

Negligible effects of blood sampling on reproductive performance and return rates of Tree Swallows

Sophia C. M. Orzechowski,^{1,2,3} J. Ryan Shipley,¹ Teresa M. Pegan,^{1,4} and David W. Winkler¹

¹Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York 14853, USA

Received 18 June 2018; accepted 8 November 2018

ABSTRACT. Blood sampling is a frequently used method of collecting genetic and physiological data in natural populations, and understanding the possible impact of blood sampling on individuals and populations is important, both for the welfare of study organisms and to avoid introducing bias into analyses using bled individuals. Most studies of birds have revealed minimal negative effects of blood sampling. However, Brown and Brown (2009. *Auk* 126: 853–861) found that blood sampling reduced return rates of Cliff Swallows (*Petrochelidon pyrrhonota*), suggesting that these results are not always generalizable and that swallows (Hirundinidae) may be particularly sensitive to blood sampling. We examined the possible effects of blood sampling on the reproductive performance of female Tree Swallows (*Tachycineta bicolor*) and the return rates of both females and offspring in a population in New York state. To reduce the chances of Type II error, we tested 15 possible effects of blood sampling on reproductive performance and return rates using generalized linear mixed models. Overall, our results suggest that blood sampling had few negative effects on within-year reproductive success or survival of adult females. The one exception was that bleeding nestlings had a negative effect on the number of young that fledged in broods of five or six nestlings. Bleeding did not negatively impact return rates of females or nestlings in our Tree Swallow population. Our results support those of other studies suggesting that blood sampling has few negative effects on within-year reproductive success or survival of adult females, and reiterate the importance of systematically testing for possible effects with datasets compiled over multiple years because such effects may be impossible to detect via direct monitoring during a single breeding season.

RESUMEN. Efectos despreciables de la toma de muestras de sangre en el desempeño reproductivo y las tasas de retorno de *Tachycineta bicolor*

La toma de muestras de sangre es un método frecuentemente usado para coleccionar datos genéticos y fisiológicos en poblaciones naturales y es importante entender el posible impacto de la toma de muestras de sangre sobre los individuos y las poblaciones, por ambas razones, el bienestar de los organismos de estudio y para evitar la introducción de sesgos a los análisis usando individuos sangrados. Muchos estudios en aves han revelado que los efectos negativos de la toma de muestras de sangre son mínimos. Sin embargo, Brown y Brown (2009. *Auk* 126: 853–861) encontraron que la toma de muestras de sangre reducía la tasa de retorno en *Petrochelidon pyrrhonota*, sugiriendo que estos resultados no son siempre generalizables y que las golondrinas (Hirundinidae) pueden ser particularmente sensibles a la toma de muestras de sangre. Examinamos los posibles efectos de la toma de muestras de sangre en el desempeño reproductivo de *Tachycineta bicolor* y las tasas de retorno en hembras y crías en una población en el estado de Nueva York. Para reducir los chances de error Tipo II, probamos 15 efectos posibles de la toma de muestras de sangre en el desempeño reproductivo y las tasas de retorno usando modelos lineales mixtos generalizados. En general, nuestros resultados sugieren que la toma de muestras de sangre tuvo pocos efectos negativos sobre el éxito reproductivo intra-anual o las tasas de supervivencia de las hembras adultas. La única excepción fue que el sangrado de los pichones tuvo un efecto negativo en el número de juveniles que salían del nido en las nidadas de cinco o seis pichones. El sangrado no tuvo ningún impacto negativo sobre las tasas de retorno de las hembras o de los pichones en nuestra población de *Tachycineta bicolor*. Nuestros resultados soportan esos otros estudios que sugieren que la toma de muestras de sangre tiene pocos efectos negativos sobre el éxito reproductivo intra-anual o sobre la supervivencia de hembras adultas y reiteran la importancia de comprobar sistemáticamente los posibles efectos en las bases de datos recopiladas a través de múltiples años por que dichos efectos pueden ser imposibles de detectar por medio de monitoreos durante una única temporada de reproducción.

²Corresponding author. Email: sorzechowski@ufl.edu

Current address:³Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611, USA.

⁴Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, USA.

Key words: bleeding, episodes of selection, experimental manipulation, fitness, investigator disturbance

One of the fundamental challenges when analyzing data is disentangling patterns from underlying biases. A potential influence on experimental results is investigator disturbance,

where the researcher and/or the experimental process have unintended effects on the study organism. These effects can vary in severity depending on experimental design, methods, and sensitivity of study organisms (Götmark 1992, Carney and Sydeman 1999, Blumstein 2006, Ibáñez-Álamo et al. 2012, Stein et al. 2017). However, quantifying these effects is essential because they can influence interpretation of results and can also impact the welfare of organisms, ranging from acute consequences to carryover effects on fitness and survival (O'Connor et al. 2014). Carryover effects of an experiment may not be apparent within the experimental timeframe and may only be detectable when the protocol or disturbance is recurrent.

With the advent of molecular biology and advances in vertebrate physiological ecology, blood sampling has become a standard part of the study protocol for many vertebrate populations, producing a wide array of information on parentage, mating systems, and metabolic, endocrine, and immuno-physiology. Several investigators have examined the possible effects of blood sampling on birds (Sheldon et al. 2008, Redmond and Murphy 2011, Guillemain et al. 2015, Bowers et al. 2016, Smith et al. 2016). Although most studies have revealed no detectable effect of bleeding on fitness, Brown and Brown (2009) found that Cliff Swallows (*Petrochelidon pyrrhonota*) that were bled were 21–33% less likely to return the following year. However, Voss et al. (2010) suggested caution in generalizing the strength of unintended experimental effects between different populations of even the same species because they are unlikely to be exposed to the same conditions. Therefore, determining the effect of experimental procedures such as blood sampling on any long-term study population to identify possible short- and long-term effects on fitness is important.

The fitness consequences of blood sampling are context-dependent, and the severity of any potential effect likely depends on when sampling occurs during the reproductive cycle. “Episodes of selection” (Arnold and Wade 1984) are based on major life-history stages in the reproductive cycle (such as hatching and fledging) and provide a useful framework for understanding how the fitness consequences of experimental procedures vary

throughout the reproductive cycle. For example, blood sampling is likely to have greater effects on the fitness of birds during incubation when they are more likely to abandon nests (Winkler 1991, Götmark 1992); in contrast, parents may be less likely to abandon nests during the nestling stage. However, if blood sampling negatively affects parents, sampling during the nestling period may have a negative impact on the number of young that fledge. This could potentially occur via a reduction in nestling provisioning rates if, for example, blood sampling affected the foraging efficiency of adults.

Tree Swallows (*Tachycineta bicolor*) have been widely studied as a model organism in vertebrate ecology (Jones 2003). Readily nesting in man-made nest boxes, these swallows can be encouraged to breed in large numbers at experimental sites, and their considerable robustness to disturbance during the nesting season has made it possible to study a broad range of topics (e.g., Jones 2003, Palacios et al. 2012, Stedman et al. 2017). Many such studies include blood sampling as part of the study protocols, and the possible effects of bleeding should be evaluated carefully. Toward that end, we analyzed the bleeding and reproductive performance records of a population of Tree Swallows near Ithaca, New York, that has been studied since 1986. Using this long-term dataset spanning 28 years (1986–2014), we examined the possible effects of sampling blood (1) on adult females during the incubation and nestling stages, (2) on nestlings, and (3) on the probability of females and nestlings returning to the study area in the subsequent breeding season. We investigated the effects of bleeding on the probability of any eggs hatching, number of eggs that hatched, probability of any nestlings fledging, and number of young fledged.

METHODS

All reproductive data are from the long-term experimental populations of Tree Swallows near Ithaca, New York. All sites were within 27.4 km of each other and are described in Winkler and Allen (1996).

