in our network including the 40 females and 20 males that were tagged at focal nests in this study. We included trips to or visits from those boxes when characterizing social activity in this study because this provided a more complete picture of overall social activity at the sites.

Our RFID network recorded a total of 596 days of activity at our 40 focal boxes (15.7 ± 10.8 days per nest; range 1–37). When

combined with other active nests included in the larger RFID network at our site, we monitored a total of 1557 active nest days. In total, our network identified 2 433 455 PIT tag records, with 81 509 (3.3 %) of those records attributed to individuals visiting boxes at which they were not the owner. After applying the time threshold (see above)—and including only the 40 focal nests in this study—we identified a total of 2642 unique visits made to focal nests and 486 trips made by focal females to other boxes at the study sites.

Corticosterone

Stored plasma samples were used to measure corticosterone concentrations using an enzyme immunoassay (EIA) kit (DetectX Corticosterone, Arbor Assays: K014-H5). We validated this kit for tree swallows by testing for parallelism in serial dilutions of tree swallow plasma pools. When comparing samples run with different starting amounts of plasma, it is important to assess parallelism because the absolute volume of plasma used can impact the final measurements, resulting in a correlation between starting volume and corticosterone concentration (Stevenson and Purushothama 2014). Ideally, samples run with different starting concentrations should yield similar measurements after applying a dilution factor. There is no standardized test for parallelism, but a typical rule-of-thumb is that the CV of back-calculated concentrations from a serial dilution should be <25–30% (Stevenson and Purushothama 2014). Previous studies on birds have used the same EIA kit either directly on plasma with a dissociation reagent provided in the kit or after an alcohol extraction to separate corticosterone from plasma constituents. We initially tried the dissociation reagent at several concentrations on a serial dilution of plasma. Although the relative concentrations for samples using the same starting amount of plasma were reliable using this method, we found that samples run with different starting volumes did not yield consistent results. Thus, samples tested with different starting amounts of plasma could not be compared directly and absolute corticosterone concentrations were unreliable.

Given these limitations, we next attempted a triple ethyl-acetate extraction of corticosterone on 2 serial plasma dilutions. We tested 2 separate plasma pools with 2-fold serial dilutions ranging from 1:10 to 1:80 of plasma to assay buffer. The average CV when comparing calculated final corticosterone concentrations for these 2 pools was 4.1%; thus, the requirement for parallelism was satisfied (Stevenson and Purushothama 2014). Finally, we optimized the extraction procedure to maximize extraction efficiency by testing a variety of vortexing methods and durations. In our final extraction protocol, we 1) added 5 μL of plasma to 45 μL of assay buffer

and vortexed at 1800 rpm for 3 s, 2) added 250 μL of ethyl acetate and vortexed at 1800 rpm for 30 s, 3) allowed the samples to settle for 5 min before placing in a $-80\text{ }^\circ\text{C}$ freezer for ~ 10 min, 4) collected the liquid alcohol containing dissolved corticosterone into a new tube while leaving the frozen plasma undisturbed, 5) repeated steps (2)–(4) 2 additional times adding the alcohol to the same collection tube each time, 6) allowed the alcohol with corticosterone to completely evaporate overnight, 7) reconstituted the dried corticosterone in 125 μL of assay buffer with an additional 30 s of vortexing, and 8) proceeded with the ELISA following the manufacturer's protocol.

For samples included in this experiment, extraction efficiency—determined by adding aliquots spiked with a known amount of steroid standard to each plate—averaged 89.7%. Samples were extracted using a starting volume of 5 μL of plasma when possible. If samples had <5 μL of plasma, we extracted the maximum amount possible and then corrected for the dilution relative to the other samples. Extracted samples were run in duplicate on EIA plates along with a 9-standard curve. The lower limit of detection when starting with 5 μL of plasma was 0.8 ng/mL. Intraplate variation based on duplicate samples was 10.6% and interplate variation based on a plasma pool run across plates was 5.7%.

Data analysis

To determine the degree to which our treatments represented a challenge, we initially examined the main effects of our 2 treatment groups on nest abandonment and adult phenotype relative to the control group. The response variables that we examined were hatching success (yes or no), adult corticosterone (baseline and stress response), and adult mass. We next asked whether individual coloration predicted resilience—measured as hatching success—to the challenging treatment within each group. With hatching success as the outcome variable, we fit a set of binomial generalized linear models with coloration (breast or back), treatment, and a treatment by coloration interaction as potential predictors (Table 1). We compared these models using AICc scores.

After establishing the effects of stress treatments, we used RFID data to ask whether coloration predicted patterns of nest visitation and trips to other boxes. Total visit numbers were correlated with the number of days that RFIDs were active at each box (Supplementary Figure S2). The number of active RFID days at a nest was driven by early nest failure, which in turn was related to treatment group and to coloration. To avoid these confounds, we focused only on daily visitation patterns in our main analysis, but note that the qualitative patterns were similar regardless of

Table 1
Candidate models of nest abandonment during incubation

Predictors	Log Likelihood	K^a	ΔAICc	w_i
Treatment \times breast brightness	−15.87	6	0.00	0.49
Treatment + breast brightness	−18.80	4	0.41	0.40
Treatment \times back Φ	−18.37	6	5.00	0.04
Treatment	−22.94	3	6.18	0.02
Treatment \times back r_a	−19.10	6	6.45	0.02
Treatment + back r_a	−22.14	4	7.07	0.01
Intercept only	−25.98	1	7.70	0.01
Treatment + back Φ	−22.72	4	8.23	0.01

All models with interactions also include the main effects. Models are generalized linear models with a binomial distribution.

^a K = number of parameters.

whether daily social activity or pooled metrics for the whole season were used.

