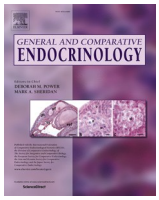
General and Comparative Endocrinology 385 (2026) 115019

Contents lists available at ScienceDirect

General and Comparative Endocrinology

journal homepage: www.elsevier.com/locate/ygcen



# Estradiol predicts pausing during mate-choice updating in a stress-buffered treefrog

Alexander T. Baugh [a,*], Mark A. Bee [b,c], Joseph Eyiolowope [a], Jingchen Hou [a], Caleb Kim [a]

[a] Department of Biology, Swarthmore College, 500 College Avenue, Swarthmore, PA 19081, USA
[b] Department of Ecology, Evolution, and Behavior, University of Minnesota - Twin Cities, 1479 Gortner Ave, St. Paul, MN 55108, USA
[c] Graduate Program in Neuroscience, University of Minnesota - Twin Cities, 321 Church Street SE, Minneapolis, MN 55455, USA

ARTICLE INFO

*Keywords:*
Anuran
Choosiness
Decision-making
Gray tree frog
Reproduction
Mate-sampling
Stress

ABSTRACT

Acute environmental stressors elevate circulating glucocorticoids on timescales of minutes to hours, raising the question of whether transient hypothalamic-pituitary-adrenal/interrenal (HPA) activation disrupts reproductive function during the brief breeding windows of seasonally breeding vertebrates. The conventional HPA-HPG crosstalk model predicts glucocorticoid-induced suppression of gonadal hormones and reproductive behavior, but seasonal breeders such as many frogs may be buffered against acute glucocorticoid elevation during peak reproductive readiness. Using an experimental cohort of Cope's gray treefrogs (*Hyla chrysoscelis*), we measured plasma 17β-estradiol ($E_2$) and tested a primary prediction of the buffering hypothesis: that acute corticosterone (CORT) manipulation will not suppress circulating $E_2$. We then explored whether endogenous $E_2$ covaried with mate-choice behavior in unmanipulated control females. Acute CORT manipulation did not alter circulating $E_2$, even at doses that increased plasma CORT by an order of magnitude, supporting the buffering hypothesis. Exploratory analyses revealed two complementary patterns: endogenous $E_2$ did not predict broad outcomes of mate choice (sexual proceptivity, locomotor effort, choice reversal probability, or initial preference for longer calls), but was positively associated with the duration of pauses made specifically at moments of decision reversal. Together, these findings support buffering of the female HPG axis against acute HPA activation during reproductive readiness while identifying decision deliberation as a fine-grained behavioral dimension that retains hormonal sensitivity during peak reproductive readiness.

## 1. Introduction

The conventional model of hypothalamic-pituitary-adrenal/interrenal (HPA) and hypothalamic-pituitary-gonadal (HPG) crosstalk predicts that strong HPA activation suppresses reproduction at multiple levels, including hypothalamic GnRH release, pituitary gonadotropin secretion, and gonadal steroidogenesis (Sapolsky et al., 2000; Wingfield and Sapolsky, 2003). However, this suppressive relationship is not universal. Spawning Pacific salmon (*Oncorhynchus* spp.), for example, sustain elevated cortisol throughout reproduction without arresting gametogenesis, and ovulated females complete spawning under acute experimental cortisol elevation (McConnachie et al., 2012; Couch et al., 2022). In breeders with tightly constrained fecundity windows, HPG output may be functionally uncoupled from HPA activation during the period of reproductive readiness — a buffering of the reproductive axis (Wingfield and Sapolsky, 2003; Bastien et al., 2018; Baugh et al., 2021) selectively favored when breeding opportunities are brief, infrequent, or deterministic for annual or lifetime fitness.

Anuran amphibians offer a powerful test system for evaluating HPA-HPG coupling. Most temperate-zone anurans breed during seasonal and weather-dependent windows, during which females become gravid, locate breeding aggregations, select mates using species-specific advertisement calls, and oviposit before exiting reproductive readiness (Wells, 2007; Gerhardt and Huber, 2002; Gall et al., 2019). Plasma corticosterone (CORT) in female anurans is elevated and highly variable during peak readiness, with concentrations spanning more than two orders of magnitude (Bastien et al., 2018; Gall et al., 2019; reviewed in Leary and Baugh, 2020). Building on these observations, we recently proposed that female reproductive behavior in seasonally breeding anurans is buffered against acute glucocorticoid elevations during peak breeding (Baugh et al., 2021). Consistent with this prediction, exogenous CORT in female Cope's gray treefrogs, *Hyla* (*Dryophytes*)

* Corresponding author at: Swarthmore College, 500 College Avenue, Singer Hall, Swarthmore, PA 19081, USA.
*E-mail address:* abaugh1@swarthmore.edu (A.T. Baugh).

https://doi.org/10.1016/j.ygcen.2026.115019
Received 21 May 2026; Received in revised form 26 July 2026; Accepted 17 September 2026
Available online 24 September 2026
0016-6480/© 2026 Elsevier Inc. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

*chrysoscelis* (AmphibiaWeb, 2024), did not affect sexual proceptivity or species-typical preferences for longer calls but did strongly modulate choosiness, and strictly at a moderate elevation (Baugh et al., 2021). Mate choosiness, which is costly in both time and locomotor effort, is operationalized as the forfeiture of an initial mate approach to pursue a suddenly more attractive, but also more distant, mate (Baugh et al., 2021). This effect was associated with variation in forebrain glucocorticoid receptor expression that was itself unaffected by the manipulation (Rodriguez-Santiago et al., 2024).

A central but untested prediction of HPA-HPG uncoupling during the period of reproductive readiness concerns the gonadal hormones. If acute HPA activation does not disrupt sexual proceptivity, choosiness, or species-typical preferences (Baugh et al., 2021), does it nonetheless suppress the circulating gonadal hormones that might support those behaviors? Under the conventional HPA-HPG crosstalk model, acute CORT elevation should reduce circulating estrogens; under the buffering hypothesis, no such suppression is expected. This is the primary prediction we test in the present study. In females, the principal HPG ligand at central and peripheral targets mediating reproductive behavior is 17β-estradiol ($E_2$), which acts on auditory, motor, and limbic circuits implicated in phonotaxis (Sisneros et al., 2004; Lynch and Wilczynski, 2008; Remage-Healey et al., 2013; Gall et al., 2019). $E_2$ is sufficient to induce phonotaxis in non-receptive túngara frogs, *Physalaemus pustulosus* (Chakraborty and Burmeister, 2009), establishing $E_2$ as a causal signal for receptivity, but whether endogenous $E_2$ variation among already-receptive females tracks any aspect of mate-choice behavior remains unexplored in anurans.