Blood collection. Blood was sampled sporadically for phylogenetic (Sheldon and Winkler 1993) and physiological (Hasselquist

et al. 2001) studies in the earlier years of study, but regular sampling for a DNA archive for parentage studies began in 2002. Since that time, we have sampled blood from females during 2038 of 3688 (55%) nesting attempts. About 60% of bled females (1528 of 2038) were sampled during incubation, and 25% (510 of 2038) during the nestling period. The rest were bled either before or after the breeding season after being captured in mist-nets. About one-third of bled birds were sampled more than once during a breeding season, again depending on experimental agendas. Because males were more difficult to capture, fewer were bled and we focus our analyses on females.

Except for one study where jugular venipuncture was used (Palacios et al. 2011, 2012), we sampled blood from brachial veins using 27 gauge 1/2 inch- (1.27 cm) or 26 gauge 3/8 inch- (0.95 cm) sized needles. In our blood-collection protocol, we swabbed the sample site lightly with petroleum jelly and inserted a needle at a shallow angle ($< 20^\circ$). We collected blood in 80- μ L heparinized capillary tubes and either immediately stored samples in sodium dodecyl sulfate lysis buffer or on ice to be centrifuged later for plasma extraction. After sampling, we placed a small swab of cotton on the puncture site and applied pressure, inspecting the vein again to ensure that bleeding had ceased before releasing birds. Nestlings were sampled using the same protocol. For corticosterone

stress series on adults or nestlings (e.g., Stedman et al. 2017), we sampled blood from both wings due to the difficulty of collecting multiple samples from the same wing in a short period of time.

We recorded the volume of blood drawn in units of capillary tubes (1 tube = 80 μ L blood). If less than a tube was taken, we recorded the fractional amount to the nearest 0.25. Blood samples from females ranged in volume from 0.25 to 6 capillary tubes (mean = 1.0 ± 0.7 [SD] tube). The mean blood volume collected represented 0.4 ± 0.3 (SD) % of mean body mass of breeding Tree Swallows, which ranged from 17 to 22 g (we used 19.5 g for calculations; Winkler et al. 2011). This range is within the Ornithological Council's recommended 1% of body mass for blood volume taken from wild birds (Fair et al. 2010). Most of the largest volumes of blood (≥ 3 tubes) were sampled via jugular bleeding for a study in 2008 (Palacios et al. 2011). Blood samples from nestlings ranged in volume from 0.25 to 2 capillary tubes (mean = 0.6 ± 0.3 [SD] tube). The mean mass of nestlings that were bled was 16.4 ± 5.0 [SD] g (mean mass of nestlings in control nests was 17.9 ± 5.0 [SD] g). Mean nestling mass decreased slightly with increasing brood size for nests where nestlings were weighed at 8, 9, or 12 days old (Fig. 1). Mean blood volume collected represented 0.30 ± 0.16 (SD) % of the mean mass of bled nestlings, with 0.8% of the mean

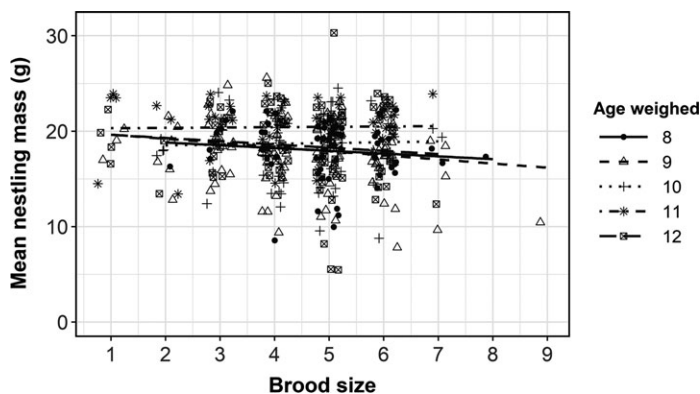


Fig. 1. Relationship between mean nestling mass and brood size at the time of banding/bleeding with regression lines for nests measured at the same age. The data are slightly jittered to enhance readability and consist of 410 records from Analysis 9 for which we had average nestling masses measured on days 8 to 12 post-hatching.

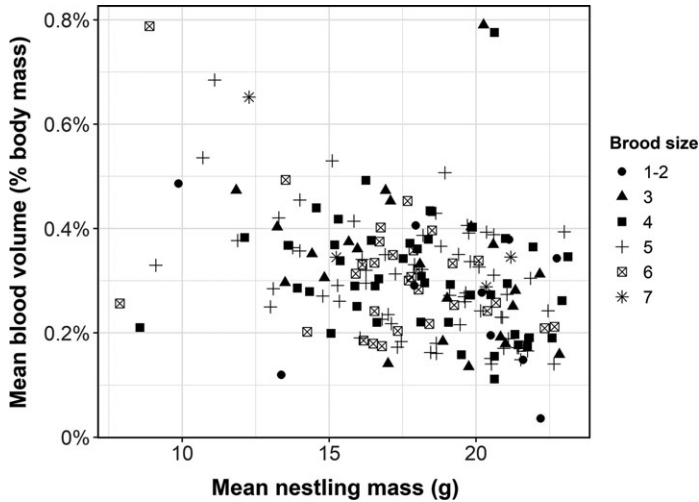


Fig. 2. Relationship between the average amount of blood taken from nestlings (in terms of percent body mass) and average nestling mass in nests where nestlings were bled. The different shapes indicate brood size at the time of sampling. The data consist of 178 records from Analysis 9 for which we had both average nestling masses and average amount of blood taken from nestlings.

nestling mass being the highest value for nests where we had information on both nestling mass and blood volume collected from nestlings (Fig. 2).

Data formatting and filtering. Our long-term dataset is useful for studying the effects of blood sampling on Tree Swallow fitness, but the experimental system was not explicitly designed to test those effects. We filtered our data to remove potentially confounding issues and ensure valid comparisons between birds at control (not bled) and experimental (bled) nests. We included only the first nest record for each female in the dataset to reduce age-related effects on reproductive output and remove the bias of variable bleeding status of individuals across different years. This removed all second nesting attempts and nesting attempts made by the same female in the same and all subsequent years. We excluded nests involved with brood manipulation experiments or experiments that involved the destructive collection of eggs or nestlings. We excluded nests where adults had either been fitted with backpack geolocator devices or some flight feathers had been experimentally shortened. We also excluded years where sample sizes of control and bled nests were highly mismatched. To ensure a fair comparison between nests where nestlings were bled

and control nests, we only included nest records where at least one nestling survived to be banded in control nests or banded and bled in experimental nests. We typically banded and bled nestlings between days 8–12 post-hatching (mean = 9.4 ± 2.3 [SD] d, calculated from Analysis 9 dataset). After this filtering process, low (< 2 or 3) and high (> 7 or 8) values of clutch and brood sizes were not represented in the experimental group in some analyses. Similarly, not all values of capture timing were represented in both groups after the filtering process described above. To be consistent, we chose not to filter these variables in any of the datasets for each analysis. We decided the parameter estimates should be based on all valid values of these variables, regardless of uncertainty at low or high values.

Our dataset included the band numbers of female Tree Swallows and relevant information about their first nest record at our study sites, including age, lay date, clutch size, hatch date, brood size, number of young that fledged, when and how many times females were captured and sampled for blood, whether and at what age nestlings were bled, and whether females or nestlings returned the following year. We aged adult females using either plumage (1-year-old females have

mostly brown upperparts) or previous banding data. We tested whether including females of unknown age (iridescent adults banded after year one) changed the overall conclusions of the analysis. Because it did not, we included those females in the final dataset. We standardized lay date by calculating the difference between a nest's lay date and the mean lay date during that year. The time when females were captured during incubation was the number of days after the date of clutch completion when females were captured (mean = 6.9 ± 3.0 [SD] d, calculated from Analysis 1 dataset in Table 1); during the nestling stage, the time of capture was the number of days after the hatch date (mean = 6.2 ± 4.8 [SD] d, calculated from Analysis 5 dataset in Table 1). Nestling age at the time of banding/bleeding was the number of days after the hatch date. Brood size in nestling-bled analyses was the number of nestlings alive at the time of banding/bleeding. For nests where nestlings were bled, the entire brood was bled in 81.6% of nests (222 of 272, calculated from Analysis 9 dataset in Table 1). In the remaining nests, one-half to four-fifths of nestlings were bled. We quantified the volume of blood taken from females during each episode of selection. For female-return analyses, we summed the total volume of blood taken from each female during the entire season, and the total number of times a female was bled throughout the season.

