Ribbit and Racket: The Effects of Anthropogenic Noise on Anuran Calling Behavior in the Chicago Region

BY

NOLAN K. BIELINSKI
B.S., Rhodes College, 2009

THESIS

Submitted as a partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biological Sciences
in the Graduate College of the
University of Illinois at Chicago, 2020

Chicago, Illinois

Defense Committee:

David H. Wise, Chair and Advisor
Emily Minor
Alan Molumby
Joseph Milanovich, Loyola University, Chicago
Liam Heneghan, DePaul University, Chicago

This dissertation is dedicated to my wife, Anna Rose.

ACKNOWLEDGEMENTS

Simon Sanders, and Zainab Mughal. A big thanks goes out to Dr. Jennifer Pajda De La O, who helped me tackle experimental and statistical issues.

Alexis Smith and Emily Portratz also played a significant role for me during my time at UIC, as we expanded our breadth of ecological inquiry with side projects that will hopefully one day get published.

Finally, I would like to thank my family for providing the stability needed outside of school for me to accomplish such a long and arduous task. To my parents and in-laws, thank you for having patience, feeding me, and leaving me alone (for the most part) when I was on the computer. For my brand-new wife, thank you for providing both motivation and distraction at exactly the right time when I needed it throughout my graduate school experience.

NKB

CONTRIBUTION OF AUTHORS

Chapter 1 of this dissertation is a published manuscript with several authors (Nolan Bielinski, Jennifer Pajda-De La O, Adrianna Gorniak, and David Wise). Each author's contribution was as follows: Nolan Bielinski came up with the research topic, experimental design, conducted all field work, oversaw soundscape analysis, conducted the statistical analysis, and wrote the publication. Jennifer Pajda-De La O helped with the statistical approach and made revisions. Adrianna Gorniak performed most of the soundscape analysis using computer software. David Wise provided guidance on ecological theory, helped review the statistics and previous drafts, and critiqued the discussion. The other two chapters are work created by Nolan Bielinski, with statistical guidance from Dr. Jennifer Pajda De La O and David Wise, and computer analysis conducted by undergraduate researchers, overseen by Nolan Bielinski.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

III. ANTHROPOGENIC NOISE PLAYBACK EXPERIMENTS ON URBAN BULLFROG AND GREEN FROG INDIVIDUALS

LIST OF ABBREVIATIONS

ARD Automated recording device

dB Decibel

FN False negative

FP False positive

LM Linear model

LMEM Linear mixed effects model

PCoA Principal coordinates analysis

PERMANOVA Permutational multivariate analysis of variance

CI Confidence interval

SD Standard deviation

SPL Sound pressure level (also known as power)

TP True positive

UPGMA Unweighted Pair-Group Methods using arithmetic Averages clustering

SUMMARY

Anthropogenic noise represents a major obstacle to anurans in urban landscapes because they rely heavily on vocalizations to find breeding grounds, choose mates, and settle territorial disputes. A large proportion of urban frogs may be affected by noise, as past research suggests that road-side wetlands are a major contribution to frog biodiversity and migration ability in developed landscapes. Thus, there is a need for improved monitoring and research techniques for studying anurans in urban habitats so we can better understand and mitigate the effects of noise on the ecology, behavior, and physiology of frogs.

This dissertation uses soundscape methods to achieve the following: (Chapter 1) Improve how automated data collection is used in urban frog call surveys. (Chapter 2) Conduct community-level noise playback experiments to investigate how patterns in call attributes change across species, genera, and natural history characteristics. (Chapter 3) Compare results for individual-level versus population-level noise playback experiments.

In Chapter 1, we address the shortcoming that automated frog call detection remains untested in the literature. In urban landscapes, higher levels of anthropogenic noise may reduce the efficacy of recognizers (software programs that can automatically identify calls). We built and evaluated the performance of standard "full" recognizers versus new "narrow-banded" recognizers that focus on a limited frequency range within its target call. These included three recognizers that focused on the lower, middle, and upper portions of the calling frequency of two species, the American bullfrog (*Lithobates catesbeianus*) and the green frog (*L. clamitans*). We checked the outputs of the recognizers to determine whether identifications were correct.

For both species, we found that narrow-banded recognizers best avoided anthropogenic noise. A narrow-banded recognizer had the highest true positive rate (TPR) at two of four sites

for the bull frog. Using the optimum recognizer per site, we increased the TPR (from 0.281 to 0.558) while decreasing the false negative rate (FNR; from 0.719 to 0.450). For the green frog, the highest TPRs were from narrow-banded recognizers at all five sites. Using narrow-banded recognizers, we increased the TPR compared to that of the full recognizer (from 0.146 to 0.345) while reducing the FNR (from 0.835 to 0.646). Finally, we found that the number of calls identified by the narrow-banded recognizers correlate with the actual number of calls recorded, meaning that the recognizer output serves as a dependable proxy for calling activity. Using narrow-banded recognizers could prove useful for automatically collecting chorusing data at urban sites.

In Chapter 2, the crux of the dissertation, we used noise playback experiments to investigate patterns of frog calling behavior at the population-level. Past literature investigating the effects of anthropogenic noise on frog calling behavior shows no overarching pattern. These varying results may be a consequence of disjoint experimental approaches with single target species. This underscores the need for research that investigates patterns across communities. We attempt to fill this need by conducting the same playback experiments on an urban community of nine frog species from the Chicago region.

We conducted 155 anthropogenic noise playback experiments on chorusing males at 21 breeding ponds. Our target species were the bullfrog, green frog, wood frog (*L. sylvaticus*), Cope's gray treefrog (*Hyla chrysoscelis*), eastern gray treefrog (*Hyla versicolor*), spring peeper (*Pseudacris crucifer*), western chorus frog (*Pseudacris triseriata*), northern cricket frog (*Acris crepitans*), and American toad (*Anaxyrus americanus*). For every species, we conducted two types of playback experiments. The "Noise Source Experiment" used three different 90 dB noise

treatments: car, airplane, and train. The "Noise Power Experiment" used the same car noise recording, but with peak amplitudes at 90 dB, 80 dB, and 70 dB.

We measured call rate, dominant frequency, call length, and pulse rate from calls during each treatment. We first compared control treatments to pooled noise treatments from the Noise Source Experiment. Next, if a response was found, the separate noise treatments (car, airplane, and train) were compared against each other. Finally, we compared the treatments from the Noise Power Experiment against the control.

We predicted that frogs would respond to noise in the following ways: (1) Frogs will reduce their call rate to avoid wasting energy when calling in unfavorable conditions for females to assess male suitors. (2) Frogs will shift their dominant frequencies higher to avoid spectral interference of the low-frequency noise. We should see a clearer signal in the low-frequency callers because they will experience the most spectral overlap. (3) We expect frogs to shorten their calls. Changes should be most prominent in the car traffic treatments because the interval between cars much shorter than that between trains and airplanes. (4) Observed changes should be more extreme in louder noise treatments. (5) Species living in permanent ponds will have greater changes in call rate and length compared to others because they have longer breeding periods and less pressure to breed at the moment. (6) Congeneric species should have more similar responses compared to heterogeneric species because call modifications should be genetically influenced.

Our multivariate analyses (PERMANOVAs, a Principal Coordinate Analysis, and clustering) did not reveal any patterns other than grouping by genus and species. Our data did not separate by hydrology preference, breeding season, or length of breeding season. We did not find a treatment $\times$ species interaction. Within species, there was also no grouping by treatment, a

consequence of only having one call attribute (call rate) consistently changing in response to anthropogenic noise.

Our univariate analyses (effect sizes with 95% CIs) revealed that eight of nine species reduced their call rate in response to noise. Moreover, we found that the Spring Peeper, the only species to not reduce its call rate, was also the only species that raised its dominant frequency, suggesting that these two strategies are dichotomous choices to solve the same problem. There was no noticeable pattern in the change of call length or pulse rate across species. There was also no evidence to that frogs were able to perceive different anthropogenic noise sources and act differently as a response.

In general, the community had a larger magnitude of responses during louder noise treatments. There were seven instances where an effect was seen at the 90 dB level only, five instances where a threshold was seen starting at the 80 dB treatment, and zero instances beginning at the 70 dB treatment. Thus, noise may only be affecting calling behavior in the loudest habitats, such as road-side ditches and retention ponds.

Our study clearly provides much-needed data on a community of species. Our study was also the only to compare three types of anthropogenic noise. Future playback studies should (1) target new communities to ask questions about taxonomy and natural history, and to test more generalizable hypotheses and (2) compare responses from urban versus natural populations of the same species to explore whether urban species have developed new strategies for dealing with anthropogenic noise.

In Chapter 3 we conducted noise playback experiments on American bullfrog (n =6) and green frog (n =5) individuals to compare behavioral changes to our results from Chapter 2, and examine a new response variable, call power, which was not possible to collect during the

population-level experiments. We conducted this comparative study because there could be a difference between the aggregate calling response, and individual responses. Call attributes from individuals with lower call rates will be underrepresented in population-level studies. The playback experiment and analysis had the same structure as the "Noise Power Experiment" from Chapter 2. We predicted to find more effects and higher magnitudes of effects compared to the previous study.

Our major finding was a threshold effect for the reduction in call length starting at 80 dB for both species, which was not seen for either species in Chapter 2. In response to anthropogenic noise, it may be that some individuals choose to shorten their calls, but in aggregate the main strategy may be to maintain the usual call length and reduce their call rate. The differences in the other compared attributes were not major.

Bullfrog individuals made louder calls (+ 7.6 dB) during the Car 90 dB treatment only. It seems that bullfrogs raised their amplitude once the treatment surpassed their natural calling power level. We did not see an in calling power for the green frogs, although there was a slight positive trend between calling power and treatment power.

This comparison revealed that patterns of some individual-level attributes may not be distinguishable in a chorusing group. Aggregate patterns seen in population-level studies may not indicate the response across all individuals. We should be hesitant to extrapolate patterns or mechanisms of individual studies to a loud chorus environment. Ideally, researchers should utilize both chorus-level and individual-level methods on the same species. Chorus-level studies should be better at capturing the chaos that females will encounter when making mate choices, whereas individual-level studies should be better at identifying changes made by individual males in an attempt stand out amidst this chaos.

Overall, the findings from this dissertation improve upon popular methods (automated detection, and playback experiments) used to study frog calls and add to the growing dataset of anuran soundscape research with a rare community-wide study.

Chapter 1 is a publication from *Herpetological Conservation and Biology* and requires approval

to be in included in my dissertation. Here is an email of approval from the editor citing their

journal guidelines:

> The following statement is from page 2 of Instructions for Authors (attached): "all
> copyrights remain with the first author of a manuscript." This means that you have the
> copyright for the article and can do with whatever you want. Thus, you can give yourself
> permission to use the article as a chapter in your dissertation. The journal has no say.
> Let me know if you need anything further. Sincerely,
> David F. Bradford, Ph.D.
> Amphibian Section Co-Editor & Board of Directors
> Herpetological Conservation and Biology

Here is the full paragraph regarding copyrights from the journal's official Instructions for

Authors:

> Further, all copyrights remain with the first author of a manuscript. Authors are free to
> archive their manuscript on their personal or institutional web sites, as well as in
> manuscript repositories. Authors who chose to distribute their manuscript by other
> means should be aware that the relative popularity of their research may suffer when
> *Herpetological Conservation and Biology* compiles and publishes download statistics for
> impact factors.

1. IMPROVING AUTOMATED DETECTION OF FROG CALLS IN NOISY URBAN HABITATS USING NARROW-BANDED RECOGNIZERS

1.1 Abstract

Tracking behavioral and demographic changes of anuran populations in urban landscapes presents difficulties due to the high amount of noise interference from anthropogenic sources. In this study, we used Song Scope software to build narrow-banded recognizers that only cover a limited portion of the full spectral range of a call and tested if these recognizers can improve automated call-detection capabilities at noisy sites. We built recognizers for two species with naturally broad-spectrum calls, the Green Frog (*Lithobates clamitans*) and American Bullfrog (*L. catesbeianus*) and tested them at five noisy ponds in the suburbs of Chicago. Narrow-banded recognizers had greater percentages of true positives compared to full-spectrum recognizers. Classification indices used to assess call recognition efficacy showed that narrow-banded recognizers were more effective at all sites for the Green Frog, and at two sites for the American Bullfrog. High-frequency recognizers had 13% fewer errors caused by anthropogenic noise ($P < 0.01$) than other recognizers. Finally, for every recognizer, true positives standardized by the maximum daily value were highly correlated with the number of calls identified manually, indicating that automated detection data is an accurate proxy for the actual number of calls at noisy sites. For acoustic taxa, we recommend that scientists consider identifying broad-spectrum calls using narrow-banded recognizers to reduce detection problems associated with noise interference between anthropogenic noises and biotic acoustic signals.

Key Words: acoustic species; anthropogenic noise; bioacoustics; American Bullfrog; *Lithobates catesbeianus*; Green Frog; *Lithobates clamitans*; soundscape

1.2 INTRODUCTION

1.2.1 Effects of anthropogenic noise on frogs.

Noise pollution is a significant disturbance to wildlife in urban landscapes. Anthropogenic noise, defined as noise from human technology and activity, represents a major obstacle to anurans because they rely heavily on vocalizations to find breeding grounds, choose mates, and settle territorial disputes (Gerhardt 1994). Anthropogenic noise is a pervasive edge effect characterized by low-frequency intermittent buzzing or humming (Pijanowski *et al.* 2011a), which may prominently overlap the frequencies of biological sounds including some anuran mating calls (Bee & Swanson 2007). Many natural habitats in urban regions (even interiors of preserves) experience significant noise pollution, with the main sources being traffic from cars, airplanes, and trains. For example, highway noise from 150 m away can reach 70 dB (Warren *et al.* 2006). Train noise, although more infrequent, can produce noise up to 90 dB within 10 m (Nolan Bielinski, pers. obs.). Areas under airport flyover paths can reach 74 dB (Warren *et al.* 2006).

Consequently, frogs living in urban landscapes experience the effects of a drastically altered soundscape. Traffic noise exposure has been shown to increase corticosterone production and decrease antimicrobial peptide production (Tennessen *et al.* 2018), change the calling attributes of males (Sun & Narins 2005; Parris *et al.* 2009; Cunnington & Fahrig 2010), limit male participation in chorusing (Kaiser *et al.* 2011), and alter the response of females to male calls (Bee & Swanson 2007). Looking across all studies investigating the effects of anthropogenic noise on frog calling behavior, no explicit pattern emerges, suggesting that responses may be species- or region-specific. This highlights the need for further monitoring and improved research techniques for studying anurans in noisy habitats. Future studies would

3

benefit from increased sample sizes and improved efficiency of acoustic analysis through the implementation of the latest soundscape software techniques.

1.2.2 Frog call surveys

Broad-scale, manual (in-person) calling surveys are often used to assess the vulnerability of frog species (Williams *et al.* 2013) and population trends over time (Gibbs & Breisch 2001) but are inadequate in many ways. For example, the North American Amphibian Monitoring Program protocol advises volunteers to collect data for 2 h total, beginning 30 min after sunset, which systematically ignores post-midnight calling activity (Bridges & Dorcas 2000; US Geological Survey. 2019. North American Amphibian Monitoring Program. Available from https://www.usgs.gov/centers/pwrc/science/north-american-amphibian-monitoring-program?qt-science_center_objects=0#qt-science_center_objects [Accessed 2 August 2019]). Moreover, volunteers dedicate only 3 min to each listening point. Considering that volunteers may disturb frogs when moving to each location, more time may be needed for the frog community to revert to natural calling behavior (Crouch & Paton 2002). In general, manual surveys also risk having high instances of misidentification and omission of species, especially when dealing with faint, infrequent, or scattered calls, or during nights when multiple species are calling simultaneously (Genet & Sargent 2003). Consequently, monitoring programs may significantly bias abundance (Lotz & Allen 2007) or presence/absence data (Genet & Sargent 2003).

Most deficiencies associated with manual surveys can be resolved using automated recording devices (ARDs). ARDs can be used to collect data throughout the entire night or even entire 24-h cycles, which increases the chance of recording rare (MacLaren *et al.* 2017) and cryptic species (Engbrecht 2011) and provides a more accurate picture of true calling phenologies (Bridges & Dorcas 2000). Breeding ponds also experience less disturbance once

recording devices are installed (Acevedo & Villanueva-Rivera 2006). After data collection in

the field, recorded calls can be reviewed by smaller cohorts and allow for both visual (spectral)

and auditory review to identify species in a more comprehensive and consistent manner, and

problematic recordings can be replayed for clarification. Furthermore, soundscape software can

be used to detect and analyze changes in several call characteristics such as dominant frequency,

pulse rate, and sound pressure level, that may be indiscernible without equipment (Cocroft &

Ryan 1995). These quantifiable attributes of the soundscape allow researchers to test hypotheses

on how noise may alter frog communities by changing favorable calling strategies.

The final benefit of ARDs is the ability to filter through large datasets to automatically

identify calls using digital signal processing algorithms called "recognizers." Recognizers look

for patterns within the time-frequency state-space to isolate patterns matching the call structure

of a target species (Wildlife Acoustics. 2011. Song Scope bioacoustics software version 4.0

documentation. Available from https://www.wildlifeacoustics.com/ [Accessed 2 August 2019]).

Using recognizers dramatically reduces the time required to convert recorded

soundscapes into quantifiable data. This pronounced increase in scale of data collection and

analysis afforded by recognizers could be particularly beneficial when attempting to document

rare species, catching the onset of breeding from explosive breeders, or monitoring changes in

range or behavior of species in response to climate change or the spread of disease across a

region.

1.2.3 Our study

Recognizer performance in noisy soundscapes remains untested in the literature. The

purpose of this study was to investigate the capabilities of recognizers with different frequency

ranges to identify frog calls in noisy habitats to advance acoustic monitoring efforts for researchers in urban soundscapes.

There are several commercially available soundscape software programs that implement recognizer-like functionality, including Song Scope, Kaleidoscope, MonitoR, and RavenPro; a comparison of these programs is offered by Knight *et al.* (2017). We chose to use Song Scope because it was rated as the best performing program (Knight *et al.* 2017), has a relatively low learning curve, is used in many soundscape studies, and is available for free.