Beyond confirming whether acute CORT suppresses $E_2$, the present study also allowed us to ask a second question: among unmanipulated, already-proceptive females, does natural variation in plasma $E_2$ predict any component of mate-choice behavior? The conventional view — based on pharmacological $E_2$ induction of proceptivity in túngara frogs (Chakraborty and Burmeister, 2009) — provides reason to expect that endogenous $E_2$ should covary with broad mate-choice measures such as proceptivity, choosiness, or species-typical preference. Alternatively, if behavior in receptive females is itself buffered, endogenous $E_2$ variation may be uninformative for binary outcomes. We tested four broad mate-choice measures (proceptivity, locomotor effort, choosiness, and species-typical preference) for $E_2$ covariation in females from the control treatments (no exogenous CORT). In addition, we examined a previously unanalyzed behavior: pause duration during decision-reversal. This measure was motivated by a broader literature on decision-making in vertebrates. In rodents and humans, transient pauses at choice points are recognized as behavioral indicators of deliberation: Tolman (1939) and colleagues termed this "vicarious trial-and-error" (VTE) and proposed that it reflects covert evaluation of alternatives, and neural recordings show that VTE pauses are accompanied by serial hippocampal representations of the available options (Redish, 2016). More generally, evidence-accumulation models predict that decision-time scales with the cognitive demand of the choice (Ratcliff and McKoon, 2008). In animal mate choice specifically, female mating latency tracks aspects of the difficulty of the sampling context. For example, in túngara frogs, females demonstrate categorical perception of male calls along a continuum of conspecific to heterospecific — stimuli judged to be the same elicit faster choices than stimuli that span a category boundary (Baugh et al., 2008). Similarly, in female *H. chrysoscelis*, latencies are longer for mating calls that are (a) masked by noise (Bee, 2007; Bee and Swanson, 2007; Bee and Schwartz, 2009; Ward et al., 2013); (b) co-presented in the presence of spectrally similar distractors (Nityananda and Bee, 2011); or (c) similar in pulse repetition rate (Ward et al., 2013). Pause duration during phonotaxis trajectory updating is therefore a candidate deliberation index, and an exploratory analysis of whether exogenous CORT and endogenous $E_2$ relate to this measure complements the broad-behavior characterization above.

## 2. Materials and methods

### 2.1. Animals, treatments, and behavior

Animal collection, experimental design, phonotaxis testing, blood sampling, and CORT enzyme immunoassays are described in Baugh et al. (2021); methods are summarized briefly here. Mating pairs of *H. chrysoscelis* (western genetic lineage) were collected in June 2018 and 2019 from wetlands in Minnesota, USA.

In a mixed within- and between-subjects design ($N = 107$ females; $N = 19$–$22$ per treatment), females were randomly assigned to one of five CORT treatment groups: no injection (No Inject), Vehicle (sesame oil), Low CORT (20 ng $g^{-1}$), Medium CORT (60 ng $g^{-1}$), or High CORT (180 ng $g^{-1}$). Each female was tested in a battery of five two-alternative phonotaxis trials (one static, two dynamic control, two dynamic experimental) before and after the designated CORT treatment ("Pre-treatment" and "Post-treatment"; SM1). Phonotaxis is a standard assay of mate choice behaviors in frogs (Gerhardt and Huber, 2002). All stimuli, arenas, and procedures are described in Baugh et al. (2021). Briefly, females were given a choice between two synthetic calls differing in pulse number (PN) only; previous work in this species has established that females generally prefer calls with higher PN (Bee, 2008; Tanner et al., 2017; Baugh et al., 2021). The two calls alternated from playback speakers located 2 m apart such that they simulated two nearby calling males differing in attractiveness to females (Fig. 1). In *static trials*, females chose between two calls with PNs that were either 2 SD above or below the population average (22 pulses vs 38 pulses); the call assigned to each speaker remained fixed over the duration of the trial. In *dynamic trials*, females were tested with the average call (30 pulses) versus an alternative having PNs that were either 2 SD above (38 pulses) or 2 SD below (22 pulses) average. In *dynamic experimental* trials, the position of the higher-PN call was switched between speakers when the female crossed an approach boundary en route to one of the speakers (Fig. 1), allowing measurement of choice reversals (i.e., choosiness). In *dynamic control* trials, no switch occurred. In all trials, we scored a female's choice of stimulus when she crossed a choice boundary in front of one of the speakers (Fig. 1).

Plasma CORT was quantified by enzyme immunoassay (Arbor Assays K014) following diethyl ether extraction; full details and validation are reported in Baugh et al. (2021). Bleed timing followed injection by approximately 60 min, a window over which exogenous CORT was confirmed to be elevated and behaviorally active at the assayed treatment doses (Baugh et al., 2021).

### 2.2. Plasma $E_2$

Plasma $E_2$ was measured using a commercial enzyme immunoassay (Arbor Assays KB30-H) following the diethyl ether extraction protocol previously validated in this species (Gall et al., 2019). Briefly, 8 μL of plasma was double extracted with 2 mL diethyl ether, dried under nitrogen, and reconstituted in commercial assay buffer at a 1:30 dilution. Samples were assayed in duplicate alongside blanks, standards, and aliquots of pooled plasma to determine coefficients of variation (CV); assay sensitivity was 2.21 pg $mL^{-1}$ and intra- and inter-assay CV were 2.1% and 10.7%. Technical repeatability was estimated by independently re-extracting and analyzing 29 experimental samples on a separate (fourth) plate; the linear correlation between replicates was high ($r = 0.96$, $R^2 = 0.93$; SM2). Experimental samples were processed across three plates (kit lot number 25A092). The manufacturer reports antibody cross-reactivities of 100% with 17β-estradiol and $< 1\%$ with other steroids screened (e.g. estrone, 17α-estradiol). Extraction efficiency for plasma $E_2$ in this species and assay system was established using the identical protocol in our laboratory (76.4 ± 3.6% SE; Gall et al., 2019); following that study, concentrations are not corrected for extraction loss, a choice that leaves all rank-based and log-scale inferential results unchanged.