The strength of both long- and short-term effects can be highly context-dependent (e.g., Winkler 1991), so we structured our potential effects and analyses to test effects of blood sampling during distinct stages or episodes of selection. We identified four episodes of selection to evaluate separately (Fig. 3). We assessed the potential immediate effects of bleeding females on (1) hatching outcomes during the incubation stage, and (2) fledging outcomes during the nestling stage. We assessed (3) entire season outcomes in the form of female return rates. Tree Swallows with poor reproductive outcomes in a given breeding season are more likely to disperse, but, in general, only a small percentage of breeding females disperse and return rate is most strongly related to survival (Winkler et al. 2004). We separately assessed the immediate effects of bleeding nestlings on nestling stage outcomes, and effects of

bleeding nestlings on (4) offspring recruitment outcomes. Outcomes for each episode consisted of one to two response variables that could potentially have been affected by bleeding during each episode. To systematically evaluate consequences of blood sampling on each outcome, we formulated 15 questions and divided our dataset into subsets appropriate for each analysis (Table 1). Each subset contained nests where females or nestlings were bled, and a control group of nests where females were captured during the corresponding stage (incubation or nestling) and nestlings were banded, but no blood sampling occurred. Data subsets for incubation-stage analyses included only females that were first bled or first captured (controls) during incubation. Subsets for female nestling-stage analyses included only females first bled or first captured during the nestling period to control for potential short-term carryover effects of bleeding during incubation.

Data analysis. We tested the possible effects of blood sampling on reproductive metrics using generalized linear mixed models (GLMMs) and an information-theoretic approach. We constructed a small set of models where the blood term (either binary bleed status or blood sample volume) was added to a base model as a main effect, or as interacting with covariates when we had reasonable predictions regarding a potential interaction. Our base model included female age and lay date, variables long established as important predictors of avian reproductive metrics (Stutchbury and Robertson 1988, Winkler and Allen 1996, Hasselquist et al. 2001). We also included clutch size (brood size in nestling-stage analyses) to account for the influence of varying clutch sizes on our reproductive metrics.

We incorporated interactions between the blood term and clutch or brood size to represent the following predictions regarding potential effects. (1) In models of hatch probability, we predicted that the effects of blood sampling females, if present, would reduce hatch probability in small clutches (1–3 eggs) compared to small clutches in the control group. In other species, smaller clutches relative to the species' typical clutch size are deserted at higher frequencies (Wiggins et al. 1994, Székely et al. 1996), and desertion rates are influenced by disturbances (Götmark

Table 1. How datasets were subsetted for each analysis, and descriptions of response variables.

Analysis	N total	N bled	N control	Data subsetted by:	Response variable
(1) Did bleeding females during incubation affect whether or not ≥ 1 egg hatched? (Binomial)	846	262	584	Nests where females were bled during incubation and not before, and nests of females not bled before or during incubation from the same years	Hatched – Binary variable indicating whether any eggs in a nest hatched
(2) Did the amount of blood taken from females during incubation affect whether or not ≥ 1 egg hatched? (Binomial)	845	242	603	Same as (1), but nests and years where bled females had known quantities for their blood samples	Hatched – Binary variable indicating whether any eggs in a nest hatched
(3) Among nests where ≥ 1 egg hatched, did bleeding females during incubation affect how many eggs hatched? (Gaussian)	728	217	511	Same as (1), but only nests where at least one egg hatched and excluding years with few controls	Brood size – Maximum number of eggs that hatched in a nest
(4) Among nests where ≥ 1 egg hatched, did the amount of blood taken from females during incubation affect how many eggs hatched? (Gaussian)	703	192	511	Same as (2), but only nests where at least one egg hatched and excluding years with few controls	Brood size – Maximum number of eggs that hatched in a nest
(5) Did bleeding females during the nestling period affect whether the nest fledged ≥ 1 young? Binomial	167	62	105	Nests where females were first bled during the nestling stage and not before, and nests of females not bled before or during the nestling stage from the same years	Fledged – Binary variable indicating whether any young in a nest fledged
(6) Did the amount of blood taken from females during the nestling stage affect whether or not ≥ 1 young fledged? Binomial	158	53	105	Same as (5), but nests where bled females had known quantities for their blood samples	Fledged – Binary variable indicating whether any young in a nest fledged
(7) Among nests that fledged ≥ 1 young, did bleeding females during the nestling period affect how many young fledged? (Gaussian)	139	51	88	Same as (5), but only nests that fledged at least one young	Fledgenum – Number of young that fledged from a nest
(8) Among nests that fledged ≥ 1 young, did the amount of blood taken from females during the nestling period affect how many young fledged? (Gaussian)	121	41	80	Same as (6), but only nests that fledged at least one young	Fledgenum – Number of young that fledged from a nest

Table 1. Continued

Analysis	N total	N bled	N control	Data subsetted by:	Response variable
(9) Did whether or not nestlings were bled affect whether nests fledged ≥ 1 young? (Binomial)	767	272	495	All nests where nestlings were bled, and nests where nestlings were not bled from the same years	Fledged – Binary variable indicating whether any young in a nest fledged
(10) Among nests that fledged ≥ 1 young, did whether or not nestlings were bled affect how many young fledged? (Gaussian)	517	232	285	Same as (9), but only including nests where young fledged	Fledgenum – Number of young that fledged from a nest
(11) Did whether females were bled in their first breeding season affect whether they returned the following season? (Binomial)	656	265	391	Nests where young were fledged and females were first bled during incubation and were bled at least once in a season, and nests where young were fledged and females were not bled in a season from the same years	Returned – Binary variable indicating whether a female ever returned the following year
(12) Did the total number of blood samples taken from females affect whether they returned the following breeding season? (Binomial)	656	265	391	Nests where young fledged and females were first bled during incubation and bled at least once in a season, and nests where young fledged and females were not bled in a season from the same years	Returned – Binary variable indicating whether a female ever returned the following year
(13) Did the quantity of blood taken from females in a breeding season affect whether they returned the next breeding season? (Binomial)	611	220	391	Same as (11 and 12), but only nests where females were first bled during incubation and had known blood quantities summed across all blood sample events within the season (and same control group)	Returned – Binary variable indicating whether a female ever returned the following year
(14) Did bleeding females affect whether ≥ 1 young returned the following year? (Binomial)	561	253	308	Same as (11 and 12), but only nests where eggs hatched and had band dates in control nests and band/bled dates in bled nests	Youngreturned – Proportional response: number of young returning to any site the following year out of the number that fledged
(15) Did bleeding nestlings affect whether or not ≥ 1 young returned the next year? (Binomial)	440	146	294	Same as (14), but tallying number of nests where all nestlings were bled or not instead of females and removing years with few controls	Youngreturned – Proportional response: number of young returning to any site the following year out of the number that fledged