The efficacy of recognizers has already been investigated with promising results in areas where anthropogenic noise is minimal (Waddle *et al.* 2009; Eldridge 2011; Brauer *et al.* 2016; Crump and Houlahan 2017; MacLaren *et al.* 2017), but it is unclear whether recognizers are useful in noisy environments. For example, the best strategy for implementing automated detection may differ greatly between sites with pristine versus degraded soundscapes. Thus, it would be beneficial for researchers to know if they should alter the parameters of their recognizer depending on the noise profile of their sites, regardless of the soundscape software being used or the taxon being studied.

We conducted this research in the suburbs of Chicago. In urban landscapes like the Chicago region, few natural habitats are completely shielded from noise pollution (USDOT. 2019. National Transportation Noise Map. Available from https://maps.bts.dot.gov/arcgis/apps/webappviewer/index.html?id=a303ff5924c9474790464cc0e9d5c9fb [Accessed 2 August 2019]). In Chicago the gridded streets and large highways bisect green spaces, airplanes crowd the skies from the nation's 2nd and 25th busiest airports (Federal Aviation Administration 2019), and train traffic is high enough to designate Chicago as busiest

rail freight gateway in the US. We specifically chose our study sites based on high noise levels

from the above-mentioned sources.

We recorded Green Frog (*Lithobates clamitans)* and American Bullfrog *(L. catesbeianus*;

hereafter, bullfrog*)* calls to test these soundscape methods because both species are widespread

in urban areas. Additionally, these species have calls with low dominant frequencies, meaning

that spectral interference with anthropogenic noise should be greater for them than other local

species (Phillips *et al.* 1999; Pijanowski *et al.* 2011b).

For both species, we created and compared multiple recognizers with different spectral

parameters. Additionally, we characterized and measured background noise from samples of

true positives (TPs), false positives (FPs), and false negatives (FNs) to describe the specific

strengths and weaknesses of each recognizer. Finally, we compared manual counts to the

respective recognizer output to determine if recognizer data can be used as a reliable proxy for

actual call rate (defined here as the number of calls over time) at noisy sites.

1.3 MATERIALS AND METHODS

1.3.1 Acoustic surveys

We deployed ARDs (SongMeter model SM4, Wildlife Acoustics, Maynard,

Massachusetts, USA) at five breeding sites (A - E) between 25 June 2016 and 15 August 2016 on

forest preserves and private land with ponds in the suburbs of Chicago (Supplemental Table S1).

We secured the ARDs to trees approximately 1 m away from the edge of each pond. We

programmed ARDs to record the first 5 min of every hour from 1800 to 0100 for four nights.

Throughout the study, the ARDs did not require any battery or memory card changes, which

means we avoided all physical disturbance around the pond edge.

1.3.2 Study sites

Site A is a road-side permanent pond located in Cherry Hill Woods (Cook County Forest Preserve). The surrounding area has a mixture of forest, open woodlands, and savanna. It is dominated by oak (*Quercus spp.*) and hickory (*Carya spp.*). Site B is a large slough located in the center of Wolf Road Woods (Cook County Forest Preserve). This area has rolling hills and a mixture of forest and open woodlands with interspersed ephemeral pools. It is dominated by oak and hickory. Sites C and D are artificial ponds located on private property in Palos Hills, Illinois. The neighborhood has scattered homes adjacent to patches of oak-dominated forest on rolling hills. Sites C and D are close to each other (118 m) so they share the same anthropogenic noise sources. However, calls from one pond could not be heard or detected at the other, meaning that recorded calls from these sites were from different individuals. Site E is a small semi-permanent pond located in Van Patten Woods (Lake County Forest Preserve). It is bordered on one side by a train track, and on all other sides by oak forest.

1.3.3 Noise profiles at sites

During installation and removal of the ARDs, we spent an hour collecting data on sound levels using a sound pressure level meter (Model DS-HWCJ04, Koolertron, Shenzhen, China). We recorded the distance to noise sources and the sound pressure level of ambient noise (general ongoing sounds) and acute noise, defined here as distinct instances of punctuated anthropogenic sounds from cars, trucks, motorcycles, trains, and airplanes (Table 1). Every site is close to anthropogenic noise sources, with site B being the farthest away from the nearest road or train track, at 402 m. The measured noise levels from acute instances of anthropogenic noise reflect these proximities. Putting these noise levels in a biological context, at 1 m Green Frogs can call at 84 dB (Bee & Perrill 1996) and bullfrogs call at 80 dB (Simmons 2004). Using the inverse

8

square law for sound intensity, this is an equivalence of 64 dB and 60 dB, respectively, at 10 m, approximately the distance where females can interpret multiple calls from potential mates, based on male territory size (Wells 1977). In our recordings we also encountered acoustic signals from insects including crickets (family Gryllidae), dog-day cicadas (*Neotibicen canicularisi*), and katydids (family Tettigoniidae), and vocalizations from birds including Red-winged Blackbirds (*Agelaius phoeniceus*), Great Blue Herons (*Ardea herodias*), and several others. We did not identify insect and bird species for our noise categorization analysis. Instead, we classified them as biological noise.

1.3.4 Supplemental recordings

During the same field season, we collected additional training data (i.e., data required to build recognizers; see below) from study sites via supplemental recordings of extra nights, and from Hegewisch Marsh (Chicago Park District) and Hickory Creek (Will County Forest Preserve).

1.3.5 Recognizer development

We performed all soundscape data analysis on Song Scope bioacoustics software (version 4.1.5; Wildlife Acoustics, Inc., Concord, Massachusetts, USA). To build recognizers, the software must be fed training data, where one manually identifies confirmed signals from a target species within a spectrogram and collects them in a Build Recognizer page, and then recognizer parameters can be adjusted and filtered to properly match the vocalizations in the training data (Wildlife Acoustics. 2011. *op. cit.*). Song Scope then uses hidden Markov models to construct a model call to compare to candidate vocalizations in new recordings. More detail on recognizers is available elsewhere (Agranat, I. 2009. Automatically identifying animal species from their vocalizations. Wildlife Acoustics, Concord, Massachusetts, USA. Available

Normally, frog calls are structurally simple enough to cover using a single recognizer. Most calls lack the spectral and temporal complexities that may warrant splitting up a call into multiple sub-signals, which is sometimes implemented to identify more complex avian songs (Gelling 2010). The Green Frog and the bullfrog have short, simple breeding calls usually consisting of one syllable, with most of the acoustic power at a low frequency. However, both calls have simultaneous medium- and high-frequency harmonics, as seen by the large signal range displayed along the y-axis of a spectrogram (Fig. 1A & B). We took advantage of these naturally large spectral ranges by building four recognizers for each species that focused on a different spectral range of their call (Table 2). We considered frequency range as the most important parameter because anthropogenic noise often overlapped part of the natural frequency range of our target calls (for a full list of every parameter, see Appendix A).

Considering the high levels of anthropogenic noise at our sites, calling males experienced spectral interference in the low frequency range. Moreover, the soundscapes also included avian and insect signals, which overlap with frog calls at higher frequencies. Thus, for the Green Frog and the bullfrog, we built recognizers that encompassed the full frequency range (a conventional recognizer), and three narrow-banded recognizers covering the highest portion, a middle portion, and the lowest portion of the frequency range. We named Green Frog recognizers 1-Full, 2-High, 3-Middle, and 4-Low; we named bullfrog recognizers 5-Full, 6-High, 7-Middle, and 8-Low (Fig. 1C & D).

1.3.6 Recognizer analysis

We investigated whether the narrow-banded recognizers could outperform the conventional full recognizers in noisy environments. For each 5 min recording (of 200 recordings total), we conducted a manual count of the number of calls from Green Frogs and bullfrogs, providing a reference, or a condition positive with which we could compare recognizer results. Next, for recordings with Green Frog calls (112 total) we ran recognizers one through four, and for recordings with bullfrog calls (98 total) we ran recognizers five through eight. We counted TPs (calls correctly identified), FPs (sounds that were wrongly identified as calls), and FNs (missed calls). We chose to not count true negatives (correctly ignored noises) because the constancy of anthropogenic noises makes distinguishing independent units of noise very difficult. We recorded the amount of time it took make these counts.

To determine if recognizers produced accurate estimations of the actual call rate at a site, we compared relationships between TPs and manual counts over the same time-series. To do this we standardized the TPs and manual counts by their respective maximum values for each day by site. We then ran a correlation analysis between the standardized manual count and TP values per recognizer. Because our data did not meet the normality assumptions for a Pearson correlation (most likely because calling behavior varied greatly with the weather from each experimental night), we used a Kendall rank correlation test to calculate Tau-b, a statistic that tests the strength of association in ranked data while also making adjustments for ties (McLeod 2011). For each correlation, x represents the standardized manual counts from every date by hour, and y represents the standardized TPs from every date by hour.

To investigate how background noise affects recognizer accuracy, we took a random subsample of a maximum of five TPs, five FPs and five FNs from every hour during the first night. Then, for each sample, we categorized the background noise type as either biological,

geological, anthropogenic, no noise, or recorder error. To test if the type of background noise affects the ability of recognizers to correctly identify calls, we ran a PERMANOVA of the Bray-Curtis dissimilarities of noise category counts for TPs, FPs, and FNs across recognizers, stratified by site (Oksanen *et al.* 2019). We also ran a difference in proportions test to determine if certain recognizers were less impeded by anthropogenic noise interference. Next, we measured the power (peak-to-peak voltage) at each of these samples and calculated 95% confidence intervals to compare noise volume levels across recognizers.

Finally, at every site, we assessed the efficacy of each recognizer by calculating the following indices:

True positive rate (TPR) = True Positives / Manual Count

Precision (PPV) = True Positives / (Manual Count + False Positives)

False Negative rate (FNR) = False Negatives / Manual Count

False discovery rate (FDR) = False Positives / (False Positives + True Positives)

These are common indices used in classification scenarios. In this case, they are describing the ability of a recognizer to correctly identify a sound in the recording as a frog call of a target species.

We created our graphs in Tableau 2019.1.3 (Tableau Software, Inc., Seattle, Washington, USA). We conducted the statistical analyses in R version 3.5.0. (R Core Team 2019).

1.4 RESULTS

1.4.1 Recorded audio data

For site A a programing error resulted in the loss of data for the 1800 hour. After this omission, we recorded 156 audio files totaling 780 min from our five sites.

1.4.2 Green Frog recognizer performance

From all sites we manually identified 11,608 Green Frog calls. Each of the four Green

Frog recognizers varied in performance based on the research site (Fig. 2). When comparing the

ratio of TPs to the manual counts per recognizer, the conventional recognizer 1-Full was never

the top-identifying recognizer.

For each recognizer, we ran separate Kendall's Tau-b correlations on standardized manual

counts and TPs. All recognizer TP's were correlated with actual calling taking place. The level

of correlation was highest for recognizer 4-Low (Tau-b = 0.721, P < 0.001) followed by 3-

Middle (Tau-b = 0.712, P < 0.001), 1-Full (Tau-b = 0.607, P < 0.001), and 2-High (Tau-b =

0.538, P < 0.001).

Next, we compared the classification indices between recognizers (Table 3). The best-

rated value (highest value for TPR and PPV and lowest for FNR and FDR) in each index.

Depending on the site, the top performing recognizers were either 2-High or 3-Middle. It is also

evident from this table that 4-Low across all sites had too many FPs to be considered effective,

considering its remarkably high FDR values.

Using the optimum recognizer per site (Table 3), we would achieve a mean TPR of

0.345, compared to a mean TPR of 0.146 from the conventional recognizer 1-Full alone.

Furthermore, this optimum set of recognizers would produce a mean FNR of 0.646, compared to

a mean FNR of .835 from 1-Full alone.

1.4.3 Noise description during Green Frog recognizer output

Site B had the highest amount of biological background noise recorded, which was from

intense insect chorusing. Additionally, sites D and E had higher proportions of geophonies

primarily from wind (Appendix B). PERMANOVA analysis for Green Frog recognizer output

indicated a weak effect of background noise on recognizer identification capability (P[Pseudo-$F_{3,16}$] $= 0.025$, $R^2 = 0.11$). There were no general patterns between noise and TP, FP, and FN classifications within recognizers (Supplemental Fig. S1).

All narrow-banded recognizers experienced significantly lower average noise volume levels than the conventional recognizer, with 3-Middle being the lowest, and 2-High being the second lowest (Fig. 3). When considering recognition errors across sites (Fig. 4), the percentage of noise classified as anthropogenic is lowest in recognizer 2-High. A difference in proportions test showed that there were 13% fewer errors attributed to anthropogenic noise for high-frequency recognizers ($P < 0.01$), demonstrating that recognizers that avoid low frequencies can reduce the amount of anthropogenic noise interference that cause identification errors.

1.4.4 Bullfrog recognizer performance

From all sites we manually identified 3,695 bullfrog calls. Site C was not used in the analysis, as no bullfrogs were present. Each of the four bullfrog recognizers varied in performance based on the research site (Fig. 5). When comparing the ratio of TPs to the manual counts per recognizer, narrow-banded recognizers were the top performer in 3 out of 4 sites.

As with the Green Frog, all bullfrog recognizer TPs were correlated with actual calling taking place. The level of correlation was highest for 7-Middle (Tau-b $= 0.688$, $P < 0.001$) followed by 8-Low (Tau-b $= 0.611$, $P < 0.001$), 5-Full (Tau-b $= 0.531$, $P < 0.001$), and 6-High (Tau-b $= 0.443$, $P < 0.001$).

According to the classification indices (Table 4), the conventional bullfrog recognizer (5-Full) had the best scores for two sites, D and E. Recognizer 7-Middle scored best for A, and 8-Low scored best for B. Using the optimum recognizer per site as determined by Table 4, we would achieve a mean TPR of 0.558, compared to a mean TPR of 0.281 from the conventional

recognizer 5-Full alone, and a mean FNR of 0.450 compared to a mean FNR of 0.719 from 5-Full alone.

1.4.5 Noise description during Bullfrog recognizer output

Site A shows a high level of biological noise interference in recognizers 5-Full and 6-High, as their frequency range overlapped regularly with bird calls (Appendix C). This was not the case for Green Frog recognizers because their call length parameter was shorter than the signals coming from birds, which disqualified them as potential target calls. A large amount of biological noise in site B interfered with the 6-High recognizer which is from high-frequency insect chorusing. PERMANOVA analysis for bullfrog recognizer output indicated a weak effect of background noise on recognizer identification capability (P[Pseudo-F$_{3,16}$] = 0.094, R^2 = 0.21745). There were no general patterns between noise and TP, FP, and FN classifications within recognizers (Supplemental Fig. S2).

All narrow-banded recognizers experienced significantly lower noise volume levels than the conventional recognizer, with 6-High being the lowest, and 7-Middle being the second lowest (Fig. 3). When considering recognition errors across all sites (Fig. 4), the percentage of noise classified as anthropogenic is lowest in recognizer 6-High. A difference in proportions test showed that anthropogenic noise produced 13% fewer errors for high-frequency recognizers ($P <$ 0.01), mirroring the results from the Green Frog recognizers.

1.5 DISCUSSION

1.5.1 Sub-optimal conditions

In contrast to previous studies using recognizers (Waddle *et al.* 2009; Eldridge 2011; MacLaren *et al.* 2017), this study implements soundscape techniques at sites that are particularly

15

noisy from multiple anthropogenic sources. Frequency overlap between the frog calls and noise was common (Appendix D).

It was a challenge to build recognizers for our target species because their calls are very short relative to other acoustic signals, for the Green Frog especially. Due to this, the recognizers have limited information to work with along the time axis, making it harder for the recognizers to identify patterns in the state-space of a spectrogram (Brauer *et al.* 2016). Some noises, like splashes in the water or snaps of branches, mimicked the shape of Green Frog calls on the spectrogram, resulting in some FPs (Appendix E). Moreover, Green Frog and bullfrog calls have broad frequency ranges and limited pulsations compared to other frogs. Therefore, the recognizers cannot home in on any specific pure tone or pulsation pattern, which are two major parameters incorporated into recognizer builds. This limited the overall accuracy (Tables 3 and 4) of the recognizers compared to the ideal scenario: a noiseless site with more structurally complex calls. Recognizer 4-Low had so many FPs, that reviewing the recognizer output essentially approached conducting a manual count over the full recording. Thus, any promising 4-Low output in terms of TPs was not considered too beneficial.

1.5.2 Error trade-off

The consequence of surveying noisy sites with ARDs is an increase in error rate, either through FPs or FNs depending on how the confidence parameters of the recognizer are adjusted. Thus, the recognizers built for this study were more error-prone compared to previous studies (Waddle *et al.* 2009; Eldridge 2011; MacLaren *et al.* 2017). It is important to adjust error trade-offs to match the goal of the study. For our study, we prioritized a reduction in FNs. Using Song Scope software, it is quick and easy to review identified calls and manually remove FPs from the recognizer data output, so a priority in automatically reducing FNs may generally be the

16

best route when building recognizers because FN's can only be checked by listening to full

recordings (Eldridge 2011).

1.5.3 Time investment

Recognizers represent a potentially noteworthy jump in data collection efficiency. For a

single recording the average time for the manual collection of data took 6 min 19 s, whereas the

automated collection plus manual removal of FPs took 3 min 28 s. Large jumps in efficiency

have also been recorded for previous recognizer studies (Knight *et al.* 2017; MacLaren *et al.*

2017). As alluded to by Waddle *et al.* (2009) and Eldridge (2011), a combined approach where

recognizers are used but then closely monitored manually may be best, especially for noisy sites.

1.5.4 Recognizer performance

According to our classification indices for Green Frog recognizers, by using either the

high- or medium-frequency recognizers we were able to improve TPR while simultaneously

reducing the error rate. Because 1-Full was never the top-identifying recognizer, the broad

frequency range of a conventional recognizer may be less effective in noisy environments. For

bullfrog recognizers, the conventional recognizer performed better, which may be attributable to

the fact that bullfrog calls are longer than Green Frog calls and therefore less limiting in the time

dimension, meaning that the recognizer has more state-space to work with to make proper

identifications.

For the Green Frog recognizers, we found that 3-Middle avoided the most noise

according to noise level, but 2-High best avoided anthropogenic noise specifically. For the

bullfrog recognizers, 6-High avoided the highest noise levels, and best avoided anthropogenic

noise. Therefore, instead of relying solely on conventional recognizers at noisy sites, choosing a

recognizer with appropriate frequency bands based off noise profile for each site can improve identification performance.