To analyze social activity, we used a candidate model approach with competing models compared by AICc scores. We fit a set of generalized linear mixed models in the R package “glmer” with a Poisson distribution and with response variables for the number of visits received or trips made on a daily basis for each nest. The candidate model set included an intercept only model, a treatment only model, and 9 models that included a color metric (breast brightness, back θ , or back r_a), a color metric and treatment, or a color metric by treatment interaction (Tables 3 and 5). All of these models also included a categorical predictor for breeding stage; the categories for breeding stage were 1) pretreatment, 2) during treatment, 3) between 2nd and 3rd capture, which included hatching to nestling day 7, and 4) late nestling stage. For trips to other boxes, only breeding stages 2, 3, and 4 were included because many females did not have PIT tags before their initial capture. Female identity was included as a random effect in each model to account for multiple days of observation at each nest. Day of the season was also included as a random effect in each model to capture variation associated with the number of PIT tags deployed and number of active RFID units in the site on each day of the season; this predictor also likely captured any variation in social activity patterns associated with factors that would impact the entire population, such as weather, food availability, etc.

We fit this same set of candidate models with visits to the box or trips to other boxes as the response variable. For visits, we fit the candidate set separately for the full dataset and for reduced datasets that included only visits by males or only visits by females. In all 3 cases, the best-supported model was the same and the sign and magnitude of effects were similar. Thus, we report only the results for the single analysis that included all visits (results of the separate male and female visit analyses are included in the Supplementary Tables S3–S6).

Finally, we asked whether coloration was related to the hormonal stress response by comparing pretreatment coloration to baseline corticosterone and to the magnitude of the corticosterone response to handling stress using simple linear regression. All analyses and figures were produced in R version 3.3.3 (R Core Development Team 2016).

RESULTS

Main effects of treatments

In total, 80% (12 of 15) of control nests and 69% (9 of 13) of low FR nests hatched, whereas only 33% (4 of 12) of high FR nests hatched. Relative to control females, hatching success was significantly lower for high FR females (Fisher’s exact test $P = 0.02$). Although low FR females had intermediate hatching success, they did not differ significantly from high FR or control females (Figure 1a; Fisher’s exact test low FR vs. control: $P = 0.67$; low FR vs. high FR: $P = 0.12$). We were able to use RFID records to determine if females returned to the nest after the first capture when treatments were applied. Of the 15 nests that did not hatch, 7 abandonments occurred immediately after treatments. Immediate abandonment appeared to be more common in the high FR group (immediate abandonments in control group: 1; low FR group: 1; high FR group: 5), but sample sizes were too small to reliably test for statistical differences between immediate and delayed abandonment by group. Females that did not abandon immediately continued to return to the box for 1–11 days before abandonment.

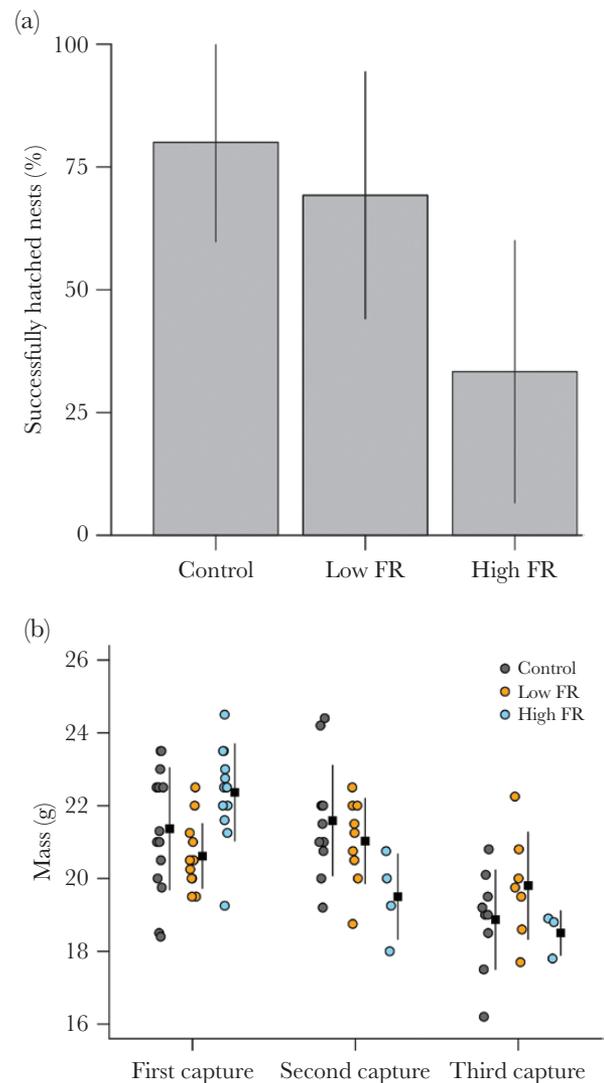


Figure 1

Main effects of feather restraint treatments. (a) Percent of nests in which at least one egg hatched by treatment group with 95% binomial confidence intervals. Hatching success was reduced in high FR treatments relative to controls. (b) Female mass at each capture by treatment group. Colored circles show raw data, while black squares and lines show group means \pm 1 SD. Controlling for initial mass, females in the high FR group lost more mass than controls between the first and second capture, but did not differ at third capture.

Although females were assigned randomly to treatment groups, females in the low FR had significantly lower mass than high FR females at the initial capture (one-way ANOVA, $df = 37$, $F = 5.2$, $P = 0.02$; Tukey HSD comparing high FR and low FR $P = 0.01$; Figure 1b). Given this difference, we tested for treatment effects on mass at the second and third capture using models that included both treatment and initial (i.e., first capture) mass. In these models, the high FR group tended to lose more mass between the first and second capture than the control group (full model $F_{3,21} = 2.68$, $P = 0.07$; high FR vs. control $t = -2.68$, $P = 0.01$; Figure 1b), but no other differences were detected between groups. Treatment groups did not differ in baseline corticosterone or stress induced corticosterone at any time point (1-way ANOVAs, all $P > 0.12$).

Coloration and the impact of flight restraint treatments

With hatching success as a binomial response variable, there was strong support for models that included either a breast brightness by treatment interaction (model weight = 0.49) or breast brightness and treatment as main effects with no interaction (model weight = 0.40). No other model in the set received substantial support (Table 1; ΔAIC_c for all other models > 4.9). In the best-fit model, restraint treatments reduced the likelihood of hatching, while greater bib brightness was associated with increased likelihood of hatching. The interaction was driven by a stronger effect of breast brightness in the high FR and, especially, low FR treatment groups, though the direction of the effect was consistent across all 3 groups (Figure 2; Table 2).