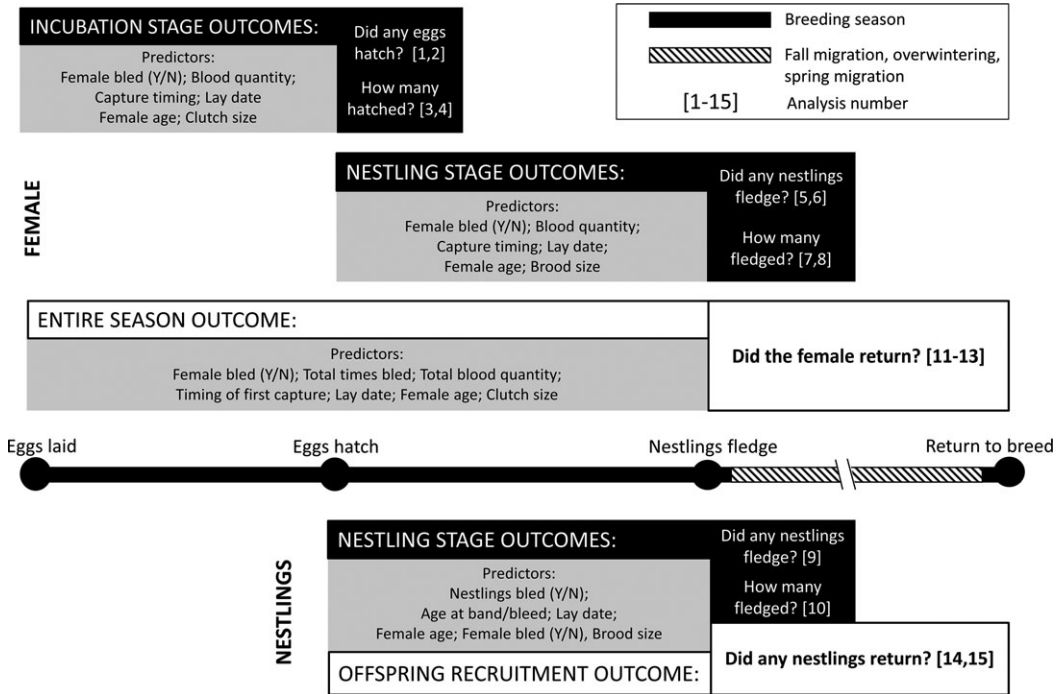


Fig. 3. A visual representation of the episodes of selection we examined, and the variables analyzed for each. Each box represents one of the four episodes of selection. Predictors are listed in gray, within-year reproductive outcomes in black, and survival outcomes in white. The corresponding numbered analyses for each episode are listed in brackets. The timeline in the middle shows where each episode lies in relation to nest phenology. Episodes related to females are shown above the timeline and the one related to nestlings is shown below.

1992). (2) In models of fledge number, we included an interaction between the bleed status of nestlings and brood size at the time of sampling because nestling mass in Tree Swallows tends to decrease with brood size due to lower per capita feeding rates (Murphy et al. 2000, Shutler et al. 2006). Nestlings from clutches smaller than the typical clutch size of 5 or 6 in our study system (80.9%, or 1436 of 1775 nests in our full dataset had clutches of 5 or 6) might be more resilient to blood sampling due to a weight advantage from higher per capita feeding rates. The potential effect of bleeding lighter nestlings in typical or larger-than-typical broods might reduce the number of nestlings that fledged compared to control nests (nestlings banded, but not bled). (3) We also included an interaction between female bleed status (or blood volume) and clutch size when modeling the number of eggs that hatched. For Tree Swallows, larger clutches require more days of

incubation (Shutler et al. 2006) and reduced hatching success for larger clutches has been reported for Barn Swallows (*Hirundo rustica*; Engstrand and Bryant 2002). We predicted that bleeding females with large clutches during incubation might reduce the number of eggs that hatched compared to large clutches in the control group. Similarly, parental provisioning rates are higher for larger broods of Tree Swallows (Murphy et al. 2000, Shutler et al. 2006, Hainstock et al. 2010). Females that work harder to provision larger broods may be particularly sensitive to the effects of bleeding. We therefore included an interaction between brood size and the blood term (female bleed status or blood volume). We also included an interaction between the blood term and capture timing in all models. For example, if sampling blood from females during incubation had an effect, we predicted that hatch success of females captured and bled early in incubation would be lower than

for early captures in the control (not bled) group. Disturbances early in the incubation stage generally have a greater effect on hatching success due to nest desertion, which is also more likely as the disturbance level increases (Winkler 1991, Götmark 1992, Verboven and Tinbergen 2002). Year was included as a random effect in every model to account for interannual variation in breeding success. We also included a null intercept-only model in each model set.

We used a binomial distribution for models of hatch, fledge, and return probabilities. In binomial models of the return rates of young Tree Swallows, we modeled the proportional response as the number of young in each nest that returned out of the total number that fledged. The ratios of the mean and variance of our dependent count variables (number of offspring hatched or fledged) were highly different from 1, violating the characteristics of a Poisson distribution. We opted to use models assuming a normal distribution of the response variables instead, which perform more optimally when count data violate Poisson assumptions (McDonald and White 2010).

We compared nested models to determine the importance of the blood term in each analysis. We ranked models using AIC_c , which contains a correction term for small sample sizes (our nestling stage datasets for female analyses were small because only 25% of records were of females first bled during the nestling stage) and converges to AIC as sample size increases (Burnham et al. 2011). We chose AIC_c to perform model selection to avoid Type II errors that might occur with a criterion like BIC, which has greater term penalization (Schwarz 1978) and may select more parsimonious models than AIC (Murtaugh 2009). If a nested model that differed by one additional term had a ΔAIC_c score within two units of the best model, we considered that term to likely be uninformative as per Arnold (2010). If a blood term was present in the model with the lowest ΔAIC_c score, we examined the effect size (standardized coefficient) and significance of the blood term and compared it to other covariates in the model. Finally, we used a pluralistic approach and directly tested, via null hypothesis testing, the importance of a blood term in the best model identified by AIC_c

(Stephens et al. 2005). We tested whether the blood term improved model fit using a Kenward–Roger conditional F -test for our Gaussian models, which is appropriate for linear mixed models fitted with restricted maximum likelihood and tends to be more robust than the likelihood ratio test (Luke 2017). Because GLMMs in lme4 are fitted with maximum likelihood, we used a likelihood ratio test (LRT) rather than an F -test for our binomial models. LRT statistics calculated on fixed effects tend to be more conservative than the default Wald P -values calculated in GLMM (Bolker et al. 2009). If the best model determined by AIC_c did not contain a blood term, we compared the best model with and without the addition of the bleeding variable as a main effect. All continuous predictor variables were standardized to enhance comparability of effects among covariates. The observed multicollinearity in our models was within acceptable limits because the variance inflation factors (VIF) were < 2.25 , below a suggested threshold of 10 (Dormann et al. 2013). All analyses were performed in R v.3.3.1 (R Core Team 2016). We ran mixed models in the lme4 package (Bates et al. 2015) and created graphs using ggplot2 and sjPlot (Wickham 2016, Lüdtke 2018). We ran Kenward–Roger F -tests in the pbkrtest package (Halekoh and Højsgaard 2014).

RESULTS

After filtering for the first record of each female, removing nests with problematic records (e.g., missing female band number) and nests with other experimental manipulations, our full dataset consisted of 1775 unique nest records from 2002 to 2014. Females were bled at some point during the breeding season at 1022 nests and, for the rest (753 nests), females were never bled. Nestlings were bled at 517 nests and, in the rest (1258 nests), nestlings were banded, but never bled.