1.5.5 Noise profiles

  Our five study sites, although all chosen because of their high levels of anthropogenic noise, have different noise profiles. Interestingly, the results of our PERMANOVAs indicated only small differences in the type of noise interfering with each recognizer. This may be a result of certain noise categories having broader spectral effects on the soundscape. It is only when looking at anthropogenic noise alone that we could see that higher-frequency recognizers avoided interference better than lower-frequency recognizers.

1.5.6 Call rate estimation

        We found that using conventional and narrow-banded recognizers at noisy sites can produce results that correlate to the true calling behavior of frogs, meaning that TP values from recognizers can be used as a reliable proxy for call rate. For both species, the correlation was strongest for a narrow-banded recognizer. Thus, scientists and land managers looking to monitor and study changes of acoustic urban species should consider narrow-banded recognizers to collect data.

1.5.7 Conclusion

        With urbanization posing a major obstacle for amphibian biodiversity, more emphasis will likely be placed on tracking frog population changes in noisy areas. Considering the noise profile of a site by using narrow-banded recognizers could prove useful for this purpose. As seen in this study and previous research, anuran recognizers can be error-prone (Waddle *et al.* 2009; Engbrecht 2011) and building recognizers has some subjective parameters (e.g., complexity and resolution parameters). However, they offer such a substantial boost in time

efficiency that they can still be useful. Additionally, if the goal of the research is to compare calling effort across sites or over time, recognizer output can be used as a proxy because the overall pattern of call rate from automated surveys still resembles the true call pattern.

For these techniques to be most reliable, it is vital to choose a recognizer based on the noise profile specific to the site. This could mean building high-frequency recognizers when sites have low frequency interference, such as train noise, or building low-frequency recognizers for sites that have high frequency interference, such as insect chorusing. Our method of splitting up call frequency ranges using multiple recognizers may also be useful with avian species in noisy habitats, but a thorough investigation into this should be conducted. Only bird species with broad-frequency calls (e.g., Red-breasted Nuthatch, *Sitta canadensis*) are suitable for this technique.

1.6 Literature Cited

Acevedo, M.A., and L.J. Villanueva-Rivera. 2006. Using automated digital recording systems as effective tools for the monitoring of birds and amphibians. Wildlife Society Bulletin 34:211–214.

Bee, M.A., and S.A. Perrill. 1996. Responses to conspecific advertisement calls in the Green Frog (*Rana clamitans*) and their role in male-male communication. Behaviour 133:283–301.

Bee, M.A., and E.M. Swanson. 2007. Auditory masking of anuran advertisement calls by road traffic noise. Animal Behaviour 74:1765–1776.

Brauer, C.L., T.M. Donovan, R.M. Mickey, J. Katz, and B.R. Mitchell. 2016. A comparison of acoustic monitoring methods for common anurans of the northeastern United States. Wildlife Society Bulletin 40:140–149.

Bridges, A.S., and M.E. Dorcas. 2000. Temporal variation in anuran calling behavior: implications for surveys and monitoring programs. Copeia 587–592.

Cocroft, R.B., and M.J. Ryan. 1995. Patterns of advertisement call evolution in toads and chorus frogs. Animal Behaviour 49:283–303.

Crouch, W.B., and P.W.C. Paton. 2002. Assessing the use of call surveys to monitor breeding anurans in Rhode Island. Journal of Herpetology 36:185–192.

Crump, P.S., and J. Houlahan. 2017. Designing better frog call recognition models. Ecology and Evolution 7:3087–3099.

Cunnington, G.M., and L. Fahrig. 2010. Plasticity in the vocalizations of anurans in response to traffic noise. Acta Oecologica-International Journal of Ecology 36:463–470.

Eldridge, J.D. 2011. A comparison of current anuran monitoring methods with emphasis on the accuracy of automatic vocalization detection software. M.Sc. Thesis, Western Kentucky University, Bowling Green, Kentucky, USA. 122 p.

Engbrecht, N.J. 2011. The status of Crawfish Frogs (*Lithobates areolatus*) in Indiana, and a tool to assess populations. M.Sc. Thesis, Indiana State Unviersity, Terre Haure, Indiana. 78 p.

Federal Aviation Administration. 2019. Air traffic by the numbers. 40 p. https://www.faa.gov/air_traffic/by_the_numbers/.

Gelling, D. 2010. Bird song recognition using gmms and hmms. M.Sc. Thesis, University of Sheffield, Sheffield, England. 39 p.

Genet, K.S., and L.G. Sargent. 2003. Evaluation of methods and data quality from a volunteer-based amphibian call survey. Wildlife Society Bulletin 31:703–714.

Gerhardt, H.C. 1994. The evolution of vocalization in frogs and toads. Annual Review of Ecology and Systematics 25:293–324.

Gibbs, J.P., and A.R. Breisch. 2001. Climate warming and calling phenology of frogs near Ithaca, New York, 1900–1999. Conservation Biology 15:1175–1178.

Kaiser, K., D.G. Scofield, M. Alloush, R.M. Jones, S. Marczak, K. Martineau, M.A. Oliva, and P.M. Narins. 2011. When sounds collide: the effect of anthropogenic noise on a breeding assemblage of frogs in Belize, Central America. Behaviour 148:215–232.

Knight, E., K. Hannah, G. Foley, C. Scott, R. Brigham, and E. Bayne. 2017. Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs. Avian Conservation and Ecology 12:14. https://doi.org/10.5751/ACE-01114-120214.

Lotz, A., and C.R. Allen. 2007. Observer bias in anuran call surveys. Journal of Wildlife Management 71:675–679.

MacLaren, A.R., S.F. McCracken, and M.R. Forstner. 2017. Development and validation of automated detection tools for vocalizations of rare and endangered anurans. Journal of Fish and Wildlife Management 9:144–154.

McLeod, A.I. 2011. Kendall: Kendall rank correlation and Mann-Kendall trend test. Version 2.2. https://CRAN.R-project.org/package=Kendall.

Oksanen, J., F.G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P.R. Minchin, R.B. O'Hara, G.L. Simpson, P. Solymos, *et al.* 2019. Vegan: community ecology package. Version 2.5-4. https://CRAN.R-project.org/package=vegan.

Parris, K.M., M. Velik-Lord, and J.M.A. North. 2009. Frogs call at a higher pitch in traffic noise. Ecology and Society 14:1. http://www.ecologyandsociety.org/vol14/iss1/art25/.

Phillips, C.A., R.A. Brandon, and E.O. Moll. 1999. Field Guide to Amphibians and Reptiles of Illinois. Illinois Natural History Survey, Champaign, Illinois, USA.

Pijanowski, B.C., A. Farina, S.H. Gage, S.L. Dumyahn, and B.L. Krause. 2011a. What is soundscape ecology? An introduction and overview of an emerging new science. Landscape Ecology 26:1213–1232.

Pijanowski, B.C., L.J. Villanueva-Rivera, S.L. Dumyahn, A. Farina, B.L. Krause, B.M. Napoletano, S.H. Gage, and N. Pieretti. 2011b. Soundscape ecology: the science of sound in the landscape. Bioscience 61:203–216.

Simmons, A.M. 2004. Call recognition in the Bullfrog, *Rana catesbeiana*: generalization along the duration continuum. The Journal of the Acoustical Society of America 115:1345–1355.

Sun, J.W.C., and P.A. Narins. 2005. Anthropogenic sounds differentially affect amphibian call rate. Biological Conservation 121:419–427.

Tennessen, J.B., S.E. Parks, L. Swierk, L.K. Reinert, W.M. Holden, L.A. Rollins-Smith, K.A. Walsh, and T. Langkilde. 2018. Frogs adapt to physiologically costly anthropogenic noise. Proceedings of the Royal Society B-Biological Sciences 285: 20182194. http://dx.doi.org/10.1098/rspb.2018.2194.

Waddle, J.H., T.F. Thigpen, and B.M. Glorioso. 2009. Efficacy of automatic vocalization recognition software for anuran monitoring. Herpetological Conservation and Biology 4:384–388.

Warren, P.S., M. Katti, M. Ermann, and A. Brazel. 2006. Urban bioacoustics: It's not just noise. Animal Behaviour 71:491–502.

Wells, K.D. 1977. Territoriality and male mating success in the Green Frog (*Rana clamitans*). Ecology 58:750–762.

Williams, P.J., N.J. Engbrecht, J.R. Robb, V.C.K. Terrell, and M.J. Lannoo. 2013. Surveying a threatened amphibian species through a narrow detection window. Copeia 552–561.

1.7 TABLES

TABLE 1. Distances to noise sources and measured noise levels during equipment setup at each study site. Ambient noise is defined as the noise level without any identifiable or distinguishable sources. Peak noise levels were used for ambient and acute noises.

Site	Nearest road (m)	Nearest train tracks (m)	Ambient noise (dBA)	Acute noises (dBA)
A	110	3,885	50	cars (65); truck (70); motorcycle (75)
B	402	1,832	51	train (64); airplane (59)
C	180	119	50	cars (52); train (81); airplane (60)
D	176	174	46	cars (48); train (72); airplane (61)
E	383	37	44	cars (45); train (83); airplane (69)

TABLE 2. Recognizer information. Recognizers 1-4 are for Green Frog (*Lithobates clamitans*) mating calls. Recognizers 5-8 are frog American Bullfrog (*Lithobates catesbeianus*) mating calls.

Recognizer Name	Target Species	Type	Frequency range (Hz)
1-Full	Green Frog	Conventional	187.5 to 3875
2-High	Green Frog	Narrow	2000 to 3875
3-Middle	Green Frog	Narrow	812.5 to 1562.5
4-Low	Green Frog	Narrow	125 to 875
5-Full	bullfrog	Conventional	187.5 to 5250
6-High	bullfrog	Narrow	2062.5 to 3000
7-Middle	bullfrog	Narrow	562.5 to 1625
8-Low	bullfrog	Narrow	187.5 to 500

TABLE 3. Classification indices by site for Green Frog (*Lithobates clamitans*) recognizers. The indices are as follows: true positive rate (TPR); precision (PPV); false negative rate (FNR); and false discovery rate (FDR). The best-rated index value in each category is bolded. The most optimal recognizer per site is bolded and starred.

Site	Recognizer	TPR	PPV	FNR	FDR
A	1-Full	0.09	0.15	0.83	0.85
	2-High	0.03	0.13	0.92	0.87
	3-Middle	0.29	0.21	0.71	0.79
	4-Low	0.33	0.09	0.67	0.91
B	1-Full	0.12	0.01	0.88	0.99
	2-High	0.07	0.04	0.93	0.96
	3-Middle	0.45	0.01	0.55	0.99
	4-Low	0.25	0.00	0.75	1.00
C	1-Full	0.15	0.23	0.84	0.77
	2-High	0.26	0.27	0.70	0.73
	3-Middle	0.27	0.24	0.68	0.76
	4-Low	0.15	0.04	0.80	0.96
D	1-Full	0.18	0.28	0.82	0.72
	2-High	0.33	0.38	0.67	0.62
	3-Middle	0.26	0.23	0.74	0.77
	4-Low	0.22	0.07	0.78	0.93
E	1-Full	0.19	0.78	0.81	0.22
	2-High	0.26	0.70	0.74	0.30
	3-Middle	0.35	0.65	0.65	0.35
	4-Low	0.19	0.30	0.81	0.70

TABLE 4. Classification indices by site for American Bullfrog (*Lithobates catesbeianus*) recognizers. The indices are as follows: true positive rate (TPR); precision (PPV); false negative rate (FNR); and false discovery rate (FDR). The best-rated index value in each category is bolded. The most optimal recognizer per site is bolded and starred.

Site	Recognizer	TPR	PPV	FNR	FDR
	5-Full	0.27	0.19	0.73	0.81
	6-High	0.26	0.07	0.74	0.93
A	*7-Middle*	0.62	0.26	0.38	0.74
	8-Low	0.51	0.35	0.49	0.65
	5-Full	0.03	0.08	0.97	0.92
	6-High	0.11	0.06	0.89	0.94
B	7-Middle	0.54	0.10	0.46	0.90
	8-Low	0.76	0.16	0.24	0.84
	5-Full	0.50	0.57	0.50	0.43
	6-High	0.47	0.44	0.53	0.56
D	7-Middle	0.46	0.59	0.54	0.41
	8-Low	0.41	0.46	0.59	0.54
	5-Full	0.32	0.22	0.68	0.78
	6-High	0.24	0.11	0.76	0.89
E	7-Middle	0.31	0.11	0.69	0.89
	8-Low	0.35	0.16	0.65	0.84
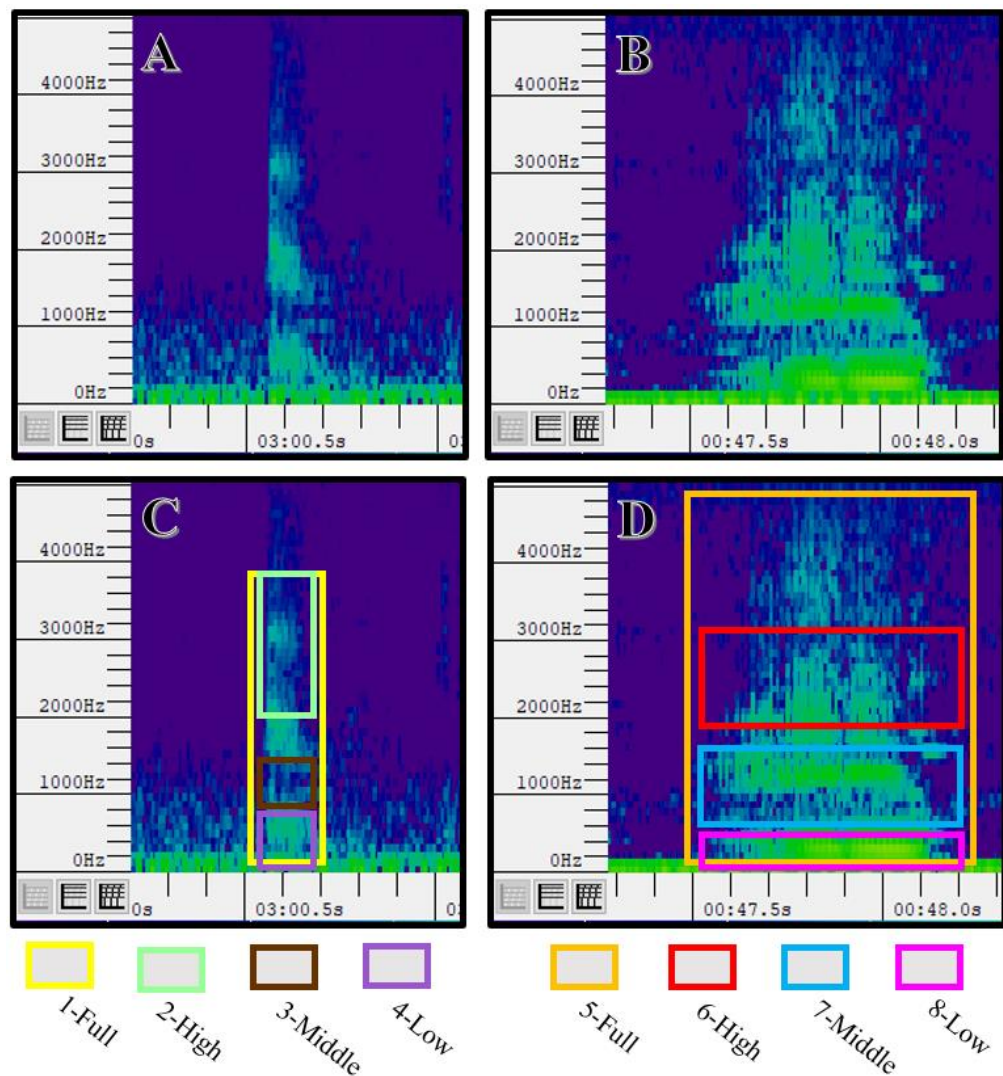
FIGURE 1. Spectrograms of mating calls for our target species with minimal background noise interference. (A) The Green Frog (*Lithobates clamitans*) has a quick "banjo-strumming" call with several simultaneous frequency harmonics. (B) The American Bullfrog (*Lithobates catesbeianus*) has a longer, sometimes vibrating call, with several simultaneous frequency harmonics. (C) Green Frog recognizers and (D) bullfrog recognizers are visually depicted over their respective calls to show the varying frequency ranges of each recognizer.