Coloration and social activity

The best-supported model for the number of unique visitors that females received at their nest included the breast brightness by treatment interaction (Table 3; model weight 0.77). All other models in our candidate set had ΔAIC_c values >4. Overall, breast brightness was associated with increased levels of nest visitation (Figure 3a; Table 3; $P < 0.001$), but the relationship differed by treatment (brightness by treatment interaction: likelihood ratio test: $\chi^2 = 11.4$, $P = 0.003$). In the control group, brighter breast feathers were associated with more nest visits, but there was little to no relationship in the low FR and high FR groups (Table 4). In addition to treatment and coloration, number of visits varied by breeding stage with the most visits occurring in the late nestling stage (Table 4).

For trips made to other boxes, the best-supported model also included the breast brightness by treatment interaction, but there was considerable uncertainty in model selection and 3 other models had ΔAIC_c values <2 (Table 5; model weight for best-supported model 0.24). In the best-fit model, females in the low FR and high FR treatments made fewer trips to other boxes than did control females (Table 6). It is important to note that, despite the model uncertainty, the reduction in trips made for low FR and—to a lesser

extent—high FR females was supported in the top 6 models with a combined model weight of 0.88. Thus, this main effect of treatment was robust, while the interaction with brightness was not as clear. Overall, breast brightness was positively correlated with trips to other boxes (Figure 3b), but the effect was only weakly supported and was driven primarily by a weak association in the low FR and marginally nonsignificant association in the high FR groups (Table 6). The number of trips made also varied by breeding stage with fewer trips made in the hatching to early nestling stage than at other times (Table 6).

Coloration and corticosterone

Of the color measurements that we considered, only back r_a was related to baseline corticosterone. Birds with more saturated dorsal coloration (higher r_a values) had lower baseline corticosterone (Table 7; Figure 4a). Only breast brightness was related to the magnitude of the stress response, with brighter white birds demonstrating a greater increase in corticosterone after handling (Table 7; Figure 4b).

DISCUSSION

We found that female tree swallows with brighter white breasts mounted a stronger physiological response to handling stress, engaged in more extensive social interactions at nest-boxes, and were more resilient to experimentally imposed flight costs. The marginal costs of similar feather restraint treatments differed for bright and dull females: brighter females were more capable of dealing with these challenges and therefore suffered a smaller marginal cost on hatching success from the same restraint. Moreover, the stressful conditions imposed by feather restraint weakened the relationship between feather brightness and social interactions, suggesting that plumage characteristics and behavior interact to produce the social environment that individuals experience. Taken together, our results suggest that stress resilience, social activity, and plumage signals form an integrated phenotype in tree swallows. Plumage reliably signals the ability to cope with an experimental challenge and is correlated with social activity. These results are consistent with the idea that receiver-dependent costs associated with bearing bright plumage play a role in the maintenance of honesty in this signal. It is important to note, however, that our study does not demonstrate direct costs of social interactions or describe the nature of these interactions and the correlations that we observed are not sufficient to conclusively demonstrate a role for receiver-dependent costs in maintaining honesty in this system. Further experimental work will be needed to elucidate the causal links generating the patterns that we observed.

Status signaling models generally predict that individuals don't "cheat" in signal production because bearing a signal that is too elaborate would result in intolerably high receiver-dependent costs (Møller 1987; Webster et al. 2017). Thus, signals themselves should be correlated with some relevant attribute of the individual (e.g., competitive ability, condition, genotype). We found that aspects of pretreatment coloration were related to both baseline and stress induced corticosterone. Furthermore, our experimental feather restraint demonstrated that females with brighter white breasts were less likely to abandon their clutch under stressful conditions. Interestingly, this relationship between color and reproductive performance was only apparent after feather restraint and was not detectable in the control group. We also detected some evidence that the relationship between coloration and social activity differed

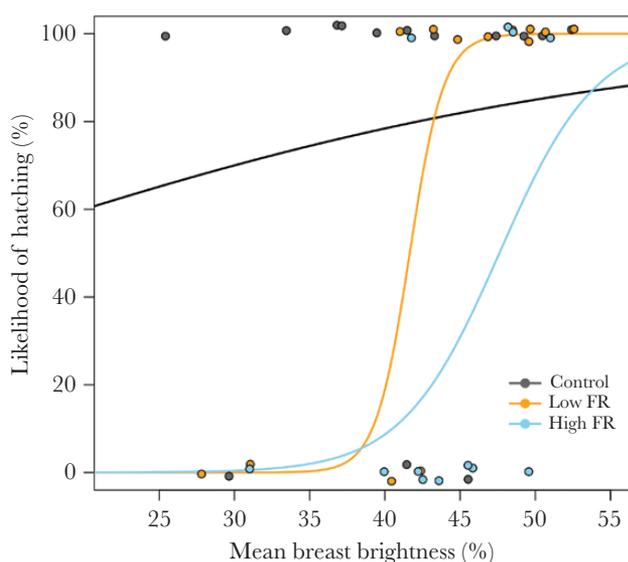


Figure 2 Likelihood of at least one egg hatching in a nest by treatment group and female breast brightness. Points show raw data and lines illustrate the predicted likelihood based on the best supported model in Table 1.

Table 2
Best-fit models for nest abandonment during incubation

Predictor	Estimate	SE	χ^2	<i>P</i>
Treatment × breast brightness ($\Delta\text{AIC}_c = 0$; $w_i = 0.49$)				
Intercept	-0.48	3.4		
Treatment			6.10	0.047
Low FR	-36.24	30.8		
High FR	-14.24	10.6		
Breast brightness	0.04	0.08	8.27	0.004
Treatment × brightness			5.85	0.054
Low FR × brightness	0.84	0.73		
High FR × brightness	0.26	0.23		
Treatment + breast brightness ($\Delta\text{AIC}_c = 0.41$; $w_i = 0.40$)				
Intercept	-5.41	2.7		
Treatment			6.10	0.047
Low FR	-1.24	1.1		
High FR	-3.12	1.2		
Breast brightness	0.17	0.07	8.27	0.004