Model-selection results. The full results of model selection can be found in Appendix S1. In most analyses (14 of 15), the best model did not include a blood sampling term (Table S1). In Analysis 1 (Did bleeding females [Y/N] during incubation affect hatch probability?), three models had ΔAIC_c scores less than 2, indicating model

uncertainty (Table S1). These models were nested and differed only in inclusion of a blood term either as a main effect or interacting with clutch size. Bleeding females during incubation was probably an uninformative variable (Arnold 2010), however, due to the observed uncertainty, we report the output of the full model with the bleed:clutch interaction term (within $\Delta\text{AIC}_c = 2$) to give estimates of the potential effects (Table 2). In Analyses 5 and 6 (Did bleeding females [Y/N] or amount of blood drawn during the nestling stage affect the probability of young fledging?), a nested model differing only in the inclusion of the blood sampling term was also within $\Delta\text{AIC}_c = 2$ of the best model, another indication that the blood term was likely uninformative. For both, we report the coefficients of the second-best model within $\Delta\text{AIC}_c = 2$ to provide an effect size of the blood term (Table 2). In Analysis 10 (Did bleeding nestlings [Y/N] affect the number that fledged?), the best model contained an interaction between brood size and bleed status of nestlings and was 4.93 AIC_c units lower than the next model (Table S1), providing strong evidence that bleeding nestlings in larger broods reduced the number of fledglings compared to large control broods (Table 2, Fig. 4).

None of the best models contained an interaction between the blood term and capture timing in any analysis. Capture timing of females was included as a main effect in the best model of hatch probability in Analyses 1 and 2 (Did bleeding females [Y/N] or amount of blood drawn during incubation affect hatch probability?), and also in the best model of fledge probability in Analyses 5 and 6 (Did bleeding females [Y/N] or amount of blood drawn affect fledge probability?; Table S1). Capture timing of females during incubation was positively associated with hatch probability (Analyses 1 and 2, Table 2) and this was confounded with daily hatch probabilities increasing for females captured closer to the hatch date. During the nestling stage, capture timing of females was similarly positively associated with the probability of young fledging (Analyses 5 and 6, Table 2). This was also confounded by increasing daily survival-to-fledge probabilities for females captured closer to the date of fledging, with nestlings more likely to survive over the

shorter intervening time interval. In Analysis 10 (Did bleeding nestlings [Y/N] affect the number that fledged?), nestling age at the time of banding/bleeding was negatively related to the number of young that fledged, although the effect was small (nestling age $\beta = -0.11 \pm 0.05$ [SE]; Table 2).

Significance of the blood term. We tested whether inclusion of the blood term significantly improved fit in the best model identified by AIC_c . In one of 15 analyses, the blood term improved model fit at a significance level of $\alpha = 0.05$ (Analysis 10; Did bleeding chicks [Y/N] affect the number of nestlings that fledged?; Table 2). The interaction between brood and nestling bleed status was significant in the best model ($F_{1,509} = 12.7$, $P = 0.0004$; brood:nestlingbleedY $\beta = -0.31 \pm 0.09$ [SE]; nestlingbleedY $\beta = -0.40 \pm 0.09$ [SE]; Table 2, Fig. 4). The mean number of nestlings that fledged differed between bled and control nests with broods of 5 (Welch two sample t -test, $t_{100} = -3.0$, $P = 0.006$, mean difference in number of fledglings in bled nests = -0.44 ± 0.31 nestlings) and 6 ($t_{40} = -4.0$, $P = 0.0007$, mean difference in number of fledglings in bled nests = -1.32 ± 0.75 nestlings). In broods of 7, the difference was not significant ($t_6 = -0.3$, $P = 0.80$, mean difference in number of fledglings in bled nests = -0.4 ± 3.8 nestlings). In Analysis 1 (Did bleeding females [Y/N] affect hatch probability?), the likelihood ratio tests did not support inclusion of the blood term (Table 2), despite the results of AIC_c model selection indicating some uncertainty regarding the best model (Table S1).

DISCUSSION

The size of our full dataset allowed us to filter out potentially biasing data and make valid comparisons between control and bled nests in each analysis. For most tests, blood-sampling variables were inconsequential predictors of within-year reproductive metrics. Similarly, bleeding did not negatively impact return rates of females or nestlings in our Tree Swallow population. In contrast, Brown and Brown (2009) found that bleeding reduced return rates of Cliff Swallows. Our results suggest that such latent effects of blood sampling depend on the species and study system. Similarly, Voss et al. (2010)

Table 2. The results of testing the significance of including the blood term in the best model identified by AIC_c (italicized), and regression coefficients of the underlined model.

Analysis	Model ± blood term	Test statistic	P	Variable	β	SE	t/z	
(1) Did bleeding females during incubation affect whether or not ≥ 1 egg hatched? (Binomial)	$\text{hatched} \sim * + \text{clutch} + \text{capture}^\dagger + \text{bled}^\dagger$	$\chi^2_1; 1.7$	0.20	Intercept	2.08	0.25	8.3	
	<u>$\text{hatched} \sim * + \text{clutch} * \text{bled}^\dagger + \text{capture}^\dagger$</u>			Lay date	0.03	0.11	0.3	
	$\text{hatched} \sim * + \text{clutch} + \text{capture}^\dagger + \text{bled}^\dagger$	$\chi^2_1; 0.9$	0.34	Female age	0.05	0.11	0.4	
				Clutch	0.65	0.14	4.6	
				Capture [†]	0.44	0.12	3.6	
				<u>$\text{hatched} \sim * + \text{clutch} * \text{capture}^\dagger + \text{bled}^\dagger$</u>	BledY [†]	-0.32	0.24	-1.3
(2) Did the amount of blood taken from females during incubation affect whether or not ≥ 1 egg hatched? (Binomial)	$\text{hatched} \sim * + \text{clutch} + \text{capture}^\dagger + \text{blood.quant}^\dagger$	$\chi^2_1; 0.2$	0.69	Clutch:BledY [†]	-0.29	0.22	-1.3	
	<u>$\text{hatched} \sim * + \text{clutch} + \text{capture}^\dagger + \text{blood.quant}^\dagger$</u>			Intercept	1.98	0.19	10.2	
	$\text{broodsize} \sim * + \text{clutch}$	$F_{1,720}; 1.7$	0.19	Laydate	0.05	0.11	0.4	
				Female age	0.004	0.12	0.03	
				Clutch	0.67	0.12	5.7	
				Capture [†]	0.47	0.12	3.9	
(3) Among nests that hatched ≥ 1 egg, did bleeding females affect how many eggs hatched? (Gaussian)	$\text{broodsize} \sim * + \text{clutch}$	$F_{1,720}; 1.7$	0.19	Blood quantity [†]	0.04	0.11	0.4	
	<u>$\text{broodsize} \sim * + \text{clutch} + \text{bled}^\dagger$</u>			Intercept	4.94	0.10	52.2	
	$\text{broodsize} \sim * + \text{clutch} + \text{bled}^\dagger$	$F_{1,638}; 0.1$	0.77	Lay date	0.05	0.04	1.3	
				Female age	0.08	0.04	2.0	
				Clutch	0.61	0.04	16.2	
				<u>$\text{broodsize} \sim * + \text{clutch} + \text{blood.quant}^\dagger$</u>	BledY [†]	-0.1	0.08	-1.3
(4) Among nests where the amount of blood taken from females during incubation affect how many eggs hatched? (Gaussian)	$\text{broodsize} \sim * + \text{clutch}$	$F_{1,638}; 0.1$	0.77	Intercept	4.95	0.09	53.5	
	<u>$\text{broodsize} \sim * + \text{clutch} + \text{blood.quant}^\dagger$</u>			Lay date	0.06	0.04	1.6	
	$\text{fledged} \sim * + \text{brood} + \text{capture}^\dagger + \text{bled}^\dagger$	$\chi^2_1; 0.2$	0.68	Female age	0.07	0.04	1.9	
				Clutch	0.61	0.04	16.4	
				<u>$\text{fledged} \sim * + \text{brood} + \text{capture}^\dagger + \text{bled}^\dagger$</u>	Blood quantity [†]	-0.01	0.04	-0.3
				(5) Did bleeding females during the nestling period affect whether the nest fledged ≥ 1 young? (Binomial)	$\text{fledged} \sim * + \text{brood} + \text{capture}^\dagger + \text{bled}^\dagger$	$\chi^2_1; 0.2$	0.68	Intercept
<u>$\text{fledged} \sim * + \text{brood} + \text{capture}^\dagger + \text{bled}^\dagger$</u>	Lay date	0.63	0.3		2.1			
$\text{fledged} \sim * + \text{brood} + \text{capture}^\dagger + \text{bled}^\dagger$	$\chi^2_1; 0.2$	0.68	Female age		0.30	0.25	1.2	
			Brood		0.07	0.22	0.3	
			Capture [†]		1.00	0.31	3.3	
			<u>$\text{fledged} \sim * + \text{brood} + \text{capture}^\dagger + \text{bled}^\dagger$</u>		BledY [†]	-0.19	0.48	-0.4