FIGURE 2. The proportion of correctly identified calls over manual counts for the Green Frog (*Lithobates clamitans*) per recognizer and site (A-E). Recognizer 1-Full (yellow) is the conventional recognizer, and all other recognizer are narrow-banded, meaning that only part of the full frequency range of typical Green Frog calls were scanned in the spectrograms.
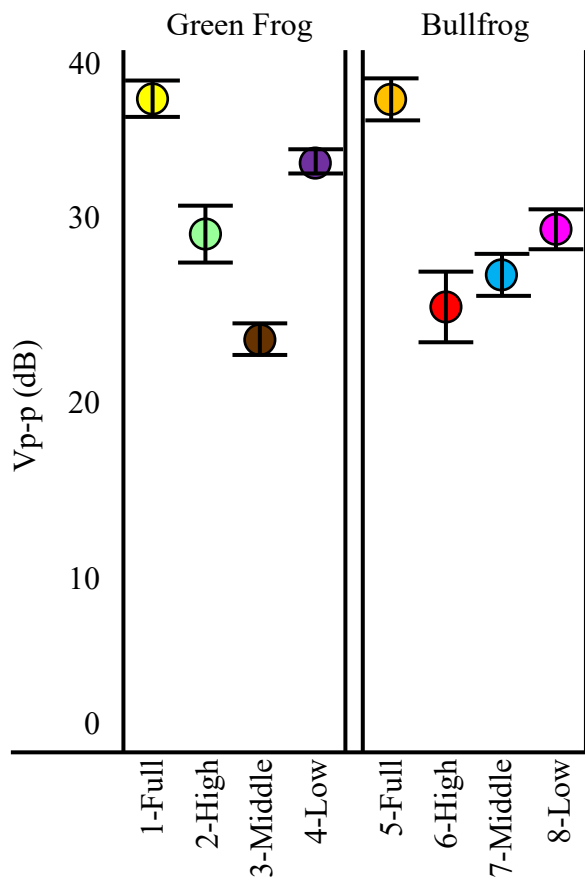
FIGURE 3. Average relative volume in peak to peak voltage (Vp-p) of background noise samples from our recognizers (solid vertical bars). Vertical lines represent 95% confidence intervals. All narrow-banded recognizers were significantly lower than the conventional recognizers (1-Full and 5-Full).
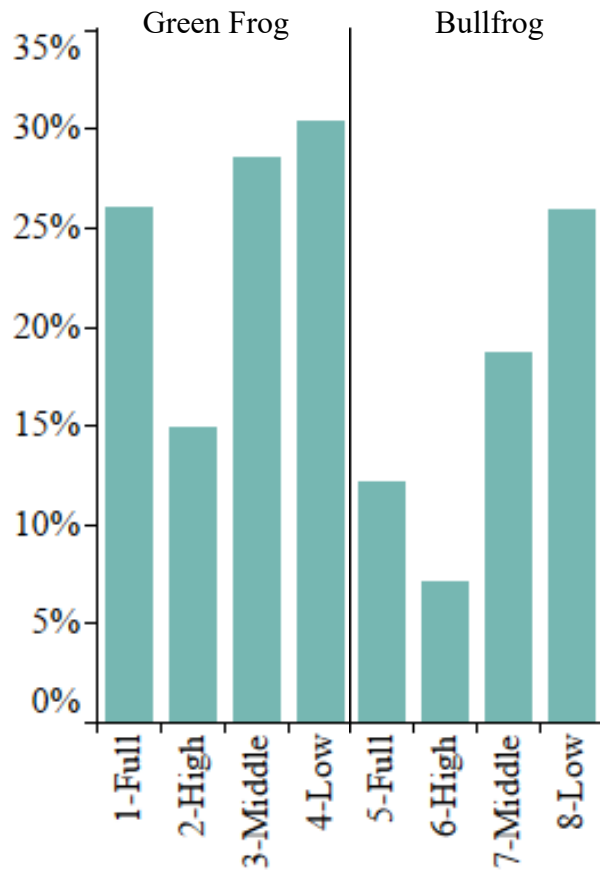
FIGURE 4. Percentage of errors (FPs + FNs) from our recognizers classified as anthropogenic noise for both Green Frogs (*Lithobates clamitans*) and American Bullfrogs (*Lithobates catesbeianus*).
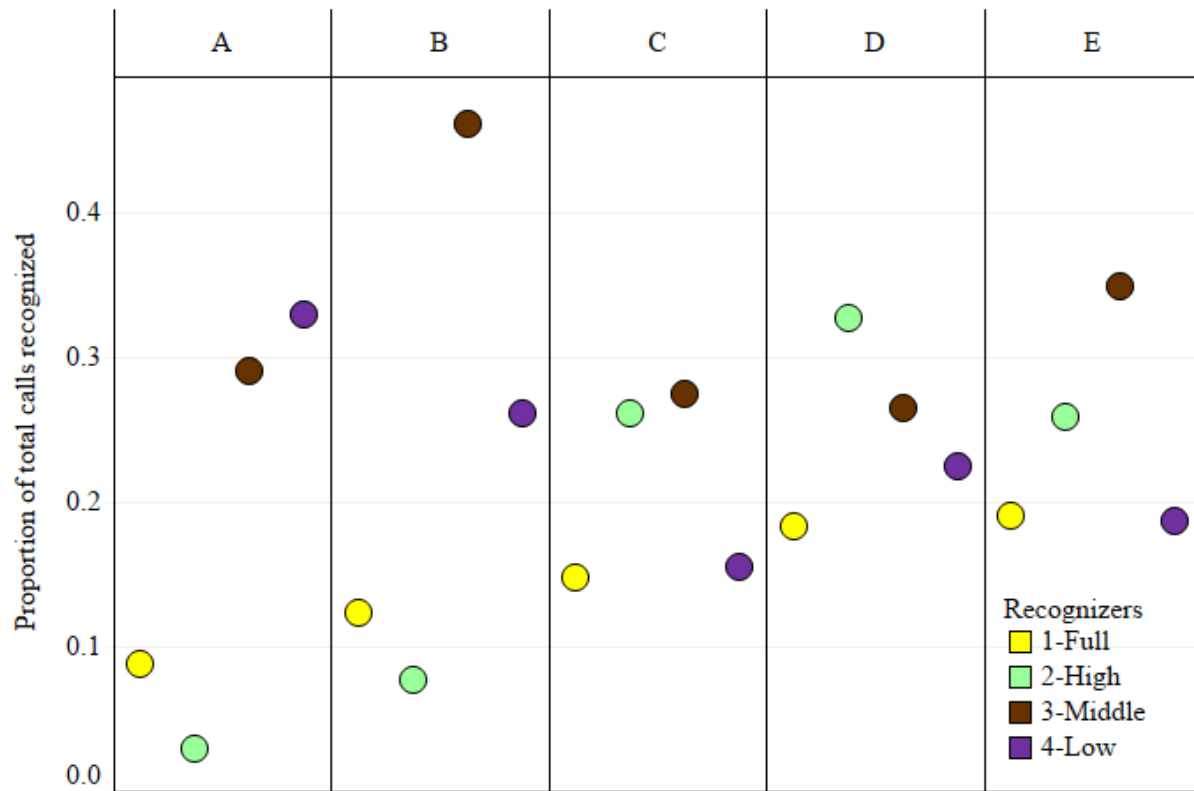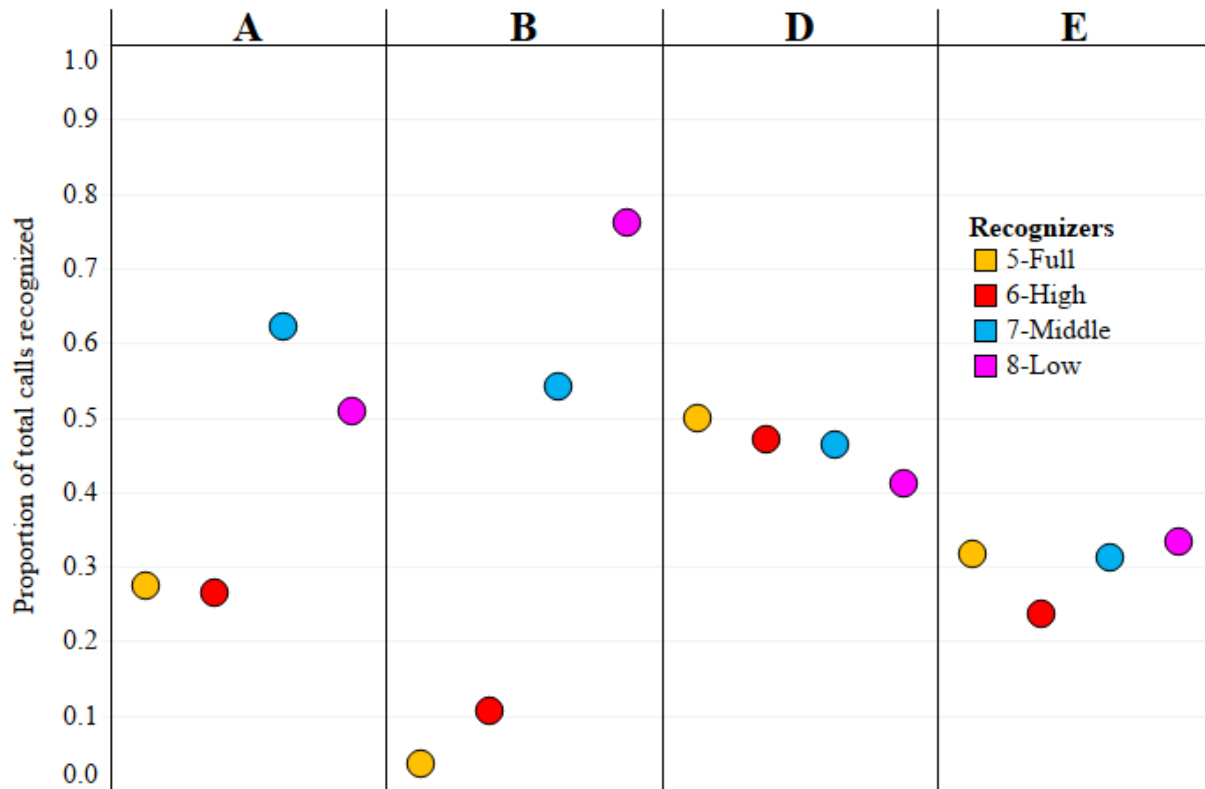
FIGURE 5. The proportion of correctly identified calls over manual counts for the American Bullfrog (*Lithobates catesbeianus*) per recognizer and site (A, B, D, and E). Recognizer 5-Full (orange) is the conventional recognizer, and all other recognizer are narrow-banded, meaning that only part of the full frequency range of typical bullfrog calls were scanned in the spectrograms.

APPENDIX A. The full set of parameters used to create our recognizers in Song Scope. Recognizers 1-4 are for Green Frog (*Lithobates clamitans*) mating calls. Recognizers 5-8 are for American Bullfrog (*Lithobates catesbeianus*) mating calls. For information on the parameters see the Song Scope manual (Wildlife Acoustics. 2011. *op. cit.*).

| Recognizer Name	1-Full	2-High	3-Middle	4-Low	5-Full	6-High	7-Middle	8-Low
Species	Green Frog	Green Frog	Green Frog	Green Frog	Bullfrog	Bullfrog	Bullfrog	Bullfrog
Type	Conventional	Narrow	Narrow	Narrow	Conventional	Narrow	Narrow	Narrow
Max. Complexity	20	20	20	20	25	25	25	25
Max. Resolution	10	10	10	10	12	12	12	12
Freq. min (Hz)	187.5	2000	812.5	125	187.5	2,062.50	562.5	187.5
Freq. max (Hz)	3,875	3875	1562.5	875	5,250	3,000	1,625	500
Max. syllable (ms)	1,128	1,128	1,128	1,128	504	472	504	496
Max. syl. Gap (ms)	0	0	0	0	232	216	232	256
Max. Song	1,400	1,400	1,400	1,400	2,576	2,064	2,456	2,808

Fast Fourier Transformation Size	512	512	512	512	512	512	512	512
Fast Fourier Transformation Overlap	2-Jan	2-Jan	2-Jan	2-Jan	2-Jan	2-Jan	2-Jan	2-Jan
Sample Rate (Hz)	4,800	4,800	4,800	4,800	4,800	4,800	4,800	4,800
Background filter	1s	1s	1s	1s	1s	1s	1s	1s
Algorithm	2	2	2	2	2	2	2	2
Quality	40	20	30	30	20	25	20	25
Score	40	65	40	60	50	50	50	65