Table 3
Candidate models for number of unique daily visits by females

Predictors	Log likelihood	<i>K</i>	ΔAIC_c	w_i
Treatment × breast brightness	-723.10	11	0.00	0.77
Breast brightness	-729.29	7	4.12	0.10
Back Φ	-729.82	7	5.17	0.06
Treatment + breast brightness	-728.78	9	7.21	0.02
Back r_a	-730.97	7	7.49	0.02
Treatment	-730.17	8	7.93	0.01
Treatment + back Φ	-729.37	9	8.39	0.01
Treatment * back Φ	-727.66	11	9.12	0.01
Treatment + back r_a	-730.07	9	9.80	0.01
Treatment * back r_a	-730.01	11	13.82	0.00
Intercept only	-763.49	3	64.38	0.00

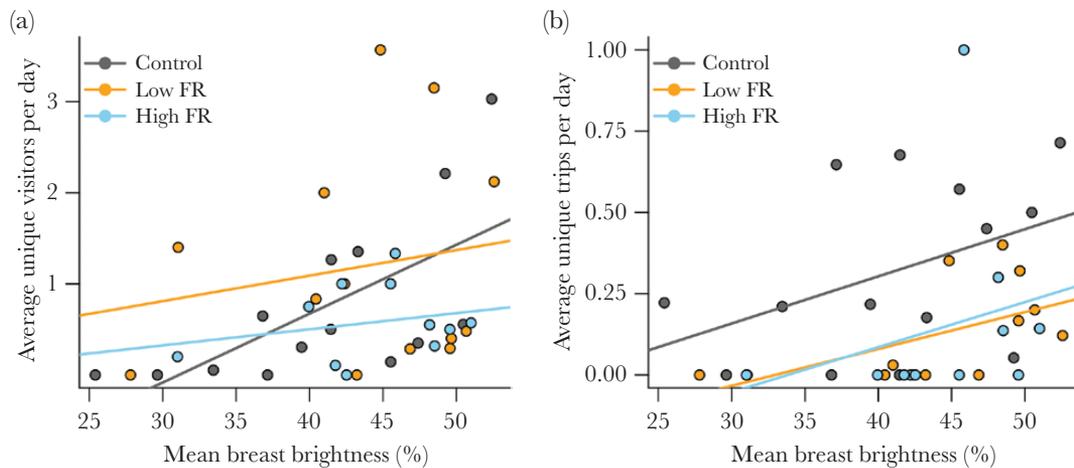
Models are GLMMs with a Poisson distribution and include nest identity and day of the season as random effects. Models that list interactions also include the main effects of each predictor. All models also include breeding stage as a predictor (see text).

across the treatment groups and that treatments themselves influenced visitation rates. Feather restraint resulted in a lower overall rate of trips to other nest-boxes for handicapped females; at the focal nest-box, the overall rate of visits did not differ by treatment, but the relationship between breast brightness and visitation rate was disrupted by feather restraint. This suggests that the relationship between coloration and social activity is context dependent and perhaps mediated by altered behavior under benign or stressful conditions. Most previous work on coloration in tree swallows has focused on the blue/green back coloration (Bentz and Siefferman 2013; Berzins and Dawson 2016; Dakin et al. 2016) and several studies suggest that this plumage patch is related to extrapair paternity (Bitton et al. 2007; Van Wijk et al. 2016; Whittingham and Dunn 2016). However, one previous study did find that females with brighter white breasts and their offspring performed better in immunological assays (Beck et al. 2015). We did not measure paternity in this study, but it is possible that the blue/green plumage patch is most important in mediating mate choice while the white plumage patch is more important in mediating aggressive interactions and signaling female quality, although the 2 functions are clearly not mutually exclusive for either plumage patch.

That repeated social interactions might function in maintaining signal honesty via the feedback of social experience controlled, in part, by the signaling phenotype itself is a classic idea in behavioral

ecology (Møller 1987). However, strong evidence for this hypothesis is relatively rare (Dey et al. 2014; Tibbetts et al. 2016; Webster et al. 2017). One reason for this is that in social species, every individual social interaction likely has a very small effect on fitness and the effects of signals on socially mediated performance may only arise through the cumulative experience of many repeated social interactions. Thus, demonstrating links between signals and the social environment depends on the ability to accurately quantify complex and diffuse social experiences for each individual in a population. Tree swallows are one of the best-studied species of birds outside of the laboratory. Although visits to other boxes have been documented previously (Lombardo 1986, 1987b; Wischhoff et al. 2015), the extent of social visitation to other boxes that we uncovered was surprising and was made possible only because of the novel use of RFID technology at each active nest-box. This technology has been used to quantify social interaction networks at central food resources (Adelman et al. 2015; Farine et al. 2015; Farine and Sheldon, 2015; Psorakis et al. 2015; Ellwood et al. 2017), but has not been used nearly as extensively in quantifying social interactions at nest locations, which may have quite different functions than social associations within a foraging flock.

One disadvantage of the RFID approach that we employed is that it is impossible to determine exactly what types of social interactions (if any) occurred at each nest-box visit that we logged. For

**Figure 3**

Relationship between mean breast brightness and the average number of unique visits received per day at the nest-box (a) or unique trips made to other nest-boxes (b) for each treatment group. Best-fit lines of each treatment group are included for illustration, but note that the main analysis was based on daily observations and included covariates and random effects not shown here.

Table 4**Best-fit model for unique number of daily nest visits**

Predictor	Estimate	SE	ζ	<i>P</i>
Treatment × breast brightness ($\Delta AIC_c = 0$; $w_i = 0.77$)				
Intercept (control, pre-T)	-1.04	0.35	-3.0	<0.01
Stage: during-T	-0.13	0.27	-0.47	0.64
Stage: hatching	0.21	0.28	0.74	0.46
Stage: nestlings	1.46	0.30	4.94	<0.001
Breast brightness	0.83	0.24	3.44	<0.001
Treatment: low FR	0.27	0.32	0.87	0.39
Treatment: high FR	-0.05	0.37	-0.14	0.89
Low FR × brightness	-1.18	0.32	-3.64	<0.001
High FR × brightness	-0.62	0.44	-1.40	0.16

The control treatment at the pretreatment breeding stage is coded as the reference value.