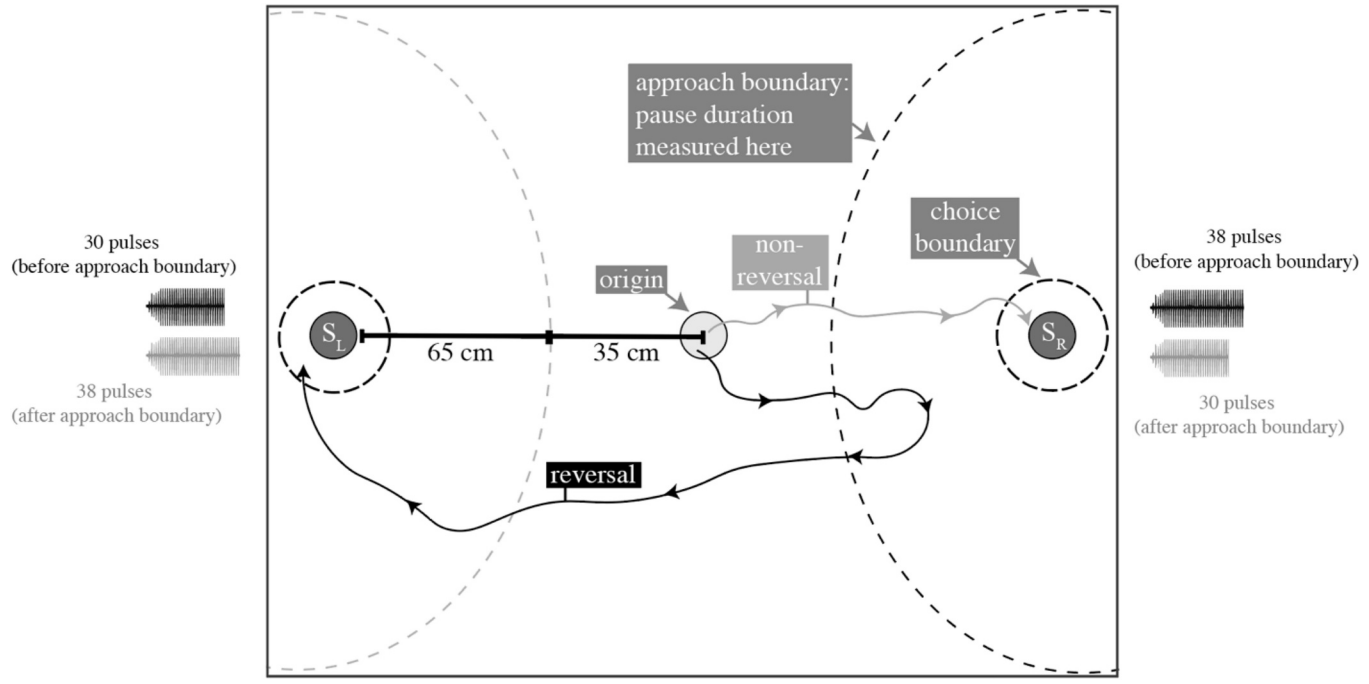



**Fig. 1.** Arena for phonotaxis testing in female *Hyla chryoscelis*. Each phonotaxis trial began with the female under an acoustically transparent cage at the origin (5 cm radius). Speakers ($S_L$, $S_R$) were surrounded by 10-cm-radius choice boundaries; approach boundaries were placed at a 65-cm radius from each speaker. Stylized traces show a reversal (black) and non-reversal (gray) path. Stylized oscillograms show the average-PN (30 pulses) and high-PN (38 pulses) calls in their initial (black) and final (gray) states in a dynamic experimental trial. Replotted from Baugh et al. (2021).

### 2.3. Pause duration

Pause duration was defined as the duration of the next cessation of locomotion occurring after the moment the female crossed an approach boundary toward one of the two speakers. Pause duration during phonotaxis was extracted from EthoVision XT (Noldus) tracking files for each dynamic trial and manually confirmed in a random subset of 30 trials. Pauses were measured whenever a female crossed an approach boundary. Because the stimulus switch in dynamic experimental trials was triggered only when a female crossed the approach boundary toward the higher-PN call (Fig. 1), pause duration in dynamic experimental trials carries a conditional interpretation: in non-reversal trials, the pause occurred as the female continued toward what was initially the higher-PN call (and which, after the stimulus switch, was now broadcasting the lower-PN call from the original speaker); in reversal trials, the pause occurred, by definition, at the inflection point of decision-updating, just prior to when she reversed course to pursue the higher-PN call from its new location. These two trial types likely reflect distinct cognitive operations and were therefore analyzed separately.

### 2.4. Statistical analysis

Our analyses had one a priori test and a set of exploratory analyses. The a priori test concerned the buffering hypothesis: acute CORT manipulation will not suppress plasma $E_2$. Plasma $E_2$ was $\log_{10}$-transformed and analyzed by one-way ANOVA across the five CORT treatment groups, with planned contrasts of (a) Vehicle versus No Inject (effect of injection per se on $E_2$) and (b) each CORT dose versus the combined control groups (No Inject + Vehicle; effect of CORT on $E_2$). No plasma was physically pooled for this contrast; "combined" refers to the statistical grouping of the two control treatments.

We also evaluated whether exogenous CORT altered pause duration. Subject-level mean $\log_{10}$ pause duration at Post-treatment testing was compared across the five treatment groups by one-way ANOVA using the same planned contrasts applied to $E_2$. As a check on random assignment we applied the same ANOVA to Pre-treatment pause duration, and we additionally tested the Pre- to Post-treatment change score. Because pause duration carries a conditional interpretation depending on whether a female reversed (§2.3), we repeated the one-way ANOVA separately within non-reversal and reversal cells. To test the interaction while retaining both trial types in a single model, we fitted a trial-level linear mixed model of $\log_{10}$ pause duration with CORT exposure (injected versus control, coded binary to preserve degrees of freedom given the small reversal cell), reversal status, and their interaction as fixed effects, and female identity as a random intercept; the interaction was evaluated by likelihood-ratio test (maximum likelihood) against the corresponding additive model, and we report the intraclass correlation as an index of within-female repeatability. Finally, because membership in the reversal cell is behaviorally self-selected rather than randomly assigned, we conducted a baseline-adjusted analysis restricted to females that reversed at both timepoints, using ANCOVA on Post-treatment $\log_{10}$ pause duration with treatment group as the factor and Pre-treatment $\log_{10}$ pause duration as the covariate.

The exploratory analyses characterized whether endogenous $E_2$ in control females (No Inject + Vehicle, $N = 42$) covaried with mate-choice behavior. We restricted these analyses to control-treatment females because their plasma $E_2$ reflects only endogenous variation, since their HPG axis was not perturbed by exogenous CORT.

We examined four broad mate-choice behaviors spanning the canonical dimensions of female mate choice in this species: proceptivity (latency to exit the origin; latency to cross the approach boundary), locomotor effort (total path length), choosiness (the proportion of dynamic experimental trials in which females reversed), and preference (the proportion of all dynamic trials in which the female's first crossing of an approach boundary was toward the higher-PN call). We used Spearman $\rho$ at the subject level to examine the relationships between $E_2$ and these measures of mate choice behavior. The use of rank-based correlations is justified because choosiness and preference were proportions bounded between 0 and 1, preference was binary at the trial level, and latencies and path-length were right-skewed. The trial-level

binary structure of preference was additionally analyzed using a binomial generalized estimating equation (GEE) with a logit link, $\log_{10} E_2$ as the predictor, and subject as the clustering variable, specifying an exchangeable working correlation structure and robust (sandwich) standard errors to account for repeated trials within females; the estimated within-female correlation was negligible, and parameter estimates were unchanged under independence and first-order autoregressive working correlations (SM3). Because preference in static trials was near the ceiling (38 of 42 control-treatment females chose the higher-PN call), we did not test $E_2$-modulated preference in static trials.

We also examined a previously unanalyzed measure: pause duration immediately following stimulus switching. Because all females were captured in amplexus and tested prior to completing oviposition, and the testing window from initial Pre-treatment trial to bleed was approximately 90 min, we treat each female's plasma $E_2$ as a within-female biological constant across her Pre-treatment and Post-treatment behavioral trials; prior work in this species indicates that the post-oviposition decline in circulating $E_2$ occurs within a 24-h window (Gall et al., 2019), not within the pre-oviposition reproductive readiness window over which our behavioral tests were conducted. We therefore pooled Pre-treatment and Post-treatment trials in the primary analysis.