(Appendix A continued)
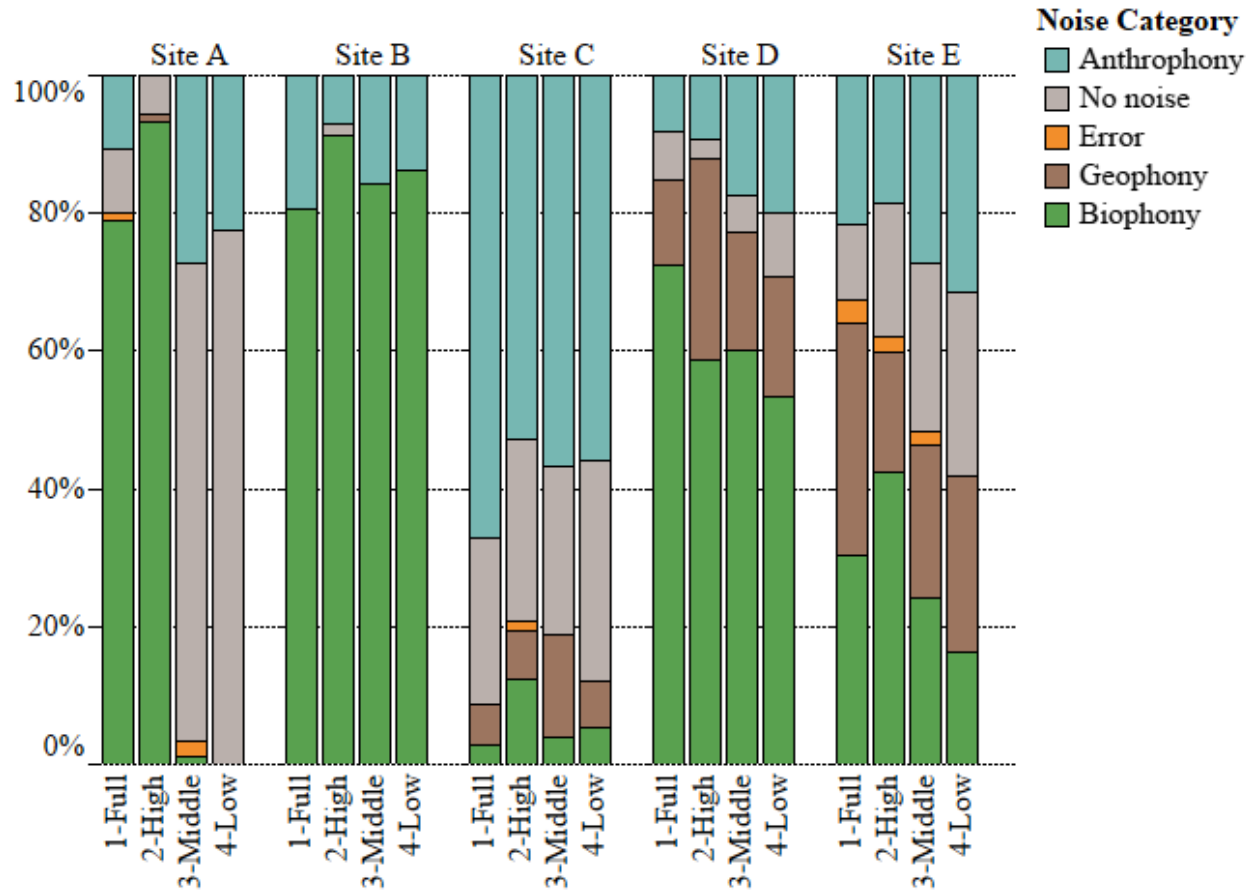
APPENDIX B. Noise category percentages from Green Frog (*Lithobates clamitans*) recognizer samples by site. The difference in the noise profiles between sites stresses the importance of considering the noise profile of a research location *a priori* when using automated detection.
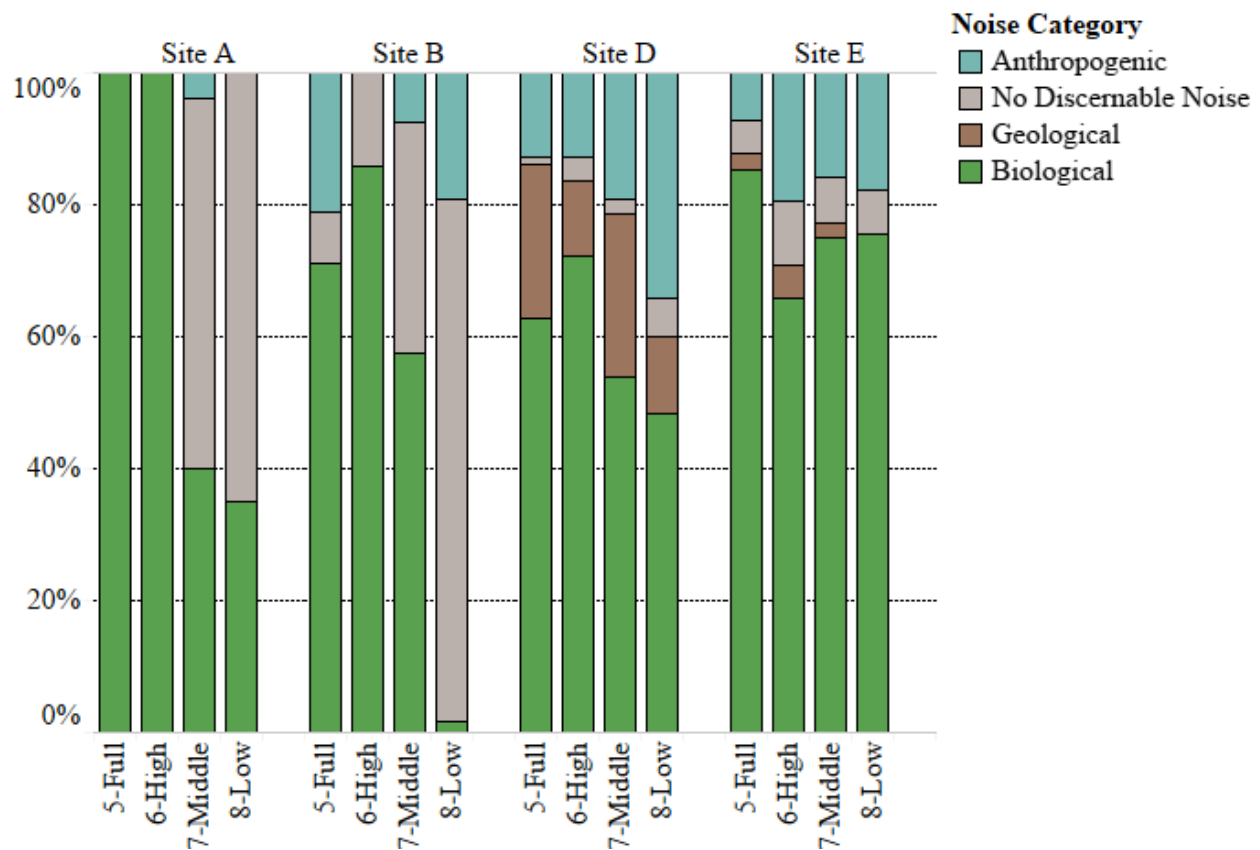
APPENDIX C. Noise category percentages from American Bullfrog (*Lithobates catesbeianus*) recognizer samples by site. The differences in the noise profiles highlights the value in our approach to independently analyze the performance of recognizers at each site and stresses the importance of considering the noise profile of a research location *a priori* when using automated detection. Site C was removed as there were no calls detected.
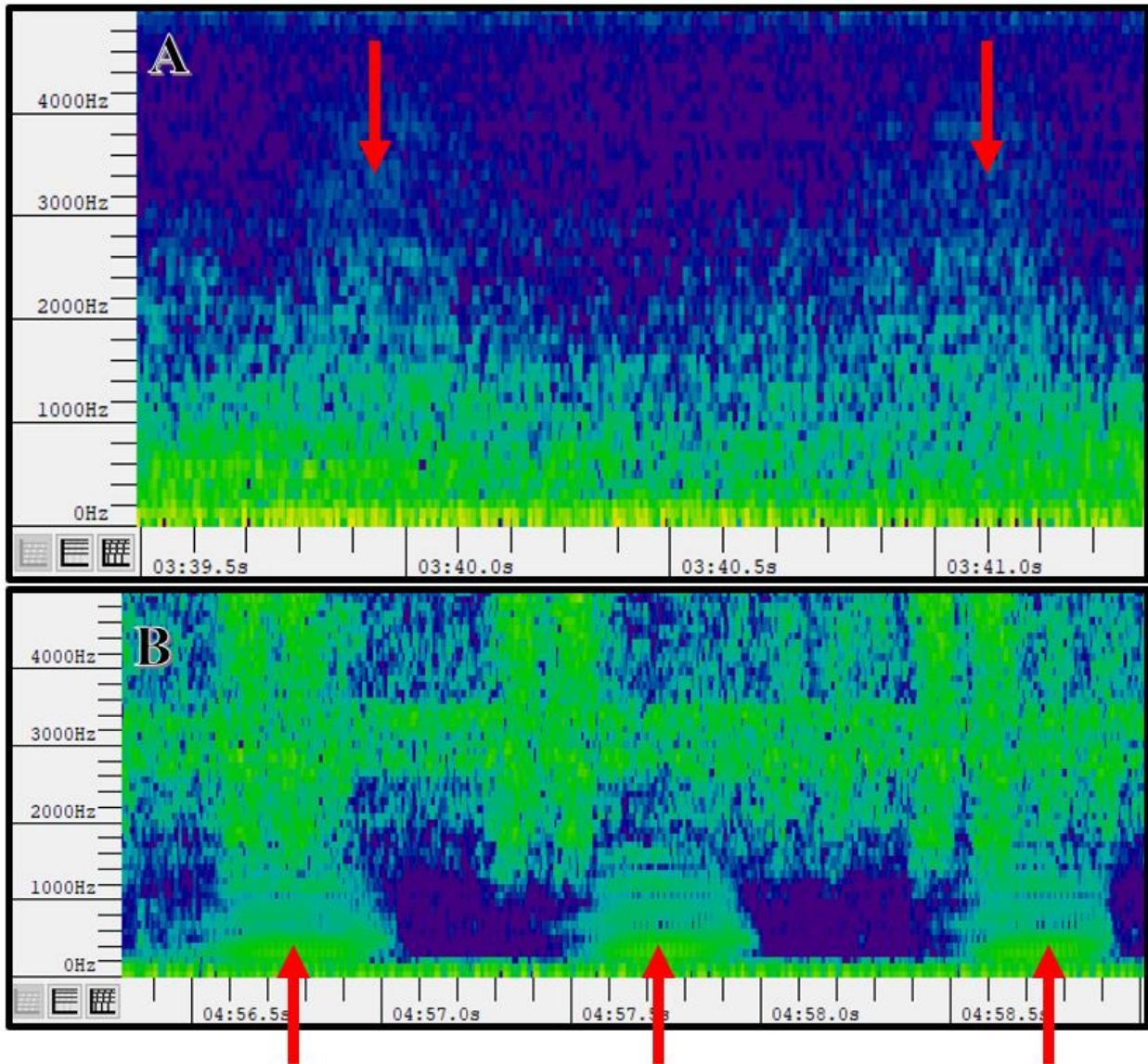
APPENDIX D. Spectrograms of American Bullfrog (*Lithobates catesbeianus*) calls in noisy environments. Red arrows indicate locations of un-interfered call energy. The benefit of using narrow-banded recognizers is clear in this figure, as only portions of calls are available to be identified. (A) Two calls almost entirely overlapped by airplane noise at site D. A full-frequency recognizer would experience a lot of noise interference, whereas a high-frequency narrow-banded recognizer may be able to identify remnants of the signal. (B) Three calls with higher-frequency overlap from intense insect chorusing at site A. Here, a lower or full-frequency recognizer would be more effective than a high-frequency recognizer to recognize calls.
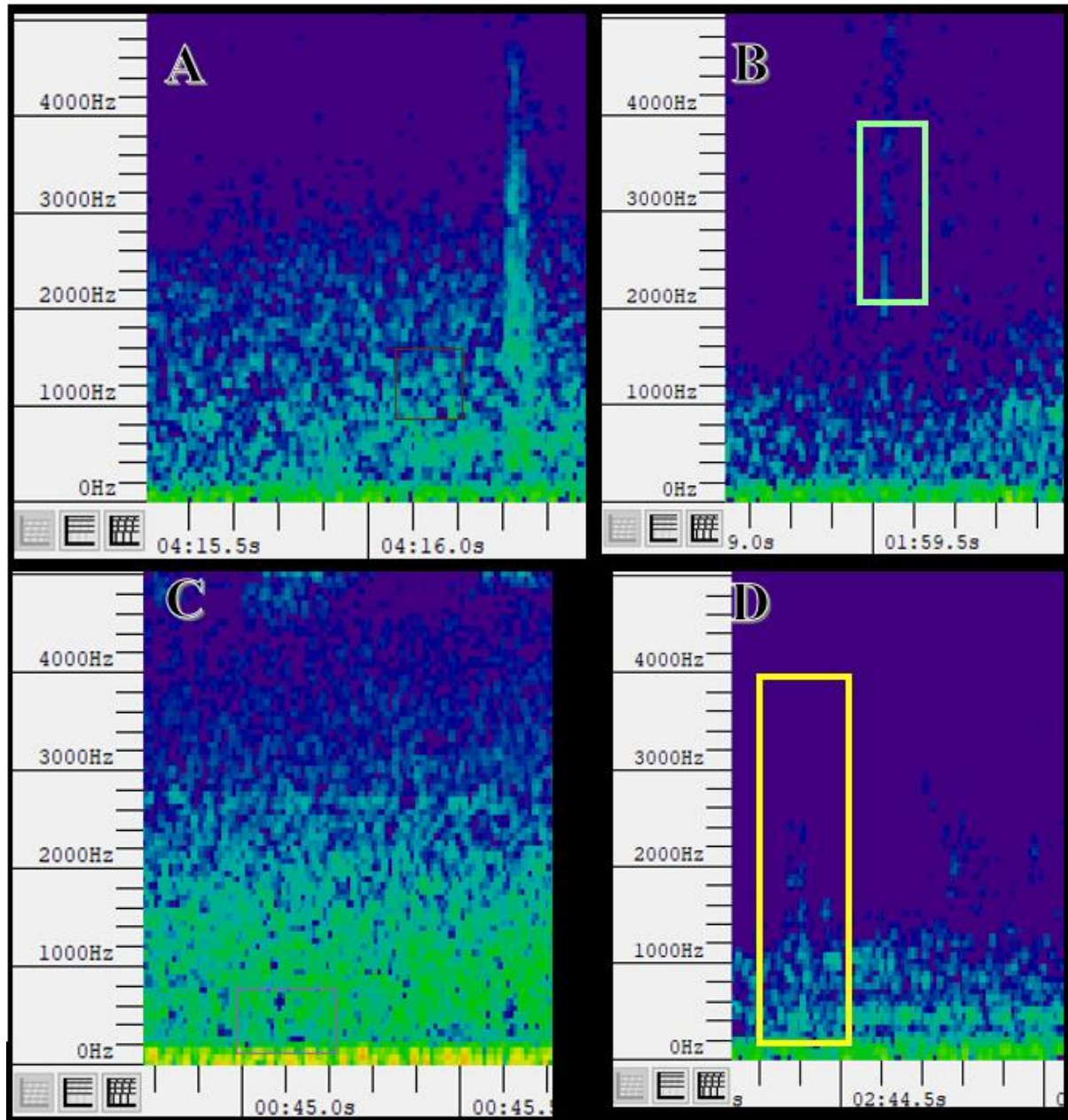
APPENDIX E. Spectrograms of false positives from Green Frog (*Lithobates clamitans*) recognizers. Colored boxes represent the state-space identified incorrectly as frog calls. (A) A false positive from recognizer 3-Middle at site C caused by car traffic. (B) A false positive from recognizer 2-High at site C caused by a snapping branch. (C) A false positive from recognizer 4-Low at site C caused by train traffic. (D) A false positive from recognizer 1-Full at site C caused by a truck.

APPENDIX F. Information about study sites included in this publication.

Site ID	Preserve Name	Site Name	Ownership	Coordinate N	Coordinate W
A	Cherry Hill Woods	Main Pond	Forest Preserve District of Cook County	41.6737	-87.873
B	Wolf Road Woods	Tomahawk Slough	Forest Preserve District of Cook County	41.7048	-87.9
C	Palos Hills	Back Pond	Private	41.6595	-87.828
D	Palos Hills	Front Pond	Private	41.6585	-87.828
E	Van Patten Woods	East Pond	Lake County Forest Preserves District	42.4748	-87.93

CH 2: ANTHROPOGENIC NOISE PLAYBACK EXPERIMENTS ON URBAN POPULATIONS OF NINE FROG

SPECIES

2.1 ABSTRACT

Given the acoustic nature of frog breeding, anthropogenic noise is an edge effect associated with urbanization that diminishes anuran habitat quality. Previous studies that use noise playback experiments to examine the effects on calling behavior yield no overarching patterns, but few studies investigated patterns across entire communities. We conducted a total of 155 playback experiments on nine anuran species across 32 breeding sites around the greater Chicago region. We investigated whether patterns in call attributes (call rate, dominant frequency, call length, and pulse rate) follow taxonomic or natural history characteristics. We also tested whether the anthropogenic noise source (cars, airplanes, or trains), or the magnitude of noise level (peaks of 70, 80, or 90 dB) affected behavioral responses. We detected multivariate clustering by genus and species but not by natural history characteristics. We found notable consistency in the changes to call rate, as eight species reduced their call rate. The only species not to reduce call rate was also the only one to increase dominant frequency, suggesting that these are two mutually exclusive strategies for dealing with anthropogenic noise. We did not detect major patterns in call length or pulse rate changes. We also did not see consistent changes between the different noise sources, indicating that our populations were unable to perceive differences in anthropogenic noise sources and act differently as a response. For noise magnitude, most behavioral changes occurred in the 90 and 80 dB treatments, indicating that noise may only be affecting calling behavior in the loudest habitats, such as roadside ditches and retention ponds. Considering that artificial ponds contribute considerably to the biodiversity and

dispersal ability of frogs in urban landscapes, noise pollution should be a concern for conservationists.

      Key Words: acoustic species; anthropogenic noise; Anura; bioacoustics; call attributes; community; playback experiments; soundscape, urban ecology

2.2 INTRODUCTION

      Amphibians are the focus of many ecological surveys, restoration efforts, and manipulative experimentation due to their worldwide decline (Phillips 1990; Stuart *et al.* 2004; Hamer & McDonnell 2008), as amphibian species are dying off at a rate at least 200 times faster than their background extinction rate (McCallum 2007; Roelants *et al.* 2007). Most documented cases of amphibian decline have been noted in anuran populations (Barinaga, 1990; Gibbs *et al.*, 2005; Steelman & Dorcas 2010). Causes of anuran decline typically emanate from habitat destruction and fragmentation (Knutson *et al.* 1999; Gibbs *et al.*, 2005; Cushman 2006). Frogs are particularly vulnerable to these causes due to their relative inability to traverse developed landscapes and their combined aquatic and terrestrial habitat requirements, in which an absence of one or the other habitat type generally indicates an uninhabitable landscape (Cushman 2006; Hamer & McDonnell 2008). As cities around the world continue to grow, animal populations living at the junction of developed and natural landscapes experience the brunt of urban expansion. This inevitably entails habitat fragmentation by roads, which can spread edge effects even further into natural areas. Frogs experience a multitude of these edge effects in urbanized landscapes which threaten their regional and global biodiversity, including altered hydrology (Calhoun *et al* 2005; Hamer & McDonnell 2008), pollution (Phillips 1990; Hamer & McDonnell

2008), and invasive species (Phillips 1990; Hamer & McDonnell 2008; Mazerolle *et al.* 2014; Sacerdote & King 2014).

Given the acoustic nature of frog breeding, anthropogenic noise is another problem associated with urbanization that weakens anuran habitat quality. Frogs rely on vocalizations for attracting and rejecting potential mates, defending territories, signaling danger, and drawing in conspecifics and heterospecifics to new breeding ponds (Benedix & Narins 1999; Bee *et al* 2000; Bee 2007; Lengagne 2008; Forti *et al.* 2017). Anthropogenic noise can obstruct these processes by interfering with frequencies and reducing the fidelity of signals, which lowers the utility of acoustic communication for both senders and receivers.

In this study, we investigated how anthropogenic noise affects changes in mating calls, which are advertisements made by males to attract females and are the most common signals that frogs use. Observable attributes of male advertisement calls can be classified as either static or dynamic properties (Gerhardt 1991). Static properties have limited plasticity and are stereotyped indications of species identity. These generally include dominant frequency and pulse rate. Dynamic properties have greater plasticity and generally undergo directional selection from females. These typically include call length and call rate. Since females directionally select for higher calling effort (defined as longer and more frequent calls), males spend extremely high amounts of energy to outperform competitors (Gerhardt 1991). Rates of energy used while males are chorusing are 6 to 21 times their resting rate (Prestwich 1994). Although females make large energy investments in producing eggs, male lipid reserves decrease even more than females over the course of a breeding season (Grafe *et al.* 1992). This loss in body condition is especially true in species with prolonged breeding seasons (Gerhardt 1994). Thus, alterations in calling behaviors as a response to increased noise levels may have far-reaching implications on

40

male competition, reproductive success, energy expenditure, speciation, and the suitability of a habitat.

2.2.1 Previous lab research

Lab research on frog communication has mainly focused on phonotaxic (based on orientation or movement to a sound source) choices or acoustic reactions of individuals to multiple auditory signals like noise or conspecific calling (Gerhardt *et al.* 2000; Schwartz 2001; Luddecke 2002; Bee 2007; Kuczynski *et al.* 2017). The strength of lab research resides in the ability to examine specific mechanisms behind precise reactions to sounds that may otherwise be hard to detect without controlling for the numerous confounding factors experienced in the field. Many lab studies collect data on individuals placed in an arena with differential playback at opposite ends while researchers monitor the positioning, movement, and calling of the target individual (Gerhardt *et al.* 2000; Shen & Xu 2016). A review of the literature on lab experiments reveals varying results. For example, Halfwerk *et al.* (2016a) found that individual male Tungara frogs (*Physalaemus pustulosus*) placed in experimental pools increased their call effort—along with amplitude and complexity (defined as number of "chucks"/number of "whines")—in in response to white noise. On the other hand, Lengagne (2008) found that the European tree frog (*Hyla arborea*) reduced its call effort in response to interference from traffic noise. Additional strategies include altering call frequency in order to avoid noise overlap and or raising call amplitude to eclipse background noise levels (Shen & Xu 2016).

Although lab techniques have specific advantages over field research, lab scenarios may entail a significant abstraction from reality, as males and females in nature operate in a very competitive environment that is raucous and more chaotic. Lab experiments on solitary individuals (Gerhardt *et al.* 2000; Halfwerk *et al.* 2016a; Shen & Xu 2016) do not appropriately

simulate the natural breeding environment. It is important to consider group-level responses, because results can vary depending on the chorus size. For example, Lengagne (2008) found diminished effects of traffic noise in group choruses compared to individual callers. Thus, the ability of individual agents to make optimal or predictable decisions may be exaggerated or skewed in controlled settings. Stratman & Hobel (2019) provides insight into this phenomenon. They found that the trait for which females show the strongest preference in arena playback experiments (call length) does not match the most detectable trait for females in the wild (call rate). Moreover, they estimate that half of mating encounters in the wild do not provide the trait variation necessary for females to reliably express a preference. Consequently, much more mating in the wild is arbitrary or non-optimal compared to what might be predicted by lab studies. It is important to note that the Stratman & Hobel (2019) study focuses on the female perspective, but a similar consideration for this phenomenon should be given for male behaviors: males may not adjust their calls optimally or as consistently in natural settings, even if lab experiments suggest that they can. This underscores the value of field-based studies.

2.2.2 Previous observational field research

Observational studies have demonstrated that in the presence of car traffic noise, certain species may increase call frequency (Parris *et al.* 2009; Cunnington & Fahrig 2010; Hoskin & Goosem 2010; Grenat *et al.* 2019), lower their call rate (Vargas-Salinas & Amezquita 2013; Vargas-Salinas *et al.* 2014), increase their call rate (Cunnington & Fahrig 2010; Hoskin & Goosem 2010), increase their pulse rate (Grenat *et al.* 2019) or produce no behavioral changes (Cunnington & Fahrig 2010). It is challenging to infer an overarching pattern across species. Moreover, it is difficult to isolate the effect of anthropogenic noise as it is inexorably linked to habitat fragmentation. Therefore, changes in behavior or population structure detected in

observational studies may not be a response to noise *per se* and may not be reliable enough to extrapolate to other species or scenarios.

Observed patterns may be affected by road mortality, especially if roads are located along dispersal or migration paths. Hels & Buchwald (2001) found that up to 10% of their study populations from nearby ponds are killed on roads annually. These effects may be greatest for highly vagile frogs (Carr & Fahrig 2001). Road salts are another potential confounding factor, as it is shown to increase juvenile mortality (Dananay *et al.* 2015). If certain subsets of frog populations are more vulnerable to these confounding factors, then changes in chorusing characteristics may manifest. For example, if size influences dispersal distance or protection from osmolality changes, then a change in the average size of a population should also lead to a change in call frequency, as body size and call frequency are correlated (Bee and Gerhardt 2001).

2.2.3 Previous field playback experiments

Field playback experimentation is a more effective approach to study the effects of noise on anuran communication, but a thorough review of past field experiments yields no overarching patterns or trajectories (Table 1). This parallels the discordant results from observational studies and lab playback experiments. The data tell a bewildering tale: In response to anthropogenic noise males have been shown to increase call rate (Sun & Narins 2005; Kaiser & Hammers 2009; Bleach *et al.* 2015; Engbrecht *et al.* 2015, Kruger & Du Preez 2016), decrease call rate (Sun & Narins 2005; Caorsi 2017), reduce call length (Hanna *et al.* 2014), lower call frequency (Hanna *et al.* 2014; Caorsi 2017), raise call frequency (Kruger & Du Preez 2016), increase power (Yi & Sheridan 2019) or produce no behavioral changes (Bleach *et al.* 2015; Forti *et al.* 2017). These varying results may be a consequence of disjoint experimental approaches or a narrow

43

theoretical framework based on a single target species, but it may also be that responses are unique according to species and localized environmental conditions. This underscores the need for research that investigates patterns across entire communities, using the same methods for each species, which allows for researchers to test more a priori predictions on taxonomy and natural history patterns.