example, in some populations, tree swallows compete so aggressively over cavities that adults and nestlings can be killed as a result (Leffelaar and Robertson 1985; Robertson and Stutchbury 1988), but they also have high rates of extrapair paternity (Winkler et al. 2011). In some other cavity breeding species, conspecific visits are known to occur frequently and are thought to be associated with prospecting and the acquisition of public information (Doligez et al. 2004a, 2004b; Thomson et al. 2013). Our approach cannot easily distinguish between these different types of social interactions that might all occur at the cavity entrance. We make a general assumption that PIT tag records captured at the nest-box are indicative of overall social interactivity and several recent experimental studies do demonstrate that females often engage in aggressive interactions at the nest-box and that the intensity of aggression can differ according to female plumage (Rosvall 2008; Coady and Dawson 2013). One recent study in barn swallows (*Hirundo rustica*) used an alternative technology—proximity data loggers—to quantify social interaction networks (Levin et al. 2016). That technology has the advantage of estimating proximity for specific pairs of individuals and of capturing interactions even if they occur away from the nest-box. However, the method is more invasive and costly and, at present, battery and technical limitations mean it cannot be used to capture datasets on the same temporal scale as the RFID

network that we employed. Moreover, that method does not isolate the specific social context in which interactions may have occurred, whereas our method is restricted to interactions in a known context (at the nest-box). Going forward, there are opportunities to use these and other emerging technologies in complementary ways to improve our understanding of feedback between the social environment, signals, and physiology.

Our study establishes a suite of correlated phenotypic traits in tree swallows and suggests that there may be dynamic feedback between internal physiology, signals, and social activity in this system. A major challenge in studying dynamic feedback under natural conditions is that both physiology and social behavior are highly flexible traits; thus, how individuals respond flexibly to different environmental challenges may be more informative than single time point measures. Our results also suggest that the nature of these links may be context dependent. Assuming that flexible hormonal and social traits are causally linked, it is still unclear to what extent causality is bidirectional. Ideally, we would collect continuous measures on both behavior and physiology to characterize within individual reaction norms in each of these labile traits and ask how each responds to perturbations in the other (Dingemanse and Dochtermann 2013; Taff and Vitousek 2016). At present, this is easier for behavioral traits, as new technologies, such as the RFID

Table 5
Candidate models for number of daily trips made to other boxes

Predictors	Log likelihood	K	ΔAIC_c	w_i
Treatment \times breast brightness	-288.25	10	0.00	0.24
Treatment	-291.55	7	0.41	0.20
Treatment + breast brightness	-290.55	8	0.46	0.19
Treatment + back r_a	-291.03	8	1.43	0.12
Treatment + back Φ	-291.51	8	2.38	0.07
Treatment \times back r_a	-289.64	10	2.78	0.06
Back r_a	-293.82	6	2.90	0.06
Back Φ	-294.80	6	4.87	0.02
Breast brightness	-294.95	6	5.15	0.02
Treatment \times back Φ	-291.30	10	6.10	0.01
Intercept only	-314.27	3	37.70	0.00

Models are GLMMs with a Poisson distribution and include nest identity and day of the season as random effects. Models that list interactions also include the main effects of each predictor. All models also include breeding stage as a predictor (see text).

Table 6
Best-fit models ($\Delta\text{AIC}_c < 2$) for number of daily trips made to other boxes.

Predictor	Estimate	SE	ζ	P
Treatment \times breast brightness ($\Delta\text{AIC}_c = 0$; $w_i = 0.24$)				
Intercept (control, dur-T)	-1.73	0.47	-3.72	<0.001
Stage: hatching	-2.13	0.47	-4.53	<0.001
Stage: nestlings	0.36	0.32	1.11	0.27
Breast brightness	0.05	0.25	0.21	0.83
Treatment: low FR	-1.93	0.61	-3.17	<0.01
Treatment: high FR	-1.22	0.71	-1.72	0.09
Low FR \times brightness	1.12	1.06	1.06	0.29
High FR \times brightness	1.32	0.75	1.77	0.08
Treatment ($\Delta\text{AIC}_c = 0.41$; $w_i = 0.20$)				
Intercept (control, dur-T)	-1.78	0.46	-3.89	<0.001
Stage: hatching	-2.15	0.47	-4.55	<0.001
Stage: nestlings	0.45	0.32	1.41	0.16
Treatment: low FR	-1.32	0.48	-2.74	<0.01
Treatment: high FR	-0.83	0.55	-1.51	0.13
Treatment + breast brightness ($\Delta\text{AIC}_c = 0.46$; $w_i = 0.19$)				
Intercept (control, dur-T)	-1.65	0.46	-3.59	<0.001
Stage: hatching	-2.14	0.47	-4.55	<0.001
Stage: nestlings	0.43	0.32	1.34	0.18
Breast brightness	0.31	0.23	1.37	0.17
Treatment: low FR	-1.54	0.51	-3.00	<0.01
Treatment: high FR	-1.04	0.58	-1.81	0.07
Treatment + back r_a ($\Delta\text{AIC}_c = 1.43$; $w_i = 0.12$)				
Intercept (control, dur-T)	-1.84	0.46	-3.99	<0.001
Stage: hatching	-2.15	0.47	-4.57	<0.001
Stage: nestlings	0.47	0.32	1.47	0.14
Back r_a	-0.22	0.22	-1.00	0.32
Treatment: low FR	-1.16	0.50	-2.30	0.02
Treatment: high FR	-0.88	0.56	-1.58	0.11

The control treatment at the treatment (incubation) breeding stage is coded as the intercept.