Plasma $E_2$ and pause duration were $\log_{10}$-transformed prior to analysis. Diagnostic tests indicated that the relationship between $\log_{10} E_2$ and $\log_{10}$ pause duration was monotonic but not strictly linear (SM4); we therefore used Spearman rank correlation. Per-subject mean $\log_{10}$ pause duration was computed separately for reversal and non-reversal trials in dynamic experimental trials (Pre-treatment and Post-treatment combined), and the relationship with $\log_{10} E_2$ within each cell was tested using Spearman ρ. Each female contributed one cell-mean for her non-reversal trials (all 42 control females); 28 of the 42 also reversed in at least one trial and additionally contributed a cell-mean for their reversal trials. The 28 dual-contributors appear in both cell-level correlations. The difference between the reversal-condition and non-reversal-condition correlations was tested by a permutation test reshuffling Reversal labels across the dataset (10,000 iterations) and recomputing the difference in Spearman ρ. To test whether the E2-pause association in reversal trials was confounded by plasma CORT, we computed partial Spearman correlations of $\log_{10} E_2$ with $\log_{10}$ pause duration controlling for $\log_{10}$ CORT. As a check on the conditional-covariation framing of the $E_2$-pause-reversal result, we also tested whether mean $\log_{10}$ pause duration itself differed between reversal and non-reversal trials using a Welch *t*-test on trial-level values.

For all significance tests we additionally report effect sizes: Cohen's d for two-group comparisons (Vehicle vs No Inject; CORT doses vs combined controls; reversal vs non-reversal pause), $\eta^2$ for one-way ANOVAs, odds ratios for the binomial GEE, and Spearman ρ as the effect-size estimate for all rank correlations. Throughout, non-independence of repeated trials within females was handled by conducting analyses at the subject level (one cell-mean per female) or by including female identity as a random intercept (linear mixed model) or clustering unit (GEE); the effective sample size for inference is therefore the number of females rather than the number of trials. Because control-treatment females received no manipulation, their Post-treatment behavior and their single plasma $E_2$ measurement both reflect endogenous state, which is why Pre- and Post-treatment control trials are analyzed together.

Sensitivity analyses examined the two control treatment groups separately, restricted analyses to Pre-treatment trials only (to address the temporal proximity of $E_2$ sampling to Post-treatment behavior), and weighted subject-level cell-means by the number of contributing trials. We did not extend the $E_2$-behavior analyses to the three CORT-injected treatment groups because each CORT-injected female's single $E_2$ measurement was sampled approximately 60 min after injection, precluding clean inference about endogenous E2-behavior coupling at the individual level (see §3.5). Analyses were conducted in R 4.4.2 (R Core Team, 2024) using RStudio 2025.05.0 + 496, and in Python 3.12 (pandas 3.0.2, SciPy 1.17.1, statsmodels 0.14).

## 3. Results

### 3.1. Acute CORT manipulation does not alter plasma $E_2$

Under the conventional HPA-HPG crosstalk model, acute CORT elevation should reduce circulating $E_2$; the buffering hypothesis predicts no such suppression. Plasma $E_2$ concentrations were similar across the five CORT treatment groups (Fig. 2A), despite the manipulation producing the expected dose-dependent elevation in plasma CORT (see Fig. 3 in Baugh et al., 2021). One-way ANOVA on $\log_{10} E_2$ with treatment group as the predictor was not significant ($F_{4,100} = 0.79$, $p = 0.53$, $\eta^2 = 0.03$). Vehicle females did not differ from No Inject controls (Welch's $t_{37.5} = -0.57$, $p = 0.57$, d = −0.18), and none of the three CORT doses differed from the combined control groups (Low: $t_{46.4} = 0.85$, $p = 0.40$, d = 0.21; Medium: $t_{35.8} = 0.16$, $p = 0.87$, d = 0.05; High: $t_{45.6} = -1.20$, $p = 0.24$, d = −0.30). Plasma $E_2$ ranged from 313 to 2902 pg $mL^{-1}$ (mean ± SD: 1217 ± 563 pg $mL^{-1}$), consistent with values previously reported in this species and population during peak reproductive readiness (Gall et al., 2019). The lack of effect of exogenous CORT injections on endogenous $E_2$ — including in the Medium-CORT group, which doubles choosiness behavior (Baugh et al., 2021), and in the High-CORT group, which reaches the upper physiological range (120 ng $mL^{-1}$; Gall et al., 2019) — rejects the conventional crosstalk prediction and supports the buffering hypothesis as applied to acute CORT in receptive female *H. chrysoscelis*.

### 3.2. Exogenous CORT alters choosiness but not pause duration

In these same animals, moderate (60 ng $g^{-1}$) CORT elevation increased choosiness — the rate at which a female reverses an initial mate choice to pursue a suddenly more attractive call — by more than 100%, whereas the other doses did not (see Fig. 6 in Baugh et al., 2021). This dose-specific behavioral effect, produced within the same post-injection window examined here, establishes a positive control against which to interpret the effect of CORT on pause duration, a behavior not previously analyzed.

Against this active manipulation, pause duration was unaffected by exogenous CORT. Subject-level mean pause duration at Post-treatment testing did not differ across the five treatment groups (one-way ANOVA on $\log_{10}$ pause: $F_{4,102} = 0.45$, $p = 0.77$, $\eta^2 = 0.017$, $N = 107$; Fig. 2B). No CORT dose differed from the combined control groups (Welch's t throughout; Low: $t_{43.9} = -0.83$, $p = 0.41$, d = −0.22; Medium: $t_{32.3} = -0.03$, $p = 0.98$, d = −0.01; High: $t_{43.2} = -1.26$, $p = 0.21$, d = −0.33), and vehicle-injected females did not differ from No Inject controls ($t_{34.6} = 0.03$, $p = 0.98$, d = 0.01). Groups were comparable at Pre-treatment testing, confirming random assignment ($F_{4,102} = 1.43$, $p = 0.23$, $\eta^2 = 0.053$), and the Pre- to Post-treatment change in pause duration likewise did not differ among groups ($F_{4,102} = 0.71$, $p = 0.59$, $\eta^2 = 0.027$).

Because pause duration carries a conditional interpretation depending on whether a female reversed (§2.3), we repeated the analysis separately within each trial type. CORT altered pause duration in neither non-reversal trials ($F_{4,89} = 0.07$, $p = 0.99$, $\eta^2 = 0.003$) nor reversal trials ($F_{4,39} = 0.47$, $p = 0.76$, $\eta^2 = 0.046$), and a trial-level linear mixed model with random subject intercepts detected no interaction between CORT exposure (injected versus control) and reversal status (likelihood-ratio $\chi^2_1 = 0.14$, $p = 0.71$). In that model, female identity accounted for 46% of the total variance in $\log_{10}$ pause duration (intraclass correlation = 0.46), indicating that pause duration is a repeatable individual attribute; the absence of a CORT effect therefore does not reflect an unreliable measure. We note that membership in the reversal cell is behaviorally self-selected rather than randomly assigned, so between-group comparisons within that cell are quasi-experimental; a baseline-adjusted analysis restricted to females that reversed at both timepoints ($N = 22$) likewise