2.2.4 Our Study

We conducted a suite of anthropogenic noise playback experiments on the community of frogs living in the Chicago region and northwest Indiana. Our research is valuable and innovative in the following ways: (1) Research was conducted on nine species, which allows us to examine patterns across genera and specific natural history traits. (2) We used three different types of anthropogenic noise (car traffic, train traffic, and airplane flyovers), whereas most previous studies only included car traffic or white noise as a proxy for all anthropogenic noise (Table 1). (3) We varied the amplitude of anthropogenic noise in order investigate threshold limits. (4) Single treatments were much longer than most playback experiments, and we incorporated "washout periods" of no noise to reduce carryover effects. (5) Experiments were conducted at the population level, which resembles the acoustic environment in which females make mate choices.

2.2.5 Predictions

We used past playback experiments and hypotheses from previous soundscape research (Pijanowski *et al.* 2011) as a guide for our predictions. We amassed a large amount of playback experiments across many species in order to create an approach that incorporates natural history into behavioral predictions, with the intention of being useful for other species outside of our target community.

Call rate

Call rate is arguably the most important call attribute to track, because it is the most detectable trait for females in the wild, and thus female preference will be expressed through discriminating between call rates (Stratman & Hobel 2019). If males can distinguish anthropogenic noise from conspecific chorusing noise—which seems likely, considering the specificity with which frogs can discriminate between different calls (Bee 2010; Bee *et al.* 2012; Schrode *et al.* 2012)—males would have no need to respond to a biologically irrelevant signal by producing expensive calls. It would thus be best for individuals to reduce call rate, in order avoid wasting energy when calling in unfavorable conditions for females to assess male suitors. Due to the intermittent nature of anthropogenic noise, frogs should be willing to temporarily forgo calling in the present for a future time with more favorable soundscape conditions. Therefore, our prediction is that call rates will be reduced.

It is a difficult task to provide an informed prediction on call rate that perfectly corresponds to previous research, because studies have shown patterns in both directions. Thus, any prediction will have some previous evidence that supports and opposes it. The most obvious explanation for the opposing hypothesis (i.e. frogs will increase their call rate in response to anthropogenic noise) is that males would be unable to distinguish between anthropogenic noise and conspecific chorusing. The soundscape produced by playback would therefore mimic a competitive scenario in which more males are calling in the vicinity. Noise treatments should then elicit increased call rates compared to the control. However, this seems to be the less likely scenario due to the inherent differences in frequency, length, and pulses between anthropogenic noise and frog calls. These traits are highly stereotypical within species, to the point where females can distinguish between conspecific and heterospecific males. Males should only seek to

intensify their call rate when the level of conspecific signals is increased. Moreover, we are testing urban populations, which means that they may a greater ability to interpret non-biological noises compared to their counterparts in quieter regions.

Call frequency

Another method frogs can use to cope with noise is to shift their call frequency away from the noise frequency to avoid spectral interference. Documented frequency shifts in response to noise have been in both directions (up: Kruger and Du Preez 2016; Shen & Xu 2016; or down: Hanna *et al.* 2014; Caorsi 2017). This is a commonly detected response in noise studies, probably because frequency alteration is already in male frogs' behavioral repertoires to adjust to natural sounds (Wong *et al.* 2009; Shen & Xu 2016) or to exaggerate size in territorial disputes (Bee *et al* 2000). Because anthropogenic noise is characteristically low frequency, we expect frogs to shift their frequencies higher in order to avoid spectral interference. We should also see a clearer signal of this shift in the naturally low-frequency callers (such as the bullfrog, green frog, American Toad, and wood frog) because they will experience the most spectral overlap.

Call length

Our predicted changes in call length mirror that of call rate, as they are both dynamic properties that are directionally selected for by females but are limited by energy. Therefore, we expect call length to be reduced. Call-length reduction should be most prominent in the car traffic treatments, as the interval between cars driving by is much shorter than the interval between trains and airplanes. In the airplane and train noise treatments, there is a longer duration of time between peak power levels (the point at which the airplane or train is closest to the

recorder), which may allow certain species to make full-length calls during lulls in the treatments without significant temporal overlap.

Threshold effect

For species that do react to the anthropogenic noise treatments, changes in attributes should be more extreme in louder noise treatments. Traffic noise at 90 dB should elicit greater call changes than 80 dB or 70 dB. It will be more difficult for females to accurately interpret mating calls during the noisier treatments, so it would benefit the males to stand out more by compensating through whatever mechanism they implement, whether it be through call rate, frequency, or length. If there are any changes, the magnitude of change will be greatest in the loudest treatment. We do not expect to see any changes in behavior during the 70 dB treatments because this is close enough to normal noise levels of a chorus (Bielinski, pers. obs.), and should not be noticeable to calling males.

Habitat preference patterns

Many natural history characteristics are dictated by the hydrologic requirements of that species. Hydrology influences calling phenology, overwintering strategies, antipredator behavior, and much more (Phillips *et al.* 1999). This may also be true of their response to anthropogenic noise. Species living in permanent ponds have longer breeding periods and therefore have less pressure in the moment to breed. During unfavorable conditions, they can switch from breeding to hunting, in order to acquire resources to continue calling later in the day or season. However, for explosive breeders in ephemeral ponds, males only have a short amount of time to breed and should be less wavering when it comes to environmental factors affecting their breeding behavior (Bevier 1997; McCauley *et al.* 2000). Explosive breeders should already be close to their maximum and most attractive calling output, and they should prioritize sticking

with this strategy despite changing acoustic conditions (Wells & Bevier 1997). Therefore, we predict that species living in permanent ponds (bullfrog, green frog, cricket frog) will have greater changes in dynamic properties compared to species living in ephemeral ponds (wood frog, spring peeper, western chorus frog).

Taxonomic patterns

Species will have different innate responses to noise interference, but the strategy with which they respond should be genetically influenced, and thus congeneric species should have more similar responses compared to heterogeneric species. We predict that some call attributes will be conserved across treatments if they are essential to distinguish between species. If congeneric species have similar mechanisms to identify species, then call recorded call attributes will cluster by genera.

2.3 MATERIALS AND METHODS

2.3.1 Study sites

This study was conducted at 21 frog breeding sites around the Chicago region and northwest Indiana (Figure 1). Most locations were in the south suburbs of Chicago, and were a mixture of city parks, county forest preserves, county-owned land, Shirley Heinze Land Trust sites, and the Indiana Dunes National Park. Sites varied greatly in vegetation and hydrology, as would be expected considering the diverse ecological requirements of our nine study species. Before experiments were conducted, we initially surveyed potential sites to look for the presence of target species, and to keep track of ephemeral pond locations, which at times can be unpredictable. Due to varying levels of regional abundance and the travel limitations of a single

48

field researcher, there were not equal numbers of sites or experiments across species. The number of playback experiments conducted per species is found in Figure 2.

2.3.2 Experimental design

We used a "crossover" design for our playback experiments, meaning that "test subjects" (ponds) received a randomized sequence of different treatments, one of which being a control (Figure 3). We incorporated four 15-min treatments: three broadcasted anthropogenic noises (car, airplane, and train noise), and one control treatment of no noise. We added 10-min periods of silence (known as "washout periods") in between the four treatments to reduce any carryover effects. In order to collect initial and final conditions, we added 15-minute periods of data collection before and after the sequence of treatments. A silent 10-min "burn-in period" preceded all data collection to make sure that the calling behavior was not affected by the disturbance called when setting up the equipment. The temporal structure of the playback experiments was as follows:

Period 1:	10 min burn-in
Period 2:	15 min initial conditions
Period 3:	15 min noise treatment
Period 4:	10 min washout
Period 5:	15 min noise treatment
Period 6:	10 min washout
Period 7:	15 min noise treatment
Period 8:	10 min washout
Period 9:	15 min noise treatment
Period 10:	10 min washout

49

Period 11:   15 min final conditions

The order of treatments was randomized for each experiment. The period × treatment combinations were not balanced across species because there were many instances when we conducted experiments on multiple species concurrently. Playback experiments did not begin until the target population (the highest priority species present at the pond) reached a call rate of 5 calls/min, which was typically an easy level to reach and was usually indicative of multiple calling males.

Noise source experiment

Throughout this study, we conducted two types of playback experiments. The first type, the "Noise Source Experiment," used three different anthropogenic noise treatments: car, airplane, and train. Each treatment was 15 minutes long with a peak of 90 dB at the loudest moments and a nadir of 60 dB during the quietest moments. Car noise was recorded from the side of LaGrange Road, a busy four-lane road that transects several forest preserves. The recorded car traffic created a random pattern of amplitude and frequency spikes and had an average SPL of 80 dB. Drops in noise level were usually very quick—less than a second—but the longest lull in noise lasted 11 seconds (Figure 4). The airplane flight path noise originated from O'Hare International Airport. Specifically, this was a looped sequence of 6 plane flyovers recorded in Labagh Woods Forest Preserve in Chicago. The sequence produced a predictable pattern of peaks and valleys for amplitude and frequency and had an average SPL of 85 dB (Figure 5). The inter-airplane duration in the recording lasted between 55 sec to 1 min 31 sec. The train railway noise was a looped sequence of a train passing by Van Patten Woods forest preserve, recorded 10 m away from the train tracks. Peaks in amplitude and frequency were

random as the train is passed by, but in between loops there was an extended lull in the noise that lasts for 1 min 8 sec (Figure 6). The train recording had an average SPL of 83 dB.

Noise power experiment

The second playback experiment, the "Noise Power Experiment" was structured to examine the effects of varying noise intensities. All three treatments used the 15-min car noise recording. The peak amplitude of the noise was adjusted to three different levels: 90 dB, 80 dB, and 70 dB, with an average SPL across the full recording of approximately 10 dB less than the peak level. The rationale behind using these three treatment levels is as follows: (1) We wanted to maintain the three-noise treatment structure of the previous experiment. (2) Noise at 90 dB was the loudest achievable level with our equipment and is about the same power level that can be registered in wetland ditches on the side of train tracks or a busy road (Bielinski, pers. obs). (3) We considered 60 dB of noise insufficient, as many of the initial measurements of the natural noise level gathered before each experiment often reached 50 to 60 dB.

2.3.3 Playback experiments

From 17 April 2017 to 23 July 2018 we conducted anthropogenic noise playback experiments on chorusing males at breeding ponds. Sites were visited during peak calling hours of the current target species, as indicated by previous research (Bielinski Dissertation Chapter 1) and natural history information from field guides (Phillips *et al.* 1999). Sites were not visited if rain was forecasted, in order to maintain a consistent amount of background noise from natural sources. We did not control for natural chorusing conditions or population numbers. Consequently, on a given day the number of chorusing individuals varied, and the presence of multiple species was possible. One to three recorders (H1 Handy Recorder, Zoom North America, Hauppauge, New York, USA) were set up at the edges of ponds, either strapped to a

51

tree, on a tripod at the pond edge or a several meters into the pond. Recorders were placed in

proximity to groups of calling individuals and distributed to capture multiple species when

possible. We also placed Bluetooth speakers (Turtle Shell 3.0, Outdoor Tech, Laguna Hills,

California, USA) along the pond's edge 10 m away from the recorder(s).

There was a considerable amount of disturbance when setting up the equipment, which

immediately led to a noticeable reduction of calls for several minutes. However, it was observed

that most individuals of any of our target species maintained their current positioning and started

calling again. If certain males did hop away, many were seen returning to their original location

over the next few minutes. This overall stationary behavior is consistent with past experiments

and observations (Ryan 1980; Gerhardt 1991; Sun & Narins 2005, Kaiser & Hammers 2009).

We implemented a 10 min "burn-in" period to ensure that frogs returned to their positions and

natural calling behavior after the possibility of disturbance when walking up on the pond and

setting up equipment. Past research indicates that 10 minutes should be substantial (Crouch &

Paton (2002). This is a longer and therefore more conservative burn-in time than suggested by

the North American Amphibian Monitoring Program and the Calling Frog Survey (US

Geological Survey. 2019. Available from https://www.usgs.gov/centers/pwrc/science/north-

american-amphibian-monitoring-program?qt-science_center_objects=0#qt-

science_center_objects [Accessed 2 August 2019]; Chicago Academy of Sciences. 2020.

Available from https://frogsurvey.org/?page_id=17 [Accessed 3 March 2020]).

2.3.4 Study species

Over the course of two years of field work, we conducted experiments on nine of the 11

anuran species in the region. Figure 7 shows the phenology, estimated relative abundance, and

habitat requirements for the region's anuran community. The two species absent from this study

include the northern leopard frog (*Lithobates pipiens*) and the Fowler's toad (*Anaxyrus fowleri*). We did not include the northern leopard frog in the analysis because were only able to record a total of five calls during the study. We did not find any breeding ponds for the Fowler's toad, which prefers sandy beaches or dune and swale habitat (Phillips *et al.* 1999) and is most likely to be found near Indiana's beaches. For each the nine species included in this study, below we summarize their natural history characteristics such as size, phenology, habitat requirements, mating behavior and call structure. Each summary is finished with a description of previous research related to playback experiments and anthropogenic noise.

Lithobates species

The bullfrog (*Lithobates catesbeianus*) is a large (9-15 cm) frog that breeds for two months, usually mid-April to mid-June. It is mostly associated with permanent ponds and is prevalent in urban areas (Phillips *et al.* 1999). Males can breed with a maximum of 3 females per year, and females usually breed with one or two males (Howard 1979). The advertisement call of the bullfrog is a single deep bass "vroom" often repeated in sets (Figure 8). Previous research has shown that bullfrogs call more often at a single site when traffic intensity is low (Vargas-Salinas 2014).

The green frog (*Lithobates clamitans*) is a large (5.7-9.5 cm) frog that has a long breeding season from early May to early September. It is mostly associated with permanent ponds, where it often is found around shallow edges and rocky outcrops (Phillips *et al.* 1999). It is prevalent in urban areas. Females usually breed with one or two males (Wells 1976). The advertisement call of the green frog is a short (about 160 msec) banjo-like twang that has a low dominant frequency (about 400 Hz) accompanied by higher frequency components (Figure 8; Bielinski Dissertation Chapter 1). Males can decrease their call frequency in competitive

settings in order to exaggerate size to other males, in an effort to protect their territory (Bee *et al.* 2000). Previous research has shown that green frogs call more often within a single site when traffic intensity is low (Vargas-Salinas *et al.* 2014). Across multiple sites, this species lowers its call rate, while increasing the amplitude and frequency of calls at sites that have higher anthropogenic noise levels. The same set of responses can be seen when using noise playback to raise the noise level of naturally low-noise sites (Cunnington & Fahrig 2010).

The wood frog (*Lithobates sylvaticus*) is a medium-sized (3.5-6 cm) frog that breeds right when the snow melts in a short two-week window somewhere between February and April. It breeds in vernal, forested pools. It is a rare species in the Chicago region. The advertisement call of the wood frog series of 5-6 cackling and clucking sounds, often resembling ducks (Figure 8; Phillips *et al.* 1999).

Pseudacris species

The western chorus frog (*Pseudacris triseriata*) is a medium-sized (up to 4 cm) frog that breeds from late February through late May. It is found in many types of habitats, including permanent ponds, wet prairies, and ephemeral forested pools (Phillips *et al.* 1999). It is prevalent in urban areas (Bielinski, pers. obs.). The advertisement call of the western chorus frog is a short (0.5-1.0 seconds) high pitched (3-3.5 kHz) trill resembling the sound of a finger running across the teeth of a comb (Figure 8).

The Spring Peeper (*Pseudacris crucifer*) is a small (up to 3.5 cm) frog that is one of the earliest to mate, from late February to mid-May. During the breeding season, it is mostly associated with ephemeral woodland pools (Phillips *et al.* 1999). The advertisement call of the spring peeper is a high pitched (~ 3 kHz) "peep" repeated several times (Figure 8). A previous

study has shown that in response to high-pitched white noise, males lowered their call frequency and made shorter calls (Hanna 2014).

Hyla species

The eastern gray tree frog (*Hyla versicolor*) is a medium-sized (up to 6 cm) frog that has a breeding season from late April to early August. It is mostly associated with shallow ponds or temporary woodland pools (Phillips *et al.* 1999). The eastern gray tree frog is morphologically identical to the Cope's gray tree frog, but the first is tetraploid and the latter is diploid. Males and females breed up to three times per year (Godwin & Roble 1983). The advertisement call of the eastern gray tree frog harsh trill of about 0.5 seconds long with a dominant frequency of 2.2 kHz (Figure 8). The call is very similar to the Cope's gray tree frog, but the eastern gray tree frog lower pitched with a slower trill rate (Gerhardt *et al.* 2007). Previous research on this species have shown contradictory results. One study found that that males do not adjust calling effort in response to increased traffic noise within a single site (Vargas-Salinas 2014), but another study found that they reduce call rate in response to car noise playback (Cunnington & Fahrig 2010).

The Cope's gray tree frog (*Hyla chrysoscelis*) has the same morphology, habitat, and breeding season as the eastern gray tree frog. It is the diploid species of the complex. Males and females breed up to three times per year (Godwin & Roble 1983). The advertisement call of the Copes gray tree frog harsh trill of about 0.5 seconds long with a dominant frequency of 2.4 kHz (Figure 8).

Acris species

The northern cricket frog (*Acris crepitans*) is a small (up to 3 cm) frog that breeds from late April to early August. It can be found at the edges of permanent pools and streams (Phillips

et al. 1999). Females breed once a year (Gray 1984). The advertisement call of the cricket frog can be described as a series of high-pitched (2.7-4.0 kHz) clicks resembling marbles being clanked together, starting out slowly and then increasing in rate (Figure 8; Ryan *et al.* 1992).

Anaxyrus species

The American Toad (*Anaxyrus americanus*) is a large (up to 10 cm) toad that has a long breeding season from early May to early September. During the breeding season, it can be found in shallow pools adjacent to forests, prairies, agricultural fields, or urban areas (Phillips *et al.* 1999). The advertisement call of the American toad is a long (up to 30 seconds) high-pitched trill (about 1700 Hz; Figure 8). Previous research has shown that American toads do not adjust calling rate during times of increased traffic noise (Vargas-Salinas *et al.* 2014). Sites with varying levels of noise maintain the same call rate, amplitude, and frequency. These attributes also stay consistent in response to car noise playbacks (Cunnington & Fahrig 2010).