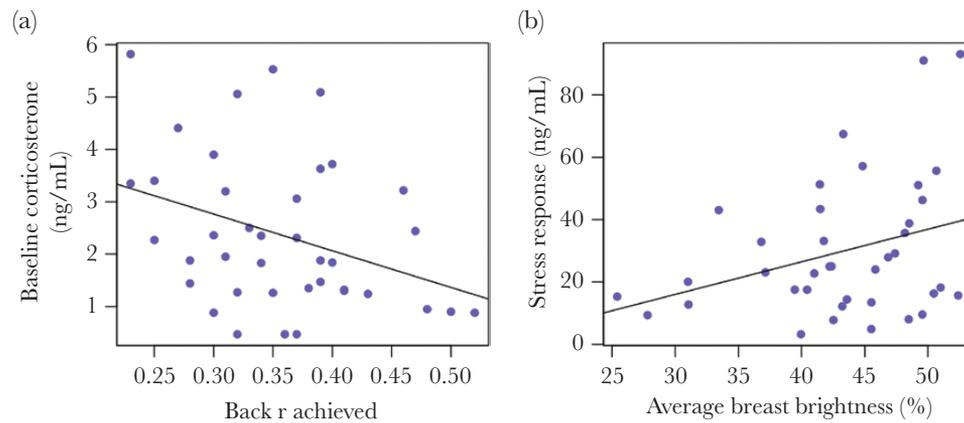
network employed here, have made it possible to capture massive behavioral datasets from free-living animals (Krause et al. 2013). Equivalent technologies are not available for most physiological measures under natural conditions, but there are some promising possibilities (Taff and Vitousek 2016).

A great deal of work in the last 3 decades has focused on uncovering the proximate mechanisms responsible for producing signaling traits that vary between individuals (Folstad and Karter 1992; Alonso-Alvarez et al. 2008; Taff and Freeman-Gallant 2014). While these studies have advanced our understanding of how signals are

produced and maintained, they largely adopt a unidirectional view of signal production in which mechanisms produce signals, which then influence reproductive success. There is growing appreciation for the fact that selection on signals can depend on dynamic feedback loops between physiology and behavior and that the signal itself plays a role in generating this selective environment (Cornwallis and Uller 2010; Safran et al. 2010; Vitousek et al. 2014). A full understanding of how selection operates on phenotypic traits that modify the selective environment that the organism experiences will require a bidirectional view of causal pathways.

Table 7**Pearson correlations between color measures and baseline or stress-induced corticosterone before treatments**

Plumage measurement	Baseline corticosterone			Stress-induced corticosterone			
	<i>r</i>	<i>t</i>	<i>P</i>	<i>r</i>	<i>t</i>	<i>P</i>	
Breast brightness	-0.03	-0.2	0.85	0.32	2.05	<0.05	
Back Φ	-0.09	-0.5	0.59	-0.03	-0.21	0.84	
Back r_a	-0.35	-2.28	0.03	0.04	0.23	0.82	

**Figure 4**Relationship between (a) back r_a and baseline corticosterone and (b) mean breast brightness and stress-induced corticosterone at initial capture.

SUPPLEMENTARY MATERIAL

Supplementary data are available at *Behavioral Ecology* online.

FUNDING

C.C.T. was funded by a postdoctoral fellowship from the Cornell Lab of Ornithology. This research was developed with funding provided by National Science Foundation (IOS 1457151) and Defense Advanced Research Projects Agency (D17AP00033). The views, opinions and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

We thank David Winkler for establishing and maintaining the tree swallow research population at Cornell's Experimental Ponds and the many assistants who helped with data collection in the field and the lab. The Vitousek lab and Webster lab provided helpful discussion and feedback on the article.

Data Accessibility: Analyses reported in this article can be reproduced using the data provided by Taff et al. (2019).

Handling editor: Andrea Griffin

REFERENCES

Adelman JS, Moyers SC, Farine DR, Hawley DM. 2015. Feeder use predicts both acquisition and transmission of a contagious pathogen in a North American songbird. *Proc Biol Sci.* 282:20151429.

Alonso-Alvarez C, Pérez-Rodríguez L, Mateo R, Chastel O, Viñuela J. 2008. The oxidation handicap hypothesis and the carotenoid allocation trade-off. *J Evol Biol.* 21:1789–1797.

Ardia D, Clotfelter E. 2006. Individual quality and age affect responses to an energetic constraint in a cavity-nesting bird. *Behav Ecol.* 18:259–266.

Beck ML, Hopkins WA, Hawley DM. 2015. Relationships among plumage coloration, blood selenium concentrations and immune responses of adult and nestling tree swallows. *J Exp Biol.* 218:3415–3424.

Bentz AB, Siefferman L. 2013. Age-dependent relationships between coloration and reproduction in a species exhibiting delayed plumage maturation in females. *J Avian Biol.* 44:80–88.

Berzins LL, Dawson RD. 2016. Experimentally altered plumage brightness of female tree swallows: a test of the differential allocation hypothesis. *Behaviour.* 153:525–550.

Berzins LL, Dawson RD. 2018. Experimentally altered plumage brightness of female tree swallows (*Tachycineta bicolor*) influences nest site retention and reproductive success. *Can Med Assoc J.* 96:600–607.

Bitton P, Dawson R. 2008. Age-related differences in plumage characteristics of male tree swallows *Tachycineta bicolor*: hue and brightness signal different aspects of individual quality. *J Avian Biol.* 39:446–452.

Bitton P-P, Dawson RD, Ochs CL. 2008. Plumage characteristics, reproductive investment and assortative mating in tree swallows *Tachycineta bicolor*. *Behav Ecol Sociobiol.* 62:1543–1550.

Bitton P-P, O'Brien EL, Dawson RD. 2007. Plumage brightness and age predict extrapair fertilization success of male tree swallows, *Tachycineta bicolor*. *Anim Behav.* 74:1777–1784.

Bonier F, Moore IT, Martin PR, Robertson RJ. 2009. The relationship between fitness and baseline glucocorticoids in a passerine bird. *Gen Comp Endocrinol.* 163:208–213.

Bonier F, Moore IT, Robertson RJ. 2011. The stress of parenthood? Increased glucocorticoids in birds with experimentally enlarged broods. *Biol Lett.* 7:944–946.

Brunton PJ, Russell JA. 2010. Prenatal social stress in the rat programmes neuroendocrine and behavioural responses to stress in the adult offspring: sex-specific effects. *J Neuroendocrinol.* 22:258–271.

Coady CD, Dawson RD. 2013. Subadult plumage color of female tree swallows (*Tachycineta bicolor*) reduces conspecific aggression during the breeding season. *Wilson J Ornithol.* 125:348–357.

Cornwallis CK, Uller T. 2010. Towards an evolutionary ecology of sexual traits. *Trends Ecol Evol.* 25:145–152.