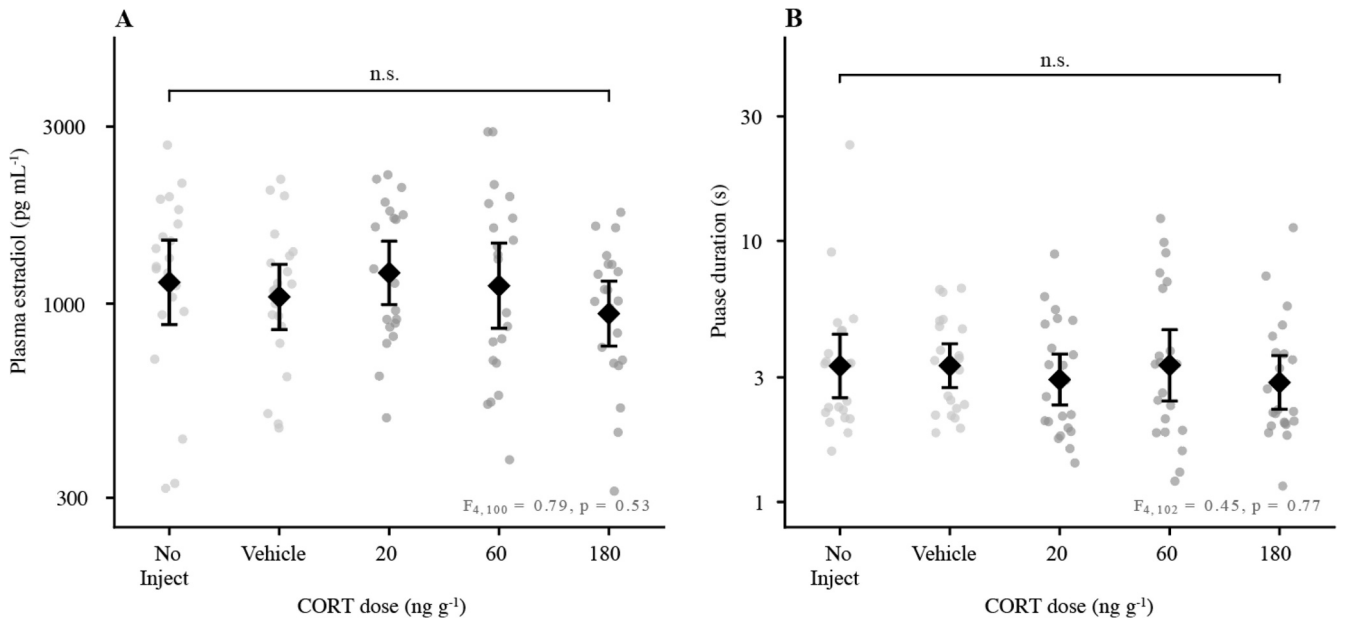



**Fig. 2.** Endogenous plasma $E_2$ and pause duration are both unaffected by acute corticosterone elevation in female *Hyla chrysoscelis*. (A) Plasma $E_2$ concentrations by treatment group ($N = 105$). (B) Subject-level mean pause duration during Post-treatment dynamic experimental trials by treatment group ($N = 107$). In both panels, individual females are shown as points; the filled diamond and bar give the geometric mean with its back-transformed, asymmetric 95% confidence interval (computed on the $\log_{10}$ scale). Doses: No Inject and Vehicle received no CORT; Low, Medium, and High received 20, 60, and 180 ng $g^{-1}$, respectively. n.s., no difference among groups (A: $F_{4,100} = 0.79$, $p = 0.53$, $\eta^2 = 0.03$; B: $F_{4,102} = 0.45$, $p = 0.77$, $\eta^2 = 0.017$). The same Medium dose more than doubled choosiness in these animals (Baugh et al., 2021; §3.2), confirming the manipulation was behaviorally active. The previously published dose-dependent CORT elevation is provided in (Baugh et al., 2021; their Fig. 3).

revealed no treatment effect (ANCOVA controlling for Pre-treatment pause: $F_{4,16} = 0.53$, $p = 0.71$).

Acute HPA activation therefore modifies whether a female revises a committed mate choice without modifying how long she pauses while doing so. Because the same doses, in the same trials, produced a large change in choosiness, this null outcome cannot be attributed to an ineffective manipulation.

### 3.3. Endogenous $E_2$ does not predict broad mate-choice measures

Having established that acute CORT does not suppress $E_2$, we explored whether endogenous $E_2$ variation within control females covaried with any of four broad mate-choice measures (SM5). In control-treatment females (No Inject + Vehicle, $N = 42$), endogenous plasma $E_2$ was unrelated to four broad measures of mate-choice behavior: per-subject probability of choice reversal in dynamic experimental trials (Spearman $\rho = 0.11$, $p = 0.50$), latency to cross the approach boundary ($\rho = 0.00$, $p = 0.98$), latency to exit the origin ($\rho = -0.19$, $p = 0.23$), and total path length during phonotaxis ($\rho = -0.04$, $p = 0.81$). Because of the pulse number preference ceiling effect (see §2.4), we performed an additional analysis of female preferences using the initial approach-boundary crossing in the four dynamic trials (two PN contrasts × control and experimental trials) at both Pre-treatment and Post-treatment time points. Following Baugh et al. (2021), we treated the speaker toward which the female first crossed an approach boundary as a behavioral indicator of her initial preference, before any stimulus alteration in experimental trials. Across all 333 dynamic trials contributed by the 42 control-treatment females, 70.9% began with initial crossings toward the higher-PN call, and per-subject proportions spanned a 0.33–1.00 range (Q1 = 0.62, median = 0.75, Q3 = 0.75) — far from a ceiling, with substantial between-female variance. Endogenous plasma $E_2$ was positively but non-significantly associated with this proportion at the subject level (Spearman $\rho = 0.21$, $p = 0.19$, $N = 42$) and at the trial level with subject-clustered standard errors (binomial GEE: $\beta = 0.81$, $p = 0.063$, Odds Ratio = 2.26, 95% CI [0.96, 5.33], $n = 333$ trials). The direction of effect is consistent with the species-typical preference pattern, but the magnitude is too modest to support a claim of $E_2$-modulated initial preference in this dataset. Together, these null and marginal results indicate that broad measures of mate choice — sexual proceptivity, locomotor effort, overall reversal probability, and initial preference for the higher-PN call — are not strongly modulated by natural variation in circulating $E_2$ in gravid females.

### 3.4. Endogenous $E_2$ predicts pause duration specifically during decision-reversal

Pause durations across all 166 dynamic experimental trials in control-treatment females had a mean ± SD of 4.14 ± 3.36 s (median 2.92 s; range 1.00–30.32 s). Mean pause duration between non-reversal trials (4.15 ± 3.66 s; $n = 124$) and reversal trials (4.12 ± 2.32 s; $n = 42$) did not significantly differ (Welch $t = 0.88$, $p = 0.38$, Cohen's d = 0.14, 95% CI [−0.21, 0.49]; $\log_{10}$-transformed); the $E_2$-pause association is therefore a difference in covariation structure, not a difference in average pause duration.

In contrast to the broad-outcome nulls reported above, an exploratory examination of pause duration revealed a striking conditional structure: endogenous $E_2$ was significantly associated with pause duration in trials where the female reversed her decision, but not in trials where she did not. Among control-treatment females, $E_2$ was positively correlated with pause duration in trials in which the female reversed her decision (Spearman $\rho = 0.51$, $p = 0.005$, $N = 28$) but not in trials in which she did not reverse ($\rho = -0.03$, $p = 0.84$, $N = 42$; Fig. 3). The difference between the two correlations ($\Delta\rho = 0.54$) was significant under a permutation test reshuffling Reversal labels across the dataset (10,000 iterations; $p = 0.041$).