Table 2. Continued

Analysis	Model \pm blood term	Test statistic	P	Variable	β	SE	t/z
(6) Did the amount of blood taken from females during the nestling stage affect whether or not ≥ 1 young fledged? (Binomial)	$\frac{\text{fledged} - * + \text{brood} + \text{capture}^\ddagger}{\text{fledged} - * + \text{brood} + \text{capture}^\ddagger + \text{blood.quant}^\ddagger}$	χ^2_1 ; 0.6	0.46	Intercept	1.88	0.28	6.7
				Lay date	0.58	0.30	1.9
				Female age	0.31	0.26	1.2
				Brood	0.07	0.23	0.3
(7) Among nests where ≥ 1 young fledged, did bleeding females during the nestling period affect how many young fledged? (Gaussian)	$\frac{\text{fledged} - * + \text{brood}}{\text{fledged} - * + \text{brood} + \text{bled}^\ddagger}$	$F_{1,132}$; 0.7	0.4	Intercept	4.14	0.18	23.4
				Lay date	0.05	0.10	0.5
				Female age	0.11	0.10	1.1
				Brood	0.82	0.10	8.4
(8) Among nests where ≥ 1 young fledged, did the amount of blood taken from females during the nestling period affect how many young fledged? (Gaussian)	$\frac{\text{fledged} - * + \text{brood}}{\text{fledged} - * + \text{brood} + \text{blood.quant}^\ddagger}$	$F_{1,116}$; 0.1	0.73	Intercept	4.08	0.19	22.0
				Lay date	0.05	0.11	0.4
				Female age	0.10	0.11	0.9
				Brood	0.82	0.11	7.5
(9) Did whether or not nestlings were bled affect whether nests fledged ≥ 1 young? (Binomial)	$\frac{\text{fledged} - 1}{\text{fledged} - \text{nestlingbled}}$	χ^2_1 ; 0.3	0.61	Blood quantity [‡]	0.04	0.11	0.4
				Intercept	2.26	0.15	14.7
				NestlingbledY	-0.13	0.25	-0.5

Table 2. Continued

Analysis	Model ± blood term	Test statistic	P	Variable	β	SE	t/z
(10) Among nests where ≥ 1 young fledged, did whether or not nestlings were bled affect how many young fledged? (Gaussian)	fledgnum ~* + brood + nestling.age + nestlingbled	$F_{1,509}$	0.0004	Intercept	4.2	0.07	59.5
	<u>fledgnum ~* + brood * nestlingbled + nestling.age *</u>	12.7		Lay date	0.09	0.05	2.1
	fledgnum ~ + brood + nestling.age			Female age	0.03	0.05	0.6
	+ brood + nestling.age + nestlingbled			Brood	1.07	0.07	18.1
(11) Did whether females were bled in their first breeding season affect whether they returned the following season? (Binomial)	returned ~ 1	$F_{1,387}$	< 0.0001	Nestling age	-0.11	0.05	-2.2
	<u>returned ~ bled</u>	18.2		NestlingbledY	-0.40	0.09	-4.5
		χ^2_1 ; 0.03	0.86	Brood:nestlingbledY	-0.31	0.09	-3.6
(12) Did the total number of blood samples taken from females affect whether they returned the following breeding season? (Binomial)	returned ~ 1			Intercept	-0.003	0.12	-0.03
	<u>returned ~ total.blednum</u>	χ^2_1 ; 0.3	0.56	BledY	0.03	0.13	0.2
(13) Did the quantity of blood taken from females in a breeding season affect whether they returned the next breeding season? (Binomial)	returned ~ 1			Intercept	0.009	0.10	0.1
	<u>returned ~ total.bloodquant</u>	χ^2_1 ; 0.02	0.89	Total times bled	0.05	0.08	0.6
(14) Did bleeding females affect whether ≥ 1 young returned the following year? (Binomial)	youngreturned ~* + brood			Intercept	0.01	0.10	0.1
	<u>youngreturned ~* + brood + bled</u>	χ^2_1 ; 0.2	0.65	Total blood quantity	0.012	0.08	0.1
				Intercept	-2.8	0.15	-18.6
				Lay date	-0.51	0.12	-4.3
			Female age	-0.15	0.10	-1.6	
			Brood	-0.17	0.10	-1.8	
			BledY	-0.08	0.18	-0.5	

Table 2. Continued

Analysis	Model ± blood term	Test statistic	P	Variable	β	SE	t/z
(15) Did bleeding nestlings affect whether or not ≥ 1 young returned the next year? (Binomial)	<i>youngreturned</i> + <i>brood</i> <u>youngreturned</u> + <u>brood</u> + <u>nestlingbleed</u>	χ^2_1 ; 0.2	0.64	Intercept Lay date Female age Brood NestlingbleedY	-2.86 -0.54 -0.16 -0.17 0.12	0.19 0.14 0.12 0.11 0.25	-18.2 -3.8 -1.3 -1.5 0.5

*Lay date + female age.

†Incubation stage.

‡Nestling stage.

Bolded variables indicate significance at $P < 0.05$; italicized models are the best according to the results of AIC_c model selection. The reported coefficients are from the underlined models that, in most cases, include all covariates in the best model plus the blood term as a main effect. The only exceptions are Analysis 10 where the best model was reported because it included a blood term interaction, and Analysis 1 where the coefficients of the fullest model < AIC_c = 2 are reported because of model uncertainty.

suggested that context is important in understanding why Brown and Brown (2009) found an effect of bleeding on return rates of Cliff Swallows; this species breeds in an arid environment where fluid loss from bleeding may have more detrimental consequences than in less water-constrained environments, and Cliff Swallows in the Brown and Brown (2009) study also suffered from unusually intense ectoparasitism. Voss et al. (2010) also suggested that the effects of bleeding may be influenced by mass and diet, and that side-effects of sampling blood from brachial veins, such as hematomas, may have a greater effect on aerial insectivores because they rely exclusively on flight to obtain food. However, like Cliff Swallows, Tree Swallows are aerial insectivores, but we detected no effects on females of blood sampling during the nestling stage when the demands of feeding nestlings require frequent provisioning visits (~ 10–13 per hour for females; Leonard and Horn 1996) and when metabolic rates of females are higher than during the incubation stage (Williams 1988). The only potential effect of bleeding females was during the incubation stage, but inclusion of the blood term in the full model was not supported by likelihood ratio tests.