- Creel S, Dantzer B, Goymann W, Rubenstein DR, Boonstra R. 2013. The ecology of stress: effects of the social environment. *Funct Ecol.* 27:66–80.
- Crocker-Buta SP, Leary CJ. 2018. Bidirectionality of hormone-behavior relationships and satellite-caller dynamics in green treefrogs. *Behav Ecol.* 29:948–957. doi: 10.1093/beheco/ary1047
- Dakin R, Lendvai A, Ouyang J, Moore I, Bonier F. 2016. Plumage colour is associated with partner parental care in mutually ornamented tree swallows. *Anim Behav.* 111:111–118.
- Dantzer B, Bennett NC, Clutton-Brock TH. 2017. Social conflict and costs of cooperation in meerkats are reflected in measures of stress hormones. *Behav Ecol.* 28:1131–1141.
- Dey CJ, Dale J, Quinn JS. 2014. Manipulating the appearance of a badge of status causes changes in true badge expression. *Proc Biol Sci.* 281:20132680.
- Dingemanse NJ, Dochtermann NA. 2013. Quantifying individual variation in behaviour: mixed-effect modelling approaches. *J Anim Ecol.* 82:39–54.
- Doligez B, Pärt T, Danchin E. 2004a. Prospecting in the collared flycatcher: gathering public information for future breeding habitat selection? *Anim Behav.* 67:457–466.
- Doligez B, Pärt T, Danchin E, Clobert J, Gustafsson L. 2004b. Availability and use of public information and conspecific density for settlement decisions in the collared flycatcher. *J Anim Ecol.* 73:75–87.
- Ellwood SA, Newman C, Montgomery RA, Nicosia V, Buesching CD, Markham A, Mascolo C, Trigoni N, Pasztor B, Dyo V, et al. 2017. An active-radio-frequency-identification system capable of identifying collocations and social-structure: validation with a wild free-ranging animal. *Methods Ecol Evol.* 8:1822–1831.
- Farine DR, Aplin LM, Sheldon BC, Hoppitt W. 2015. Interspecific social networks promote information transmission in wild songbirds. *Proc Biol Sci.* 282:20142804.
- Farine DR, Sheldon BC. 2015. Selection for territory acquisition is modulated by social network structure in a wild songbird. *J Evol Biol.* 28:547–556.
- Folstad I, Karter AJ. 1992. Parasites, bright males and the immunocompetence handicap. *Am Nat.* 139:603–622.
- Hamilton WD, Zuk M. 1982. Heritable true fitness and bright birds: a role for parasites? *Science.* 218:384–387.
- Hasselquist D, Wasson M, Winkler D. 2001a. Humoral immunocompetence correlates with date of egg-laying and reflects work load in female tree swallows. *Behav Ecol.* 12:93–97.
- Hasselquist D, Wasson MF, Winkler DW. 2001b. Humoral immunocompetence correlates with date of egg-laying and reflects work load in female tree swallows. *Behav Ecol.* 12:93–97.
- Husak JF, Moore IT. 2008. Stress hormones and mate choice. *Trends Ecol Evol.* 23:532–534.
- Killen SS, Marras S, Metcalfe NB, McKenzie DJ, Domenici P. 2013. Environmental stressors alter relationships between physiology and behaviour. *Trends Ecol Evol.* 28:651–658.
- Krause J, Krause S, Arlinghaus R, Psorakis I, Roberts S, Rutz C. 2013. Reality mining of animal social systems. *Trends Ecol Evol.* 28:541–551.
- Leffelaar D, Robertson R. 1985. Nest usurpation and female competition for breeding opportunities by tree swallows. *Wilson Bull.* 97:221–224.
- Levin II, Fosdick BK, Tsunekage T, Aberle MA, Bergeon Burns CM, Hund AK, Safran RJ. 2018. Experimental manipulation of a signal trait reveals complex phenotype-behaviour coordination. *Sci Rep.* 8:15533.
- Levin II, Zonana DM, Fosdick BK, Song SJ, Knight R, Safran RJ. 2016. Stress response, gut microbial diversity and sexual signals correlate with social interactions. *Biol Lett.* 12:20160352.
- Lombardo M. 1986. Attendants at tree swallow nests. I. Are attendants helpers at the nest? *Condor.* 88:397–303.
- Lombardo M. 1987a. Attendants at tree swallow nests III. Parental responses to live and stuffed-model attendants. *Condor.* 89:768–778.
- Lombardo M. 1987b. Attendants at tree swallow nests. II. The exploratory-dispersal hypothesis. *Condor.* 89:138–149.
- Madliger CL, Love OP. 2016. Do baseline glucocorticoids simultaneously represent fitness and environmental quality in a declining aerial insectivore? *Oikos.* 125:1824–1837.
- Maia R, Eliason CM, Bitton P-P, Doucet SM, Shawkey MD. 2013. pavo: an R package for the analysis, visualization and organization of spectral data. *Methods Ecol Evol.* 4:906–913.
- Møller A. 1987. Variation in badge size in male house sparrows *Passer domesticus*: evidence for status signalling. *Anim Behav.* 35:1637–1644.
- Ouyang JQ, Lendvai Á, Dakin R, Domalik AD, Fasanello VJ, Vassallo BG, Haussmann MF, Moore IT, Bonier F. 2015. Weathering the storm: parental effort and experimental manipulation of stress hormones predict brood survival. *BMC Evol Biol.* 15:219.
- Perez EC, Mariette MM, Cochard P, Soulage CO, Griffith SC, Vignal C. 2016. Corticosterone triggers high-pitched nestlings' begging calls and affects parental behavior in the wild zebra finch. *Behav Ecol.* 27:1665–1675.
- Psorakis I, Voelkl B, Garroway C, Radersma R, Aplin L, Crates R, Culina A, Farine D, Firth J, Hinde C, et al. 2015. Inferring social structure from temporal data. *Behav Ecol Sociobiol.* 69:857–866.
- Robertson R, Stutchbury B. 1988. Experimental evidence for sexually selected infanticide in tree swallows. *Anim Behav.* 36:749–753.
- Romero LM, Dickens MJ, Cyr NE. 2009. The Reactive Scope Model—a new model integrating homeostasis, allostasis, and stress. *Horm Behav.* 55:375–389.
- Rosvall KA. 2008. Sexual selection on aggressiveness in females: evidence from an experimental test with tree swallows. *Anim Behav.* 75:1603–1610.
- Safran RJ, Adelman JS, McGraw KJ, Hau M. 2008. Sexual signal exaggeration affects physiological state in male barn swallows. *Curr Biol.* 18:R461–R462.
- Safran RJ, Vitousek MN, Hauber ME, Ghalambor CK. 2010. Sexual selection: a dynamic state of affairs. Response to the comments of Cornwallis and Uller in the article: towards an evolutionary ecology of sexual traits. *Trends Ecol Evol.* 25:429–430; author reply 430.
- Senar JC, Domènech J, Uribe F. 2001. Great tits (*Parus major*) reduce body mass in response to wing area reduction: a field experiment. *Behav Ecol.* 13:725–727.
- Spée M, Marchal L, Lazin D, Le Maho Y, Chastel O, Beaulieu M, Raclot T. 2011. Exogenous corticosterone and nest abandonment: a study in a long-lived bird, the Adélie penguin. *Horm Behav.* 60:362–370.
- Spencer KA, Verhulst S. 2007. Delayed behavioral effects of postnatal exposure to corticosterone in the zebra finch (*Taeniopygia guttata*). *Horm Behav.* 51:273–280.
- Stevenson LF, Purushothama S. 2014. Parallelism: considerations for the development, validation and implementation of PK and biomarker ligand-binding assays. *Bioanalysis.* 6:185–198.
- Stoddard MC, Prum RO. 2008. Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. *Am Nat.* 171:755–776.
- Taff CC, Freeman-Gallant CR. 2014. An experimental test of the testosterone mediated oxidation handicap hypothesis in a wild bird. *Horm Behav.* 66:276–282.
- Taff CC, Vitousek MN. 2016. Endocrine flexibility: optimizing phenotypes in a dynamic world? *Trends Ecol Evol.* 31:476–488.
- Taff CC, Zimmer C, Vitousek MN. 2018. Efficacy of negative feedback in the HPA axis predicts recovery from acute challenges. *Biol Lett.* 14:20180131.
- Taff CC, Zimmer C, Vitousek MN. 2019. Data from: achromatic plumage brightness predicts stress resilience and social interactions in tree swallows (*Tachycineta bicolor*). *Behav Ecol.* doi:10.5061/dryad.46mm022
- Thomson R, Sirkkiä P, Villers A, Laaksonen T. 2013. Temporal peaks in social information: prospectors investigate conspecific nests after a simulated predator visit. *Behav Ecol Sociobiol.* 67:905–911.
- Tibbetts EA, Crocker K, Huang ZY. 2016. Socially selected ornaments influence hormone titers of signalers and receivers. *Proc Natl Acad Sci U S A.* 113:8478–8483.
- Van Wijk S, Bélisle M, Garant D, Pelletier F. 2015. A reliable technique to quantify the individual variability of iridescence in birds. *J Avian Biol.* 47:227–234.
- Van Wijk S, Bourret A, Bélisle M, Garant D, Pelletier F. 2016. The influence of iridescent coloration directionality on male tree swallows' reproductive success at different breeding densities. *Behav Ecol Sociobiol.* 70:1557–1569.
- Vitousek MN, Stewart RA, Safran RJ. 2013. Female plumage colour influences seasonal oxidative damage and testosterone profiles in a songbird. *Biol Lett.* 9:20130539.
- Vitousek MN, Taff CC, Hallinger KK, Zimmer C, Winkler DW. 2018a. Hormones and fitness: evidence for trade-offs in glucocorticoid regulation across contexts. *Front Ecol Evol.* 6:1–14.
- Vitousek MN, Taff CC, Stedman J, Zimmer C, Ardia DR, Salzman TC, Winkler DW. 2018b. The lingering impact of stress: brief acute glucocorticoid exposure has sustained, dose-dependent effects on reproduction. *Proc Biol Sci.* 285:20180722.
- Vitousek M, Zonana D, Safran R. 2014. An integrative view of the signaling phenotype: dynamic links between signals, physiology, behavior and social context. *Curr Zool.* 60:739–754.