The reversal-condition signal was directionally consistent across sensitivity analyses. The positive $E_2$-pause correlation in females that reversed held when the two control groups were analyzed separately (No Inject $\rho = 0.44$, $p = 0.087$, $N = 16$; Vehicle $\rho = 0.87$, $p = 0.0002$, $N = 12$), in Pre-treatment trials alone ($\rho = 0.55$, $p = 0.008$, $N = 22$), and

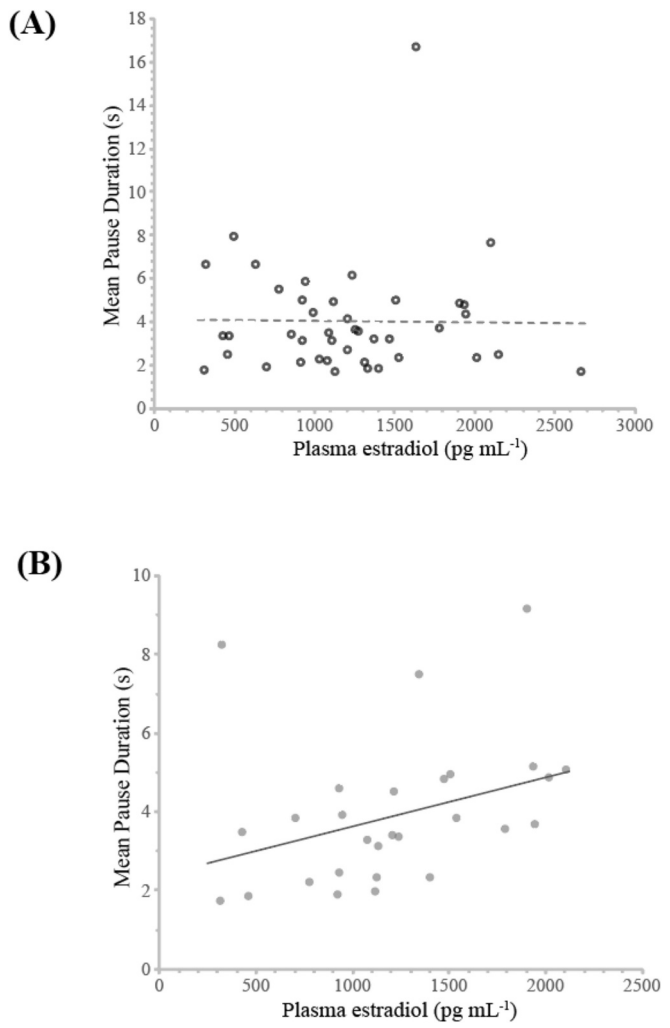



**Fig. 3.** Endogenous plasma $E_2$ concentrations versus mean pause duration in control-treatment females (No Inject + Vehicle) for (a) non-reversal trials (upper, open; dashed line: OLS fit; Spearman $\rho = -0.03$, $p = 0.84$, $N = 42$) and (b) reversal trials (lower, closed; solid line: OLS fit; Spearman $\rho = 0.51$, $p = 0.005$, $N = 28$). Each point represents one subject's mean across her contributing trials in that cell. Depicted linear regression lines show OLS fits for visualization; statistics are Spearman ρ. Rank-transformed and $\log_{10}$-transformed depictions of these relationships are provided in SM4.

after weighting subjects by the number of contributing reversal trials (weighted Spearman $\rho = 0.51$, permutation $p = 0.012$). Among females that reversed, 17 of 28 females contributed only one reversal trial (median 1, range 1–3); the result therefore reflects a between-subjects pattern in which females with higher endogenous $E_2$ paused longer specifically when reversing, rather than a within-subject pattern.

### 3.5. Endogenous $E_2$ is not confounded with plasma CORT or body condition

Within control-treatment females, plasma $E_2$ was not correlated with plasma CORT (Spearman $\rho = 0.18$, $p = 0.25$), and plasma CORT was not associated with pause duration in either reversal ($\rho = -0.04$, $p = 0.83$, $N = 28$) or non-reversal trials ($\rho = 0.17$, $p = 0.28$, $N = 42$). Controlling for $\log_{10}$ CORT left each correlation essentially unchanged: in females that reversed, partial $\rho = 0.52$ ($p = 0.004$) compared with the bivariate $\rho = 0.51$; in females that did not reverse, partial $\rho = -0.07$ ($p = 0.68$) compared with the bivariate $\rho = -0.03$ (§3.4). Body size and body condition were likewise not significantly associated with pause duration in either trial type (body mass, SVL, and residual body mass: all $p > 0.07$; SM6), and the $E_2$–pause association in reversal trials was unchanged or slightly stronger when each was controlled (partial $\rho = 0.51$ to 0.60, all $p < 0.007$). The $E_2$-pause-reversal association is therefore not attributable to covariation with plasma CORT, body size, or body condition.

## 4. Discussion

### 4.1. Buffering of the HPG axis

The buffering hypothesis was supported. Acute experimental CORT elevation did not alter plasma $E_2$, even at doses reaching the upper physiological range (Gall et al., 2019) and including the moderate dose that doubles mate choosiness in this species (Baugh et al., 2021). If the conventional HPA-HPG crosstalk model applied during peak reproductive readiness in female *H. chrysoscelis*, we would expect at least a downward trend in plasma $E_2$ at the higher CORT doses. The lack of such a trend aligns with results from spawning Pacific salmon, in which ovulated females complete spawning under acute cortisol elevation (McConnachie et al., 2012; Couch et al., 2022), supporting the broader hypothesis that buffering of HPG output is selectively favored when reproductive opportunities are brief, infrequent, or deterministic for annual or lifetime fitness (Wingfield and Sapolsky, 2003; Bastien et al., 2018). We measured plasma $E_2$ at a single timepoint approximately 60 min post-injection, a window over which exogenous CORT was confirmed to be elevated and behaviorally active (Baugh et al., 2021). Effects operating on longer timescales cannot be excluded, but the window relevant to acute environmental stressors during a single breeding night — and a topic of conservation-physiology relevance — is the one captured here.

The absence of an acute $E_2$ response is consistent with current models of glucocorticoid action in amphibians, in which rapid behavioral effects are mediated primarily through neural nongenomic mechanisms, whereas many endocrine effects of glucocorticoids are thought to reflect slower genomic actions (Carr, 2024). Rapid, membrane-initiated actions of corticosterone on the amphibian nervous system are well documented — corticosterone suppresses newt courtship within minutes via neuronal membrane receptors (Moore and Miller, 1984; Orchinik et al., 1991) — and such nongenomic effects on neural targets provide a plausible route for the behavioral effects of CORT observed in this system, including the modulation of choosiness reported previously (Baugh et al., 2021; §3.2). Suppression of ovarian steroidogenesis, however, would be expected to involve genomic regulation of the HPG axis and/or ovarian steroidogenic pathways rather than the rapid membrane-initiated mechanisms implicated in acute behavioral responses (Whirledge and Cidlowski, 2017). Because genomic regulation of steroidogenic output is inherently slower than membrane-initiated signaling, a change in circulating $E_2$ would not necessarily be detectable within the 60-min sampling window; in mammals, stress-induced changes in gonadal steroidogenic output emerge over several hours (Lin et al., 2014). A 60-min window is therefore unlikely to reveal genomically mediated changes in plasma $E_2$ even where they eventually occur, and our data cannot exclude slower suppression on longer timescales. What our findings do establish, however, is that within the acute window relevant to a single breeding night — during which corticosterone demonstrably alters choosiness in this population (Baugh et al., 2021; §3.2) — plasma $E_2$ and pause duration are both unaffected by CORT treatment. Whether local gonadal signaling contributes to this stability is an open question: kisspeptin systems are expressed in the amphibian gonad (Chianese et al., 2013), and their role in the female is untested.