2.3.5 Acoustic analysis

Initial recordings were split up into two 30-second samples (at the beginning and end of each period) per treatment using the Audacity digital audio editor (version 2.2.2). We collected data only from the two 30-second samples (as opposed to the full 15 minutes) in order to reduce the overall workload, which was substantial considering the number of experiments conducted and attributes analyzed. We did not analyze any calls during the washout periods.

We used Song Scope bioacoustics software (version 4.1.5; Wildlife Acoustics, Inc., Concord Massachusetts, USA) to listen to and view these samples in spectrograms and to count the number of calls produced by each species. Recognizers were used at times to find calls within recordings, but all calls were manually checked (Bielinski Dissertation Chapter 1). Incomplete calls at the beginning or end of the subsample were still counted. After counting the

calls, we selected 3 complete calls at random from each sample (a maximum of 6 calls per period) to calculate the following call attributes: length, dominant frequency and pulse rate. Many frog calls produce multiple levels of spectral shaping, called formants, which result from acoustic resonance of the frog's anatomy. For the dominant frequency, we gathered data solely on the loudest of the formants.

Call attributes were measured by using spectrogram, log-transformed waveform, and slice plots of our recordings available in the Song Scope program. Attributes were chosen based on their utilization by previous studies on this topic, and because they are needed to investigate our soundscape hypotheses (Pijanowski *et al.* 2011; Villanueva-Rivera 2014).

A few species-specific adjustments had to be made in the protocol in order to ameliorate problems deriving from unique properties of calls. For the American toad, because their calls are so long, if any calls were cut off by the beginning or end of the 30-second samples, we extended the duration of sampling to gather data on the entire calls. For the wood frog, some of our sampling missed calling for the entire experiment. When this was the case, we collected call data from the entire 15-min treatments.

2.3.6 Statistical analysis

In general, our statistical strategy was to first compare control treatments to pooled noise treatments (car, airplane, and train together) from the Noise Source Experiment, in order to determine if the frogs do respond to anthropogenic noise in general. Next, if a response was found, the separate noise treatments were compared against each other to determine if there is a differential response to noise depending on its source. Finally, we compared the treatments form the Noise Power Experiment against the control (usually around 55 to 65 dB) in order to

determine if a threshold effect exists (i.e., is there a noise amplitude between 70 dB and 90 dB at which changes in calling behavior appear?).

Multivariate Statistics

We first ran several PERMANOVAs on the pooled data with all species. The data consisted of all response variables, which included the number of calls, the average dominant frequency, the average call length, and the average pulse rate per sample. We scaled (via z-scoring) the data because of the differences in units, and then created a distance matrix using Gower's symmetrical distance, which is appropriate for mixed-type data where zeros in the data indicate some type of response. Each PERMANOVA tested for a treatment by species interaction, given an offset of the amount of calling in the initial conditions before the experiment started. The offset was included because the soundscape of a pond can vary dramatically based on how much calling is occurring. The first PERMANOVA tested for differences in calling between the pooled noise treatments compared to the control. The second PERMANOVA tested for differences in calling between the three treatments in the Noise Source Experiments. The last PERMANOVA tested for differences in calling between the treatments of the Noise Power Experiments along with the control. We ran each PERMANOVA with 999 permutations.

Next, we ran a Principal Coordinate Analysis (PCoA) of the full multivariate dataset with all species pooled. In order to investigate the effect of taxonomy and natural history on the positioning of call attributes in multivariate space, we used color, shape, and encircling to group the data by the following factors of interest: treatment, species, genus, breeding season, habitat preference, and breeding season length (Table 2). Clumping by any of these factors would

indicate within-group similarities of call attributes, and possibly hint at underlying ecological or physiological mechanisms dictating alterations of calling behavior.

Finally, we conducted a cluster analysis using the Gower's distance matrix. First, the data was clustered using several common clustering methods and their fit was compared via cophenetic distances, as described in Borcard *et al.* (2011). From this criterion, we determined that the Unweighted Pair-Group Methods using arithmetic Averages (UPGMA) best fit the data. This is a bottom-up clustering method that considers previously grouped objects as a single object by averaging the within-group distances. In order to determine how many interpretable clusters there are in the dendrogram, we compared silhouette widths for k number of groups from 2 to 30. The initial optimal number of groups (two) was ignored as it would not be helpful in investigating applicable ecological patterns. We decided on 10 groups. We then created dendrograms with an overlay of the suggested groups and colored the leaves by factors the of interest: species, genus, breeding season, habitat preference, breeding season length and treatment. We compared the suggested groups to the spread of these traits. Our multivariate analyses suggested that we look at univariate responses to noise per species.

Univariate Statistics

We ran statistical models to determine the effect size and 95% confidence intervals (CIs) across the different treatment combinations for each response variable, separated by species. Figure 9 shows a diagram of the types of models considered for this study. To model the call rates, most distributions were non-normal, so we could not use linear models. Additionally, the datasets were zero-inflated to the extent that conventional generalized linear models using Poisson or negative binomial families were also a poor fit. Instead, we used hurdle models to fit the data. A hurdle model is a split (two part) generalized linear model specialized for zero-

inflated data. The first part of the model ("zero model") uses a binomial distribution to model

zeros and non-zeros in the data. The second part of the model ("count model") takes the non-

zero data and fits them to a negative binomial or a Poisson distribution (Zeileis *et al.* 2008;

Zeileis & Kleiber 2018). The fit of each Hurdle model was evaluated using rootograms (square-

root-transformed histograms with a fitted model curve), quantile-quantile plots, and residual

plots.

Distributions of the other variables were not zero-inflated. We used linear models (LMs)

or linear mixed-effect models (LMEMs) that incorporated blocking by each experimental night

as a random factor. This blocking was only used when it improved the model fit and is justified

by the fact that we did not control for changes in the environment (temperature, humidity, wind,

initial soundscape) from one night to the next. These confounding variables could influence the

calling behavior of the frogs but are not the focus of this study. For some LMs and LMEMs, the

distribution of the data was skewed and required square-root or log transformation. The fit of

each LM and LMEM was evaluated using residual plots and quantile-quantile plots.

For every response variable by species combination, after the data were fit to a model, we

plotted the effect size of each treatment with 95% CIs. Doing so at times required back-

transformation of model coefficients and confidence intervals using exponent, square, or inverse

logit functions, depending on the model. We used a summary table to compare the effect sizes

and strength of evidence (i.e., 95% CIs and P-values) across species and genera.

For the sake of the digestibility of such a large amount of results, we condensed effect

sizes and into several different categories. For effect sizes, changes of 0-10% were considered

"small" and symbolized in figures with a skinny arrow. Changes of 11-50% were considered

"intermediate" and symbolized with a medium-sized arrow. Changes of greater than 50% were

considered "large" and symbolized with a large arrow with thick outlines. These categories were

determined after plotting all effect sizes and spotting natural breaks in the distribution. P-values

in figures are symbolized as follows: $P < 0.1$ was given no stars, $P < 0.05$ was given one star, P

< 0.01 was given two stars, and $P < 0.001$ was given three stars. Results with P-values > 0.1

were not included in our summary figures, but are available in the Appendices (Section 2.9). Our

statistical approach that combines effect size along with an indication of the P-value allowed us

to take a more nuanced view of the results and avoid the pitfalls of reliance on dichotomous null-

hypothesis significance testing (Wasserstein *et al.* 2019).

We conducted all statistical analyses on R version 3.5.0. (R Core Team 2019).

2.4 RESULTS

In total, we conducted 155 playback experiments (387.3 h recorded), counted 51,537

calls, and analyzed the attributes of 3,725 calls. The public can access our raw data online

through the Center for Open Science data repository (https://osf.io/xwzjk/). As mentioned earlier,

data collection was not balanced across our nine target species. The number of experiments

conducted per species varied from two to 40 (Figure 2). Our target community represents a large

range in dominant frequency (Figure 10). For the univariate modelling, all output for species $\times$

response variable combinations that are not mentioned in the results are in the appendices. These

include box plots and effect-size plots.

2.4.1 Multivariate Analysis

For the comparison of call attributes between pooled noise and no noise, the

PERMANOVA indicated a small effect of treatment ($R^2 = 0.011$, $P[\text{Pseudo-F}_{1,462}] = 0.005$) and

a large effect of species ($R^2 = 0.58$, P[Pseudo-$F_{9,462}$] < 0.001). There was no evidence of a treatment × species interaction ($R^2 = 0.003$, P[Pseudo-$F_{9,462}$] = 0.777).

For the comparison of car, airplane, and train treatments, the PERMANOVA indicated no effect of treatment ($R^2 = -0.004$, P[Pseudo-$F_{9,225}$] = 0.948), a large effect of species ($R^2 = 0.665$, P[Pseudo-$F_{9,225}$] < 0.001,) and no evidence of a treatment × species interaction ($R^2 = -0.010$, P[Pseudo-$F_{18,225}$] = 0.994).

For the Noise Power Experiment, the PERMANOVA indicated a small effect of treatment ($R^2 = 0.009$, P[Pseudo-$F_{3,248}$] = 0.010) and a large effect of species ($R^2 = 0.808$, P[Pseudo-$F_{8,248}$] < 0.001). There was no evidence of a treatment × species interaction ($R^2 = -0.001$, P[Pseudo-$F_{24,248}$] = 1.000).

The PCoA shows clear separation by species, and by genus (Figure 11). When overlaying the ordination with other factors of interest, there is no obvious grouping by hydrology preference (Appendix B), breeding season timing (Appendix C), or breeding season length (Appendix D). The variation in the call attribute data is explained more by taxonomy than by response to noise treatments, which parallels the results of the PERMANOVAs. Within each species there is no clumping by treatment (Appendix E). This also agrees with the absence of an interaction in the PERMANOVAs .

Clustering the multivariate data using the UPGMA method yielded a dendrogram with grouping by genus and species, although the recommended groups did not follow exactly along these designations. There was no additional clustering added by the rest of our grouping factors Figure 12).

The results from the multivariate approach included a lot of statistical noise and indicated a need for a univariate analysis. The small or absent effect of the treatments from the

PERMANOVAs could possibly be because the frogs did not alter most call attributes, or that

most species did not react to the treatments. However, the multivariate approach does not

provide information on exceptions to these patterns. Thus, we decided to investigate the species

× response variable combinations in a univariate manner to uncover any exceptional behavior.

2.4.2 Univariate Analysis

Pooled noise versus control

Overall patterns between the pooled noise and control treatments show a reduction in call

count, but limited changes to dominant frequency, length and pulse rate (Figure 13). A summary

of all model outputs where $P < 0.10$ can be seen in Table 3. All species except the spring peeper

reduced their call rate. Moreover, the spring peeper raised its dominant frequency, which was the

only instance of altered frequency in response to noise. Three species changed their call length

in response to noise. The western chorus frog and the northern cricket frog shortened their calls,

whereas the Cope's gray treefrog lengthened its calls. Two species changed their pulse rate in

response to noise. The Cope's gray tree frog increased its pulse rate, whereas the American toad

decreased its pulse rate. Box plots and effect size graphs of every comparison are located in the

Appendices.

Car versus airplane versus train treatments

We looked for differences in call attributes between the car, airplane, and train treatments

only if we identified an effect of noise (above) or an effect of the car treatment (below). Figure

14 summarizes the nine instances when differences in call attributes were detected and Table 4

includes model outputs where $P < 0.10$.

Two species had different call rates depending on the type of noise broadcasted. The

American bullfrog called less during the airplane treatment compared to the car and train

treatments (Figure 15). The American toad called more during the airplane treatment compared to the car and train treatments (Figure 16). Only the spring peeper altered its dominant frequency, with a decrease in frequency during the airplane treatment compared to the car treatment (Figure 17).

Four species showed changes in call length in response to the different noise treatments. The western chorus frog lengthened its calls during the airplane treatment compared to the car and train treatments (Figure 18). The Cope's gray tree frog shortened its calls during the airplane treatment compared to the car treatment (Figure 19). The northern cricket frog shortened its calls during the airplane treatment compared to the car treatment (Figure 20). The American toad shortened its calls during the airplane treatment compared to the train treatment (Figure 21). Only the Cope's gray tree frog altered its pulse rate, with a rate increase during car treatment compared to the airplane and the train treatments (Figure 22).

Car at 70, 80 and 90 dB versus Control

Results from the Noise Power Experiment reveal that louder treatments altered calling behavior more than quieter treatments (Figure 23). Additionally, most changes were seen in call rate, although all measured response variables displayed some changes. A summary of all model outputs where $P < 0.10$ can be seen in Table 5.

Six of the nine species reduced their call rate in response to car noise, with two species (the Cope's gray treefrog and American toad) calling less during the 80 dB treatment and continuing through the 90 dB treatment. The remaining four species (the bullfrog, wood frog, western chorus frog, and northern cricket frog) called less only during 90 dB treatment. There were no changes recorded during the 70 dB treatment.

Four species altered their dominant frequencies in response to the car noise treatments. However, only one of these species (the spring peeper) altered its dominant frequency during the 90 dB treatment, which is consistent with results from the two previous comparisons. The western chorus frog and eastern gray treefrog lowered their call frequency in the 70 dB treatment only, and the American toad lowered its call frequency in the 80 dB treatment only.

Although six of the nine species altered their call length in response to the car noise treatments, there was no clear pattern across species or treatments. The western chorus frog shortened its calls during the 80 and 90 dB treatments. The Cope's gray treefrog lengthened its calls during the 70 and 90 dB treatments. The Northern cricket frog shortened its calls in the 90 dB treatment only. For the three remaining species (the green frog, wood frog, and American toad), we detected changes in call length, but they were not in response to the 90 dB treatment.

For the five species from which we measured pulse rate, four species altered pulse rate in response to the car noise treatments. Both the western chorus frog and the northern cricket frog increased their pulse rate in the 80 and 90 dB treatments. The American toad lowered its pulse rate in the 90 dB treatment. The cope's gray tree frog increased its pulse rate only during the 70 dB treatment. Box plots and effect size graphs of every car noise comparison are located in the Appendices.

2.5 DISCUSSION

2.5.1 Call rate

The lack of interaction effect present in the noise versus control PERMANOVA indicated that although species responded to the treatments, in general they responded in the same way. In other words, the overall response to noise looked similar from species to species. The drop in call

rate for eight of our nine species (with one showing no change) was the main contributing factor to this pattern.

The observed consistency in lower call rates across eight species is notable considering that previous studies have shown mixed results (increased calling: Sun & Narins 2005; Kaiser & Hammers 2009; Bleach *et al.* 2015; Engbrecht *et al.* 2015, Kruger and Du Preez 2016; decreased calling: Sun & Narins 2005; Caorsi 2017). A decrease in call rate in response to noise may be adaptive on the individual level, as resources are not wasted creating expensive calls during times when females may have difficulty hearing or optimally comparing multiple suitors. This response might only be present in populations that are habituated to higher noise levels (Cunnington & Fahrig 2010)

All populations of this study are relatively "urban," given that study locations were all within 60 miles of Chicago's city limits, so it possible that these species have developed the same strategy over time in response to chronic noise, whereas populations in a "rural" or pristine setting may react differently. Previously published increased call rates in response to noise could be from frog populations with limited exposure to anthropogenic noise. In these populations noise may invoke calling if it triggers the same competitive mechanism that is activated during conspecific chorusing. In other words, there may be a divergence of behaviors in urban versus non-urban frog populations. This hypothesis cannot be addressed with our experimental approach; however, Halfwerk *et al.* (2019) found that urbanization does drive signal changes. Calls from urban populations of Tungara frogs were more complex and attractive to both urban and forested females compared to calls from forested populations. When these two groups were translocated, the urban frogs were able to shift back towards the normal phenotype, whereas the forest frogs did not make any changes. Therefore, the urban frogs had higher phenotypic

66

plasticity. It would be worthwhile to continue looking for similar behavioral divergences in other species.

2.5.2 Call frequency

Frequency shifts were unexpectedly rare in our study, considering that many past publications have documented shifts (Parris *et al.* 2009; Cunnington & Fahrig 2010; Hoskin & Goosem 2010; Hanna *et al.* 2014; Kruger and Du Preez 2016; Caorsi 2017; Grenat *et al.* 2019). It is important to note that the Spring Peeper, the only species to not reduce its call rate, was also the only species that raised its dominant frequency. This suggests that these two strategies (lowering call rate; raising dominant frequency) are dichotomous choices to solve the same problem. Species that raise their frequency may not need to limit their calling. This dichotomous tradeoff between call rate and dominant frequency was further supported by our results of the Noise Power Experiment.

Shifts in dominant frequency cannot be too large because it is a relatively static property that is anchored to body size and is used by females to distinguish between species (Gerhardt 1991; Gerhardt 1994). Thus, if changes in frequency in response to noise were too extreme, then calls may not attract females, regardless of how well this modified call stands out amongst different background noises. It would be useless for low-frequency species (like the bullfrog) to raise their call frequency if they do not have the plasticity to break free of the range of noise frequency. Notably, our noise treatments still have energy signals up to about 3 kHz (although vast majority is concentrated in lower frequency bands), so if males cannot call with a higher pitch than this level, or if females cannot identify conspecifics at this level, they should not alter their frequency at all. The spring peeper has dominant frequency of 2.9 kHz, the third highest in our study to the western chorus frog northern cricket frog. This dominant frequency may have

enough spectral distance from the traffic noise that an increase in frequency will allow the species to further stand out, while still being identifiable to females. Therefore, only species calling at the spectral margins of a noise signal may have the choice of altering dominant frequency in lieu of changing their call rate. More research is needed to investigate why northern cricket frog (with a dominant frequency of ~3kHz) or the western chorus frog (with a dominant frequency ~3-3.5 kHz) reduced their call rates instead of shifting their dominant frequency.

2.5.3 Call length and pulse rate

There was no noticeable pattern in the changes of call length or pulse rate across species. The inconsistency shown in the effects could be a manifestation of the variation within and among males, attributable to the underlying competition for females. Males can adjust their calls depending on the competitive environment and potential to mate with nearby females. (Bee *et al.* 2000; Martinez-Rivera & Gerhardt 2008; Jiang *et al.* 2015; Zhu *et al.* 2017). For example, longer calls are more attractive to eastern gray treefrogs (Gerhardt *et al.* 2000). Males can compete in an "arms race" to produce the best signal(s), with the evolutionary tradeoff of using more energy. This competition is always happening in a chorus setting, regardless of the playback treatment. Thus, we attribute the effects of call length and pulse rate to noise-independent variation.