Our results suggest the effect of bleeding nestlings on the number of young that fledge was limited to broods of 5 and 6. One reason why nestlings in larger broods might be more sensitive to bleeding is that feeding rates are lower per nestling and therefore masses may be lower (Leffelaar and Robertson 1986, Shutler et al. 2006, Hainstock et al. 2010, Winkler et al. 2011). When comparing nests where nestlings were weighed at the same age, we found some evidence that mean nestling mass was slightly lower as brood size increased. Nestlings that were comparatively lighter might have been more sensitive to the effects of having blood drawn, although overall we found the mean amount of blood taken was still less than 1% of mean nestling mass. In a small subset of nests where nestlings were bled, the proportion of nestlings with a catalogued blood sample was between 0.5 and 0.8 rather than 1. In most of these cases, bleeding all nestlings was likely attempted, but not successful for some nestlings, which is why they did not have a blood sample ID in our blood database. Nestlings

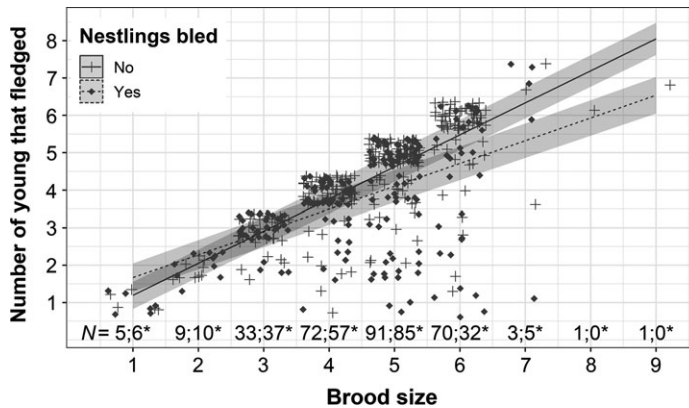


Fig. 4. Predicted effect of bleeding nestlings on the number of young that fledged when including an interaction between nestling bleed status and brood size at banding/bleeding. The lines are shaded with 95% confidence intervals. The original data are overlaid on top and slightly jittered to enhance readability. Sample size (number of nests) is reported for each brood size at the time of banding/bleeding. The values with an asterisk indicate the number of nests where nestlings were bled.

without a blood sample in an otherwise-bled brood were therefore most likely nestlings subjected to similar stresses as those that actually had their blood sampled. There was imprecision in the estimated effect of bleeding nestlings in larger-than-typical broods because the number of bled and control broods larger than six was small or zero. Regardless of that imprecision, or of what the potential mechanism(s) leading to the reduced number of young that fledged may be, we conclude that bleeding nestlings appeared to have measurable effects on fledging success for broods of five and six nestlings. Accordingly, in future studies of Tree Swallows, steps should be taken to reduce this observed effect of bleeding nestlings in these broods, such as adopting a more conservative rule of thumb for the amount of blood drawn (Voss et al. 2010). The effect of bleeding nestlings was limited to fledging outcomes because we found no effect of bleeding nestlings on the number that returned the following year relative to the total number that fledged.

One result that could potentially be interpreted as a negative effect of handling females (not necessarily bleeding) early during the incubation or nestling stage is that the timing of female captures was positively correlated with the probabilities of hatching and fledging. However, this was confounded with daily increasing probabilities of hatching or fledging for females captured closer to those endpoints. An interaction between capture timing

and the blood term was not well supported in any analysis, suggesting no difference between the reproductive metrics of early female captures in the bled or control group (captured for banding or identification, but not bled). The slight negative relationship between nestling age and the number of young that fledged in one of our analyses suggests that banding nestlings closer to fledge date had an adverse, although small, effect that was the same for bled and non-bled nestlings because an interaction between nestling age and nestling bleed status was not present in the best model. This result may be an artifact of our banding protocol; if nestlings in a brood appeared stunted and too small to band, we waited several days – sometimes beyond the usual banding window – before again attempting to band. Nestlings in these nests were in poor condition to begin with and, therefore, might have depressed the number of young fledging in later-banded nests.

Our analysis was limited to within-year effects and return rates of young and adult Tree Swallows in the following year. Our results do not rule out the possibility of longer-term accumulating carryover effects if, for example, individuals were bled in multiple years. We did not include a comparison of multi-year histories or interactions or inter-seasonal changes in reproductive effort and success over individual lifespans because there were so many possible combinations of manipulations, cohort, and return status (i.e.,

gaps between capture-years) that sample sizes would have been inadequate.

Although the effect of blood sampling depends on the species, locality, and the episode of selection when samples are taken, our results support those of other studies suggesting few negative effects of blood sampling on within-year reproductive success or survival of adult females (Sheldon et al. 2008, Redmond and Murphy 2011, Guillemain et al. 2015, Bowers et al. 2016, Smith et al. 2016). Our inability to detect an effect on adults suggests that collecting blood samples from breeding female Tree Swallows is comparable to that of routine capture and banding, which studies have also shown have little or no effect on birds when implemented properly (Angelier et al. 2011, Spotswood et al. 2012). Our study illustrates the importance of using large datasets to systematically test for possible effects that may be impossible to detect via direct monitoring during a single breeding season. Examining the potential effect of bleeding nestlings in other systems will be important, as will taking steps in our experimental system to ensure the welfare of nestlings in large broods and prevent blood-sampling-related mortality of nestlings.

ACKNOWLEDGMENTS

All handling of birds and blood collection were approved by the Institutional Animal Care and Use Committee of Cornell University (IACUC Protocol #2001-0051) and in accordance with NYS Department of Environmental Conservation Scientific License # 1477, and Federal Bird Banding Permit number 20576. This work was supported by NSF grants DEB-1242573, DEB-0717021, IBN-9207231, and IBN-0131437 to D. W. Winkler and IOS-0744753 to DWW and C. and D. Vleck as well as funds to DWW from the U. S. Department of Agriculture Hatch program, the U. S. Environmental Protection Agency, and Cornell University. We thank more than 75 crew workers over the years who have sampled blood from Tree Swallows. Jack Hruska, Eric Gulson-Castillo, and Kelly Hallinger helped proof the Tree Swallow blood sampling database, and we are grateful to Jabi Zabala, Mo Verhoeven, and Amos Belmaker for improving earlier versions of the manuscript and providing statistical insight.

LITERATURE CITED

- ANGELIER, F., H. WEIMERSKIRCH, AND O. CHASTEL. 2011. Capture and blood sampling do not affect foraging behavior, breeding success, and return

rate of a large seabird: the Black-browed Albatross. *Polar Biology* 34: 353–361.

- ARNOLD, T. W. 2010. Uninformative parameters and model selection using Akaike's Information Criterion. *Journal of Wildlife Management* 74: 1175–1178.
- ARNOLD, S. J., AND M. J. WADE. 1984. On the measurement of natural and sexual selection: applications. *Evolution* 38: 720–734.
- BATES, D., M. MAECHLER, B. BOLKER, AND S. WALKER. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67: 1–48.
- BLUMSTEIN, D. T. 2006. Developing an evolutionary ecology of fear: how life history and natural history traits affect disturbance tolerance in birds. *Animal Behaviour* 71: 389–399.
- 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.
- BOWERS, E. K., S. K. SAKALUK, AND C. F. THOMPSON. 2016. No effect of blood sampling or phytohaemagglutinin injection on postfledging survival in a wild songbird. *Ecology and Evolution* 6: 3107–3114.
- BROWN, M. B., AND C. R. BROWN. 2009. Blood sampling reduces annual survival in Cliff Swallows (*Petrochelidon pyrrhonota*). *Auk* 126: 853–861.
- BURNHAM, K. P., D. R. ANDERSON, AND K. P. HUYVAERT. 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behavioral Ecology and Sociobiology* 65: 23–35.
- CARNEY, K. M., AND W. J. SYDEMAN. 1999. A review of human disturbance effects on nesting colonial waterbirds. *Waterbirds* 22: 68–79.
- DORMANN, C. F., J. ELITH, S. BACHER, C. BUCHMANN, G. CARL, G. CARRÉ, J. R. GARCÍA MARQUÉZ, B. GRUBER, B. LAFOURCADE, P. LEITÃO, T. MÜKEMÜLLER, C. MCCLEAN, P. OSBORNE, B. REINEKING, B. SCHRÖDER, A. K. SKIDMORE, D. ZURELL, AND S. LAUTENBACH. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36: 27–46.
- ENGSTRAND, S., AND D. BRYANT. 2002. A trade-off between clutch size and incubation efficiency in the Barn Swallow *Hirundo rustica*. *Functional Ecology* 16: 782–791.
- FAIR, J., E. PAUL, AND J. JONES (eds.). 2010. Guidelines to the use of wild birds in research. Ornithological Council, Washington, D.C.
- GÖTMARK, F. 1992. The effects of investigator disturbances on nesting birds. In: *Current Ornithology* 9 (D. M. Power, ed.), pp. 63–104. Plenum Press, New York, NY.
- GUILLEMAIN, M., J. CHAMPAGNON, M. L. GOURLAY-LAROUR, F. CAVALLO, A. L. BROCHET, J. HARS, G. MASSEZ, T. GEORGE, P. Y. PERROI, AND V. JESTIN. 2015. Blood and cloacal swab sampling for avian influenza monitoring has no effect on survival rates of free-ranging ducks. *Ibis* 157: 743–753.
- HAINSTOCK, M. H., M. C. SMITH, J. CARR, AND D. SHUTLER. 2010. Parental investment and brood