### 4.2. Endogenous $E_2$ and pausing during decision-updating

Our exploratory analyses identified two complementary patterns: endogenous $E_2$ was unrelated to the broad measures of mate choice but was specifically associated with pause duration during decision-

reversal. We treat both findings as hypothesis-generating rather than confirmatory, but together they suggest a working framework that warrants direct testing in future work. Pause duration in reversal trials brackets a moment of decision-updating: the female has invested locomotor effort toward an initially preferred mate whose call has just been altered to a less-attractive variant, and she must decide whether to abandon her trajectory and pursue the now-more-attractive caller from the opposite speaker. In non-reversal trials, by contrast, the pause occurs while she is committed to her course and does not bracket a decision-updating event. The conditional structure we observe — $E_2$ covaries with pause duration only when decision-updating is engaged — is consistent with $E_2$ modulating a deliberative process rather than a general motor or motivational state.

This finding parallels mammalian decision-making research. VTE pausing in rodents tracks the difficulty of the choice and is underpinned by hippocampal-prefrontal circuits implicated in deliberation (Redish, 2016), and $E_2$ modulates effort-based decision making and reward sensitivity in female rats (Uban et al., 2012). Our result extends this picture to anurans: in already-receptive females, endogenous $E_2$ does not appear to drive the dichotomous outcome of mate preference but covaries with the temporal structure of instances where a choice is being updated. This is consistent with $E_2$ modulating evidence accumulation during conflict, rather than gating entry into mate-search behavior (Chakraborty and Burmeister, 2009). Further experiments are needed to probe the potential role of estrogen receptors in modulating decision making circuits in the midbrain and forebrain.

A receptor-level dimension of buffering further connects to this picture. Rodriguez-Santiago et al. (2024) reported that forebrain GR expression was unaltered by acute CORT manipulation in this same study population, while hindbrain and midbrain GR were affected; nonetheless, only forebrain GR positively predicted inter-individual variation in choosiness. The receptor pool that predicts the behavioral phenotype is the same pool that is insulated from the manipulation. Combined with our present finding that plasma $E_2$ is also unaltered by acute CORT, this suggests that both the gonadal hormone ($E_2$) and the central glucocorticoid effector relevant to mate choosiness (forebrain GR) are insulated from acute HPA activation. The forebrain may thus serve as a convergence point for two distinct buffered signals — gonadal E2, acting via estrogen receptors, and glucocorticoid sensitivity, set by forebrain GR expression — with stable E2 input modulating the temporal structure of decision-updating against a background of stable glucocorticoid signaling capacity. This buffering may be actively maintained rather than passive: reviewing glucocorticoid effects on anuran mate choice, Carr (2024) notes that sensory and forebrain regions in female frogs may be protected from elevated corticosterone through local metabolism by 11β-hydroxysteroid dehydrogenase (Baugh et al., 2021; Shearer et al., 2019). Our finding that plasma E2 is likewise insulated from acute CORT is consistent with buffering operating at multiple levels — peripheral gonadal output and central receptor availability alike.

## 5. Conclusions

The female reproductive endocrine and behavioral phenotype of breeding *H. chrysoscelis* is buffered against acute HPA activation: plasma $E_2$ is unaffected by exogenous CORT. Exploratory analyses in unmanipulated females suggest that natural variation in endogenous $E_2$ is uninformative for several broad measures of mate choice behavior but covaries with pause duration during decision-reversal, hinting that gonadal hormone variation may modulate decision deliberation even when broader behaviors are buffered. These exploratory findings are hypothesis-generating and motivate direct experimental tests of estradiol's role in the temporal structure of mate-choice updating.

## CRediT authorship contribution statement

**Alexander T. Baugh:** Writing – review & editing, Writing – original draft, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Mark A. Bee:** Writing – review & editing, Resources, Methodology, Funding acquisition, Conceptualization. **Joseph Eyiolowope:** Writing – review & editing, Investigation. **Jingchen Hou:** Writing – review & editing, Investigation. **Caleb Kim:** Writing – review & editing, Investigation.

## Declaration of competing interest

The authors declare no competing interests.

## Acknowledgments

We thank S. Silver, N. Ambiel, S. Gray-Gaillard, N. LaScala, J. Lavigne and N. Lee for assistance with collecting frogs, behavioral testing and record keeping and S. Gupta for assistance in collecting frogs. We thank J. Moriarty and the Three Rivers Park District and Michael Goodnature and Ramsey County Parks and Recreation for after-hours access to frog ponds. Two anonymous reviewers improved this manuscript and we are thankful for their contributions. Funding was provided by the Swarthmore College Research Fund to ATB and the National Science Foundation to MAB (IOS 1452831). Animal collections were made under Special Permits 23543 and 28347 from the State of Minnesota Department of Natural Resources. This study was approved by the Institutional Animal Care and Use Committee at the University of Minnesota (Protocol 1701-34456 A, approved 03-March-2017).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygcen.2026.115019.

## Data availability

Data file is available in SM7.

## References

AmphibiaWeb, 2024. *Hyla chrysoscelis*: Cope's Gray Treefrog. https://amphibiaweb.org/species/768 (accessed 22 July 2026).

Bastien, B., Farley, G., Ge, F., Malin, J.S., Simon-Plumb, C.L., Pulley, D.M., Yang, C., Baugh, A.T., 2018. The waiting and mating game: condition dependent mate sampling in female gray treefrogs (*Hyla versicolor*). Front. Ecol. Evol. 6, 140. https://doi.org/10.3389/fevo.2018.00140.

Baugh, A.T., Akre, K.L., Ryan, M.J., 2008. Categorical perception of a natural, multivariate signal: mating call recognition in túngara frogs. Proc. Natl. Acad. Sci. USA 105, 8985–8988.

Baugh, A.T., Gall, M.D., Silver, S.C., Bee, M.A., 2021. Moderately elevated glucocorticoids alter mate choice behavior in a treefrog. Horm. Behav. 130, 104950. https://doi.org/10.1016/j.yhbeh.2021.104950.

Bee, M.A., 2007. Sound source segregation in grey treefrogs: spatial release from masking by the sound of a chorus. Anim. Behav. 74, 549–558. https://doi.org/10.1016/j.anbehav.2006.12.012.

Bee, M.A., 2008. Parallel female preferences for call duration in a diploid ancestor of an allotetraploid treefrog. Anim. Behav. 76, 845–853. https://doi.org/10.1016/j.anbehav.2008.01.029.

Bee, M.A., Schwartz, J.J., 2009. Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. J. Acoust. Soc. Am. 126 (5), 2788–2801. https://doi.org/10.1121/1.3224707.

Bee, M.A., Swanson, E.M., 2007. Auditory masking of anuran advertisement calls by road traffic noise. Anim. Behav. 74, 1765e1776. https://doi.org/10.1016/j.anbehav.2007.03.019.