- Webster MS, Ligon RA, Leighton GM. 2017. Social costs are an underappreciated force for honest signaling in animal aggregations. *Anim Behav.* 143:167–176. doi:10.1016/j.anbehav.2017.1012.1006
- Whittingham LA, Dunn PO. 2016. Experimental evidence that brighter males sire more extra-pair young in tree swallows. *Mol Ecol.* 25:3706–3715.
- Wingfield JC. 2003. Control of behavioural strategies for capricious environments. *Anim Behav.* 66:807–816.
- Wingfield JC, Maney DL, Breuner CW, Jacobs JD, Lynn S, Ramenofsky M, Richardson RD. 1998. Ecological bases of hormone-behavior interactions: the “emergency life history stage”. *Am Zool.* 38:191–206.
- Winkler DW, Allen PE. 1995. Effects of handicapping on female condition and reproduction in tree swallows (*Tachycineta bicolor*). *Auk.* 112:737–747.
- Winkler DW, Hallinger KK, Ardia DR, Robertson RJ, Stutchbury BJ, Cohen RR. 2011. Tree Swallow (*Tachycineta bicolor*), version 2.0. In: Rodewald PG, editor. *The birds of North America*. Ithaca (NY): Cornell Lab of Ornithology.
- Winkler DW, Luo MK, Rakhimberdiev E. 2013. Temperature effects on food supply and chick mortality in tree swallows (*Tachycineta bicolor*). *Oecologia.* 173:129–138.
- Wischhoff U, Marques-Santos F, Ardia D, Roper J. 2015. White-rumped swallows prospect while they are actively nesting. *J Ethol.* 33:145–150.
- Zahavi A. 1975. Mate selection—a selection for a handicap. *J Theor Biol.* 53:205–214.
- Zimmer C, Taff CC, Ardia DR, Winkler DW, Vitousek MN, 2019. On again, off again: acute stress response and negative feedback together predict resilience to experimental stressors. *Functional Ecology.* doi: 10.1111/1365-2435.13281