- value in Tree Swallows, *Tachycineta bicolor*. *Behavior* 147: 441–464.
- HALEKOH, U., AND S. HØJSGAARD. 2014. A Kenward-Roger approximation and parametric bootstrap methods for test in linear mixed models – the R package pbkrtest. *Journal of Statistical Software* 59: 1–30.
- HASSELQUIST, D., M. F. WASSON, AND D. W. WINKLER. 2001. Humoral immunocompetence correlates with date of egg-laying and reflects work load in female Tree Swallows. *Behavioral Ecology* 12: 93–97.
- IBÁÑEZ-ÁLAMO, J. D., O. SANLLORENTE, AND M. SOLER. 2012. The impact of researcher disturbance on nest predation rates: a meta-analysis. *Ibis* 154: 5–14.
- JONES, J. 2003. Tree Swallows (*Tachycineta bicolor*): a new model organism? *Auk* 120: 591–599.
- LEFFELAAR, D., AND R. J. ROBERTSON. 1986. Equality of feeding roles and the maintenance of monogamy in Tree Swallows. *Behavioral Ecology and Sociobiology* 18: 199–206.
- LEONARD, M., AND A. HORN. 1996. Provisioning rules in Tree Swallows. *Behavioral Ecology and Sociobiology* 38: 341–347.
- LÜDECKE, D. [online]. 2018. sjPlot: Data visualization for statistics in social science. R package version 2.6.1. <<https://CRAN.R-project.org/package=sjPlot>> (Accessed 20 October 2018).
- LUKE, S. G. 2017. Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods* 49: 1494–1502.
- MCDONALD, T. L., AND G. C. WHITE. 2010. A comparison of regression models for small counts. *Journal of Wildlife Management* 74: 514–521.
- MURPHY, M. T., B. ARMBRECHT, E. VLAMIS, AND A. PIERCE. 2000. Is reproduction by Tree Swallows cost free? *Auk* 117: 902–912.
- MURTAUGH, P. 2009. Performance of several variable-selection methods applied to real ecological data. *Ecology Letters* 12: 1061–1068.
- O'CONNOR, C., D. R. NORRIS, G. T. CROSSIN, AND S. J. COOKE. 2014. Biological carryover effects: linking common concepts and mechanisms in ecology and evolution. *Ecosphere* 5: 1–11.
- PALACIOS, M. G., J. E. CUNNICK, D. W. WINKLER, AND C. M. VLECK. 2012. Interrelations among immune defense indexes reflect major components of the immune system in a free-living vertebrate. *Physiological and Biochemical Zoology* 85: 1–10.
- , D. W. WINKLER, K. C. KLASING, D. HASSELQUIST, AND C. M. VLECK. 2011. Consequences of immune system aging in nature: a study of immunosenescence costs in free-living Tree Swallows. *Ecology* 92: 952–966.
- R CORE TEAM. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- REDMOND, L. J., AND M. T. MURPHY. 2011. Multistate mark-recapture analysis reveals no effect of blood sampling on survival and recapture of Eastern Kingbirds (*Tyrannus tyrannus*). *Auk* 128: 514–521.
- SCHWARZ, G. E. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–464.
- SHELDON, L. D., E. H. CHIN, S. A. GILL, G. SCHMALTZ, A. E. M. NEWMAN, AND K. K. SOMA. 2008. Effects of blood collection on wild birds: an update. *Journal of Avian Biology* 39: 369–378.
- SHELDON, F. H., AND D. W. WINKLER. 1993. Intergeneric phylogenetic relationships of swallows estimated by DNA-DNA hybridization. *Auk* 110: 798–824.
- SHUTLER, D., R. G. CLARK, C. FEHR, AND A. W. DIAMOND. 2006. Time and recruitment costs as currencies in manipulation studies on the costs of reproduction. *Ecology* 87: 2938–2946.
- SMITH, R. J., T. J. BENSON, AND M. CAREY. 2016. Blood sampling does not negatively affect apparent survival, recapture probability, or within-season site fidelity in Field Sparrows (*Spizella pusilla*). *Auk* 134: 240–246.
- SPOTSWOOD, E. N., K. R. GOODMAN, J. CARLISLE, R. L. CORMIER, D. L. HUMBLE, J. ROUSSEAU, S. L. GUERS, AND G. G. BARTON. 2012. How safe is mist netting? Evaluating the risk of injury and mortality to birds. *Methods in Ecology and Evolution* 3: 29–38.
- STEDMAN, J. M., K. K. HALLINGER, D. W. WINKLER, AND M. N. VITOUSEK. 2017. Heritable variation in circulating glucocorticoids and endocrine flexibility in a free-living songbird. *Journal of Evolutionary Biology* 30: 1724–1735.
- STEIN, A., M. J. YOUNG, P. J. SEDDON, J. T. DARBY, AND Y. VAN HEEZIK. 2017. Investigator disturbance does not reduce annual breeding success or lifetime reproductive success in a vulnerable long-lived species, the Yellow-eyed Penguin. *Biological Conservation* 207: 80–89.
- STEPHENS, P. A., S. W. BUSKIRK, G. D. HAYWARD, AND C. M. DEL RIO. 2005. Information theory and hypothesis testing: a call for pluralism. *Journal of Applied Ecology* 42: 4–12.
- 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.
- SZÉKELY, T., J. N. WEBB, A. I. HOUSTON, AND J. M. MCNAMARA. 1996. An evolutionary approach to offspring desertion in birds. In: *Current Ornithology* 13 (V. Nolan and E. D. Ketterson, eds.), pp. 271–331. Plenum Press, New York, NY.
- VERBOVEN, N., AND J. TINBERGEN. 2002. Nest desertion: a trade-off between current and future reproduction. *Animal Behaviour* 63: 951–958.
- VOSS, M., D. SHUTLER, AND J. WERNER. 2010. A hard look at blood sampling of birds. *Auk* 127: 704–708.
- WICKHAM, H. 2016. ggplot2: elegant graphics for data analysis. Springer-Verlag, New York, NY.
- WIGGINS, D. A., T. PÄRT, AND L. GUSTAFSSON. 1994. Correlates of clutch desertion by female Collared Flycatchers (*Ficedula albicollis*). *Journal of Avian Biology* 25: 93–97.
- WILLIAMS, J. 1988. Field metabolism of Tree Swallows during the breeding season. *Auk* 105: 706–714.
- WINKLER, D. W. 1991. Parental investment decision rules in Tree Swallows: parental defense, abandonment, and the so-called Concorde Fallacy. *Behavioral Ecology* 2: 133–142.
- , AND P. E. ALLEN. 1996. The seasonal decline in Tree Swallow clutch size: physiological

constraint or strategic adjustment? *Ecology* 77: 922–932.

- , K. K. HALLINGER, D. R. ARDIA, R. J. ROBERTSON, B. J. STUTCHBURY, AND R. R. COHEN. 2011. Tree Swallow (*Tachycineta bicolor*). In: *The Birds of North America* (P. G. Rodewald, ed.). Cornell Lab of Ornithology, Ithaca, NY.
- , 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.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website.

Appendix S1. Results from AICc model selection and additional information about each analysis.