Carr, J.A., 2024. Stress and reproduction in amphibians. In: Norris, David O., Lopez, Kristin H. (Eds.), Hormones and Reproduction of Vertebrates, vol. 2, pp. 121–150 (second edition), Academic Press. ISBN 9780443160202. https://doi.org/10.1016/B978-0-443-16020-2.00002-4.

Chakraborty, M., Burmeister, S.S., 2009. Estradiol induces sexual behavior in female túngara frogs. Horm. Behav. 55, 106–112. https://doi.org/10.1016/j.yhbeh.2008.09.001.

Chianese, R., Ciaramella, V., Fasano, S., Pierantoni, R., Meccariello, R., 2013. Kisspeptin receptor, GPR54, as a candidate for the regulation of testicular activity in the frog *Rana esculenta*. Biol. Reprod. 88, 1–11. https://doi.org/10.1095/biolreprod.112.103515.

Couch, C.E., Colvin, M.E., Chitwood, R.L., Peterson, J.T., Schreck, C.B., 2022. Scope of the cortisol stress response in Chinook salmon during maturation. Gen. Comp. Endocrinol. 323–324, 114030. https://doi.org/10.1016/j.ygcen.2022.114030.

Gall, M.D., Bee, M.A., Baugh, A.T., 2019. The difference a day makes: breeding remodels hearing, hormones and behavior in female Cope's gray treefrogs (*HylaHyla chrysoscelis*). Horm. Behav. 108, 62–72. https://doi.org/10.1016/j.yhbeh.2019.01.001.

Gerhardt, H.C., Huber, F., 2002. Acoustic Communication in Insects and Anurans: Common Problems and Diverse Solutions. University of Chicago Press, Chicago.

Leary, C.J., Baugh, A.T., 2020. Glucocorticoids, male sexual signals, and mate choice by females: implications for sexual selection. Gen. Comp. Endocrinol. 288, 113354. https://doi.org/10.1016/j.ygcen.2019.113354.

Lin, H., Yuan, K.M., Zhou, H.Y., Bu, T., Su, H., Liu, S., Zhu, Q., Wang, Y., Hu, Y., Shan, Y., Lian, Q.Q., Wu, X.Y., Ge, R.S., 2014. Time-course changes of steroidogenic gene expression and steroidogenesis of rat Leydig cells after acute immobilization stress. Int. J. Mol. Sci. 15, 21028–21044. https://doi.org/10.3390/ijms151121028.

Lynch, K.S., Wilczynski, W., 2008. Reproductive hormones modify reception of species-typical communication signals in a female anuran. Brain Behav. Evol. 71, 143–150. https://doi.org/10.1159/000111460.

McConnachie, S.H., Cook, K.V., Patterson, D.A., Gilmour, K.M., Hinch, S.G., Farrell, A.P., Cooke, S.J., 2012. Consequences of acute stress and cortisol manipulation on the physiology, behavior, and reproductive outcome of female Pacific salmon on spawning grounds. Horm. Behav. 62, 67–76. https://doi.org/10.1016/j.yhbeh.2012.05.001.

Moore, F.L., Miller, L.J., 1984. Stress-induced inhibition of sexual behavior: corticosterone inhibits courtship behaviors of a male amphibian (*Taricha granulosa*). Horm. Behav. 18, 400–410. https://doi.org/10.1016/0018-506x(84)90026-6 (PMID: 6097527).

Nityananda, V., Bee, M.A., 2011. Finding your mate at a cocktail party: frequency separation promotes auditory stream segregation of concurrent voices in multi-species frog choruses. PLoS One 6 (6), e21191. https://doi.org/10.1371/journal.pone.0021191.

Orchinik, M., Murray, T.F., Moore, F.L., 1991. A corticosteroid receptor in neuronal membranes. Science 252 (5014), 1848–1851. https://doi.org/10.1126/science.2063198 (PMID: 2063198).

R Core Team, 2024. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Ratcliff, R., McKoon, G., 2008. The diffusion decision model: theory and data for two-choice decision tasks. Neural Comput. 20, 873–922. https://doi.org/10.1162/neco.2008.12-06-420.

Redish, A.D., 2016. Vicarious trial and error. Nat. Rev. Neurosci. 17, 147–159. https://doi.org/10.1038/nrn.2015.30.

Remage-Healey, L., Jeon, S.D., Joshi, N.R., 2013. Recent evidence for rapid synthesis and action of estrogens during auditory processing in a songbird. J. Neuroendocrinol. 25, 1024–1031. https://doi.org/10.1111/jne.12055.

Rodriguez-Santiago, M., Ruppert, A., Gall, M.D., Hoke, K., Bee, M.A., Baugh, A.T., 2024. In your CORT: corticosterone and its receptors in the brain underlie mate choosiness in female Cope's gray treefrogs (*Hyla chrysoscelis*). Horm. Behav. 159, 105477. https://doi.org/10.1016/j.yhbeh.2024.105477.

Sapolsky, R.M., Romero, L.M., Munck, A.U., 2000. How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and preparative actions. Endocr. Rev. 21, 55–89. https://doi.org/10.1210/edrv.21.1.0389.

Shearer, F.J.G., Wyrwoll, C.S., Holmes, M.C., 2019. The role of 11β-hydroxy steroid dehydrogenase type 2 in glucocorticoid programming of affective and cognitive behaviours. Neuroendocrinology 109, 257–265. https://doi.org/10.1159/000499660.

Sisneros, J.A., Forlano, P.M., Deitcher, D.L., Bass, A.H., 2004. Steroid-dependent auditory plasticity leads to adaptive coupling of sender and receiver. Science 305, 404–407. https://doi.org/10.1126/science.1097218.

Tanner, J.C., Ward, J.L., Shaw, R.G., Bee, M.A., 2017. Multivariate phenotypic selection on a complex sexual signal. Evolution 71, 1742–1754. https://doi.org/10.1111/evo.13264.

Tolman, E.C., 1939. Prediction of vicarious trial and error by means of the schematic sowbug. Psychol. Rev. 46, 318–336. https://doi.org/10.1037/h0057054.

Uban, K.A., Rummel, J., Floresco, S.B., Galea, L.A.M., 2012. Estradiol modulates effort-based decision making in female rats. Neuropsychopharmacology 37, 390–401. https://doi.org/10.1038/npp.2011.176.

Ward, J.L., Buerkle, N.P., Bee, M.A., 2013. Spatial release from masking improves sound pattern discrimination along a biologically relevant pulse-rate continuum in gray treefrogs. Hear. Res. 306, 63–75. https://doi.org/10.1016/j.heares.2013.09.006.

Wells, K.D., 2007. The Ecology and Behavior of Amphibians. University of Chicago Press, Chicago.

Whirledge, S., Cidlowski, J.A., 2017. Glucocorticoids and reproduction: traffic control on the road to reproduction. Trends Endocrinol. Metab. 28, 399–415. https://doi.org/10.1016/j.tem.2017.02.005.

Wingfield, J.C., Sapolsky, R.M., 2003. Reproduction and resistance to stress: when and how. J. Neuroendocrinol. 15, 711–724. https://doi.org/10.1046/j.1365-2826.2003.01033.x.