A hypothetical reduction in either call length or pulse rate would save energy, similar to a reduction in call rate. We expected to see a reduction in call length (which we did see in the western chorus frog and northern cricket frog in both the experiments), but it could be that for most of our species, lowering the call rate suffices to save energy during the noise treatments.

2.5.4 Noise type

For the Noise Type Experiment, the evidence for a differential response to different anthropogenic noises was not strong. The Noise Type PERMANOVA showed that the community did not respond differently to the three different types of anthropogenic noise, which implies that simple properties of noise like frequency and amplitude play a larger role in affecting frog call behavior than smaller details like how quick or sporadic the noise spikes are, or whether there are sudden bursts of energy at a new frequency (like train horns).

We detected few effects in the univariate analysis, and the pattern of these effects was not consistent across species. For example, in the two species where call rate did change, the bullfrog called least during the airplane treatment, whereas American toad called least during the train treatment. There were also clear irregularities for call length across species, none of which lined up with our prediction that calls would be shortest during the car treatment. Due to these inconsistencies, there is no evidence to support that any changes between treatments is due to the ability of frogs to perceive different anthropogenic noises *and* act differently as a response. If this were the case, we could have seen more calling during the airplane and train treatments, because there are longer breaks between the trains and planes compared to the cars. Therefore, we consider the measured changes in call attributes as natural variation within the population or individuals (Bielinski Dissertation Chapter 3), or due to a response to unaccounted factors like wind (Halfwerk *et al.* 2016b), predation risk (Gomes *et al.* 2016) or male-male competition (Bee *et al.* 2000; Bee & Gerhardt 2001).

2.5.5 Multivariate grouping patterns

The multivariate ordination and clustering analysis did not reveal any patterns other than grouping by genus and species. This provides evidence for the changes in call behavior being genetically influenced. Our data did not separate by hydrology preference, breeding season, or

length of breeding season. The multivariate analysis was useful to determine whether there was a treatment × species interaction, which we did not find. Within species, there was also no grouping by treatment, a consequence of only having one call attribute consistently changing in response to anthropogenic noise, with the rest of the response variables only adding statistical noise to the analysis. The univariate analysis provided the most insight into how frogs respond to anthropogenic noise because only one variable (call rate) was consistently affected.

2.5.6 Taxonomic observations

Across all three comparisons, species in the genus *Lithobates* show the lowest amount of phenotypic plasticity. They did show a reduction in the number of calls, but the rest of the call attributes were relatively unchanged. *Lithobates spp.* can alter their call frequency, especially during territorial disputes (Bee *et al.* 2000). Thus, there we see no evidence here that *Lithobates spp.* misinterpreted our noise treatments as competing males. Male *Lithobates* may be correctly identifying the treatments as non-biological, and their subsequent response is to call less.

The Cope's gray treefrog appears to possess more phenotypic plasticity compared to the eastern gray treefrog. These two species are a diploid-tetraploid species complex that diverged several times independently in the past three to fifteen million years (Ptacek *et al.* 1994). They are physically "identical" and only have slight variations in their calls (Phillips *et al.* 1999). The Cope's gray tree frog exhibited greater change in the number of calls, call length, and pulse rate in both the noise versus control comparison and the Noise Power Experiment. It may be that we did not monitor the proper call attributes for the eastern gray tree frog, as Bee (2015) states that eastern gray tree frog females tend to asses pulse duration and inter-pulse interval (traits that were not monitored here), whereas Cope's gray tree frog females assess pulse rate. These patterns fit our findings.

2.5.7 Experimental versus actual noise levels

Questions may arise as to how often urban frogs realistically encounter noise levels from this study. The SPLs of 80 and 90 dB may be overestimations of the noise level needed to alter frog calling for two reasons: 1) The average SPL for the full 15-minute car recording was about 10 dB less than the peak SPL (i.e. the 90 dB treatment had peaks at 90 dB but an average of 80 dB). 2) Our noise was broadcasted from a point source (speakers), compared to actual traffic, which is a line source. Noise from a point source diminishes by 6 dB per doubling of distance, whereas noise from a line source diminishes by 3 dB per doubling of distance (Rossing 2007). Therefore, frogs located further into a pond (and farther from noise sources) would experience less noise from our 90 dB broadcast at the pond edge compared to traffic that produces an equivalent SPL at that same point.

Noise levels from previous research indicate that a considerable proportion of ponds in urban areas may experience noise at the levels shown to modify behavior in this study. Parris *et al.* (2009) found that noise levels from their 47 sites around Melbourne, Australia had average SPLs (across an 18-hr interval) between 43 and 79 dB. Moreover, Cunnington & Fahrig (2010) identified 39 "high noise sites" around Ontario, Canada, which had an average SPL of 73 dB (SD = 4.9).

2.5.8 Threshold Effects

In the Noise Power Experiment, the community had a larger magnitude of response to louder noise treatments, as indicated by the effect of car noise level in the PERMANOVA. For the univariate analysis, we interpreted responses as "threshold effects" only if they continue through the 90 dB level. In other words, effects that occur during the 70 and/or 80 dB treatments but disappear in the loudest treatment are probably not caused by noise *per se*. For example,

changes in dominant frequency in the western chorus frog and eastern gray tree frog in only the 70 dB treatment are more likely due to natural variation in calls within a population, and not due to a response to noise. We see evidence of some unsubstantiated variation in the 70 and 80 dB treatments (i.e. with no measured difference between the 90 dB treatment and control) across all response variables besides call rate.

Changes to call attributes were largest during the 90 dB treatments, which is consistent with our prediction and previous research (Lengagne 2008; Halwerk *et al.* 2016). There were seven instances where an effect was seen at the 90 dB level only, five instances where a threshold was seen at the 80 dB treatment (meaning that the effect was seen in the 90 dB treatment as well), and zero instances where we saw a threshold effect beginning at the 70 dB treatment. The 70 dB treatment is relatively low considering that the SPL of a chorus can be between 75-85 dB, depending on the number of callers (Bee 2015). The lack of response to the 70 dB treatment is consistent with the idea that males are already accustomed to a certain level of noise from choruses. Invoked behavioral changes at the 90 and 80 dB level mean that for our target populations, noise may only be affecting calling behavior in the loudest habitats.

Past research on frog distributions across developed landscapes suggests that road-side wetlands (those that are likely to be noisiest) are a major contribution to maintaining frog populations. A regional survey of 53 sites in suburban landscapes in Baltimore County, Maryland, found that 89% of the wetlands with breeding frogs are artificial or stormwater ponds, which are usually situated beside roads. Species found in this survey that are common with our study include the wood frog, American toad, Cope's gray tree frog, and green frog (Brand & Snodgrass 2010). Additionally, studies show that road-side ditches can help maintain frog diversity and migration (Homyack *et al.* 2014; Mazerolle 2005; Soomets *et al.* 2017).

Considering that traffic noise that spreads through 250 ft of forest can still reach 73 dB (Reethof & Heisler 1976), it is not hard to imagine that many pools in urban landscapes may be affected by anthropogenic noise levels used in this study.

2.5.9 Further implications

Changes in calling behavior as a response to anthropogenic noise over many generations could alter the evolutionary trajectory of urban frog populations compared to their rural counterparts. Because frog breeding is so acoustically focused, there could be a divergence between populations depending on the soundscape of habitats. Previous studies already show early evidence of this. Homola *et al.* (2019) investigated the single nucleotide polymorphisms of urban and rural wood frogs and found evidence of directional selection associated with urbanization and Halfwerk *et al.* (2019) found differences in calling plasticity between populations of urban and forested Tungara frogs.

Noise may also be altering landscape-level population dynamics. After frogs metamorphose, individuals can disperse over long distances in order to find a new breeding pond (Berven & Grudzien 1990, Pizzatto *et al.* 2017). Once a new breeding pond is found, frogs tend to have strong breeding site fidelity and return to the pond year after year. In order to find new ponds (disperse) or return to old ponds (migrate), frogs use phonotaxis of conspecific and heterospecific calls (Gerhardt & Klump 1988; Bee 2007, James *et al.* 2015).

In a noisy landscape, frogs will have a more difficult time homing in on ponds because noise will both cause resident frogs to call less and interfere with calls that do occur. This could be especially harmful for frogs that prefer ephemeral breeding, as they live in a forested landscape dotted with pools that blink in and out of existence. This "ponds as patches" point of view (Murray *et al.* 2015) stresses the importance of a heterogenous landscape that is required

for frogs to thrive (Knutson *et al.* 1999). If the landscape is noisy, many of these ephemeral pools may be functionally out of reach to mobilized frogs.

Soundscape should be of interest for reintroduction programs where frogs of once-extirpated populations are brought back to a landscape either from another natural site or a captive breeding program. Wood frog reintroduction has already happened in the Chicago region and this will likely be a more common strategy in the future as cities continue to grow and amphibians continue to decline (Sacerdote 2009; Passos *et al.* 2017). Wildlife managers would like the restored population to eventually spread out from their initial reintroduction point. Therefore, when assessing potential reintroduction sites, conservationists should consider the soundscape of each area because noise could reduce habitat quality and slow down the inter-pond dispersal of frogs.

2.5.10 Research contributions

Table 1 shows the contribution of our playback experiments in the context of past research of the effects of anthropogenic noise on frog calling behavior. Our study clearly provides much-needed data on a regional community of species. Many field and lab playback experiments cannot look for community-level responses because they tend to test one or two species. This makes it harder to discover common patterns across species because it is difficult to compare studies with varying methodologies (Brooke *et al.* 2000; Kaiser & Hammers 2009; Hanna *et al.* 2014; Engbrecht *et al.* 2015; Shen & Xu 2016). Our study was also the only to compare three types and three SPL levels of anthropogenic noise, while using a larger dataset (155 experiments) than any of the other publications listed. Future playback studies should: (1) target new communities to ask questions about taxonomy and natural history, and (2) compare

responses from urban versus natural populations of the same species to explore whether urban species have developed new strategies for dealing with anthropogenic noise.

Acknowledgments

2.6 LITERATURE CITED

Barinaga, M. 1990. Where have all the froggies gone. Science 247:1033-1034.

Bee, M.A. 2007. Selective phonotaxis by male wood frogs (rana sylvatica) to the sound of a chorus. Behavioral Ecology and Sociobiology 61:955-966.

Bee, M.A. 2010. Spectral preferences and the role of spatial coherence in simultaneous integration in gray treefrogs (hyla chrysoscelis). Journal of Comparative Psychology 124:412-424.

Bee, M.A. 2015. Treefrogs as animal models for research on auditory scene analysis and the cocktail party problem. International Journal of Psychophysiology 95:216-237.

Bee, M.A., and H.C. Gerhardt. 2001. Neighbour-stranger discrimination by territorial male bullfrogs (rana catesbeiana): I. Acoustic basis. Animal Behaviour 62:1129-1140.

Bee, M.A., S.A. Perrill, and P.C. Owen. 2000. Male green frogs lower the pitch of acoustic signals in defense of territories: A possible dishonest signal of size? Behavioral Ecology 11:169-177.

Bee, M.A., A. Velez, and J.D. Forester. 2012. Sound level discrimination by gray treefrogs in the presence and absence of chorus-shaped noise. Journal of the Acoustical Society of America 131:4188-4195.

Benedix, J.H., and P.M. Narins. 1999. Competitive calling behavior by male treefrogs, eleutherodactylus coqui (anura : Leptodactylidae). Copeia:1118-1122.

Berven, K.A., and T.A. Grudzien. 1990. Dispersal in the wood frog (rana-sylvatica) - implications for genetic population-structure. Evolution 44:2047-2056.

Bevier, C.R. 1997. Breeding activity and chorus tenure of two neotropical hylid frogs. Herpetologica:297-311.

Bleach, I.T., C. Beckmann, C. Both, G.P. Brown, and R. Shine. 2015. Noisy neighbours at the frog pond: Effects of invasive cane toads on the calling behaviour of native australian frogs. Behavioral Ecology and Sociobiology 69:675-683.

Borcard, D., F. Gillet, and P. Legendre. 2011. Numerical ecology with r.

Brand, A.B., and J.W. Snodgrass. 2010. Value of artificial habitats for amphibian reproduction in altered landscapes. Conservation Biology 24:295-301.

Brooke, P.N., R.A. Alford, and L. Schwarzkopf. 2000. Environmental and social factors influence chorusing behaviour in a tropical frog: Examining various temporal and spatial scales. Behavioral Ecology and Sociobiology 49:79-87.

Calhoun, A.J., N.A. Miller, and M.W. Klemens. 2005. Conserving pool-breeding amphibians in human-dominated landscapes through local implementation of best development practices. Wetlands Ecology and Management 13:291-304.

Caorsi, V.Z., C. Both, S. Cechin, R. Antunes, and M. Borges-Martins. 2017. Effects of traffic noise on the calling behavior of two neotropical hylid frogs. Plos One 12.

Carr, L.W., and L. Fahrig. 2001. Effect of road traffic on two amphibian species of differing vagility. Conservation Biology 15:1071-1078.

Crouch, W.B., and P.W.C. Paton. 2002. Assessing the use of call surveys to monitor breeding anurans in rhode island. Journal of Herpetology 36:185-192.

Cunnington, G.M., and L. Fahrig. 2010. Plasticity in the vocalizations of anurans in response to traffic noise. Acta Oecologica-International Journal of Ecology 36:463-470.

Cushman, S.A. 2006. Effects of habitat loss and fragmentation on amphibians: A review and prospectus. Biological Conservation 128:231-240.

Dananay, K.L., K.L. Krynak, T.J. Krynak, and M.F. Benard. 2015. Legacy of road salt: Apparent positive larval effects counteracted by negative postmetamorphic effects in wood frogs. Environmental Toxicology and Chemistry 34:2417-2424.

Engbrecht, N.J., J.L. Heemeyer, C.G. Murphy, R.M. Stiles, J.W. Swan, and M.J. Lannoo. 2015. Upland calling behavior in crawfish frogs (lithobates areolatus) and calling triggers caused by noise pollution. Copeia 103:1048-1057.

Forti, L.R., A. Forti, R. Marquez, and L.F. Toledo. 2017. Behavioural response evoked by conspecific distress calls in two neotropical treefrogs. Ethology 123:942-948.

Gerhardt, H.C. 1991. Female mate choice in treefrogs - static and dynamic acoustic criteria. Animal Behaviour 42:615-635.

Gerhardt, H.C. 1994. The evolution of vocalization in frogs and toads. Annual Review of Ecology and Systematics 25:293-324.

Gerhardt, H.C., and G.M. Klump. 1988. Phonotactic responses and selectivity of barking treefrogs (hyla gratiosa) to chorus sounds. Journal of Comparative Physiology A 163:795-802.

Gerhardt, H.C., C.C. Martínez-Rivera, J.J. Schwartz, V.T. Marshall, and C.G. Murphy. 2007. Preferences based on spectral differences in acoustic signals in four species of treefrogs (anura: Hylidae). Journal of Experimental Biology 210:2990-2998.

Gerhardt, H.C., S.D. Tanner, C.M. Corrigan, and H.C. Walton. 2000. Female preference functions based on call duration in the gray tree frog (hyla versicolor). Behavioral Ecology 11:663-669.

Gibbs, J.P., K.K. Whiteleather, and F.W. Schueler. 2005. Changes in frog and toad populations over 30 years in new york state. Ecological Applications 15:1148-1157.

Godwin, G.J., and S.M. Roble. 1983. Mating success in male treefrogs, hyla chrysoscelis (anura: Hylidae). Herpetologica:141-146.

Gomes, D.G.E., R.A. Page, I. Geipel, R.C. Taylor, M.J. Ryan, and W. Halfwerk. 2016. Bats perceptually weight prey cues across sensory systems when hunting in noise. Science 353:1277-1280.

Grafe, T.U., R. Schmuck, and K.E. Linsenmair. 1992. Reproductive energetics of the african reed frogs, hyperolius viridiflavus and hyperolius marmoratus. Physiological Zoology 65:153-171.

Gray, R.H. 1984. Effective breeding size and the adaptive significance of color polymorphism in the cricket frog (acris crepitans) in illinois, USA. Amphibia-Reptilia 5:101-107.

Grenat, P.R., F.E. Pollo, M.A. Ferrero, and A.L. Martino. 2019. Differential and additive effects of natural biotic and anthropogenic noise on call properties of odontophrynus americanus (anura, odontophryinidae): Implications for the conservation of anurans inhabiting noisy environments. Ecological Indicators 99:67-73.

Halfwerk, W., M. Blaas, L. Kramer, N. Hijner, P.A. Trillo, X.E. Bernal, R.A. Page, S. Goutte, M.J. Ryan, and J. Ellers. 2019. Adaptive changes in sexual signalling in response to urbanization. Nature ecology & evolution 3:374.

Halfwerk, W., A.M. Lea, M.A. Guerra, R.A. Page, and M.J. Ryan. 2016a. Vocal responses to noise reveal the presence of the lombard effect in a frog. Behavioral Ecology 27:669-676.

Halfwerk, W., M.J. Ryan, and P.S. Wilson. 2016b. Wind- and rain-induced vibrations impose different selection pressures on multimodal signaling. American Naturalist 188:279-288.

Hamer, A.J., and M.J. McDonnell. 2008. Amphibian ecology and conservation in the urbanising world: A review. Biological Conservation 141:2432-2449.

Hanna, D.E.L., D.R. Wilson, G. Blouin-Demers, and D.J. Mennill. 2014. Spring peepers pseudacris crucifer modify their call structure in response to noise. Current Zoology 60:438-448.

Hels, T., and E. Buchwald. 2001. The effect of road kills on amphibian populations. Biological Conservation 99:331-340.

Homyack, J.A., C.J. O'Bryan, J.E. Thornton, and R.F. Baldwin. 2014. Anuran assemblages associated with roadside ditches in a managed pine landscape. Forest Ecology and Management 334:217-231.

Hoskin, C.J., and M.W. Goosem. 2010. Road impacts on abundance, call traits, and body size of rainforest frogs in northeast australia. Ecology and Society 15.

Howard, R.D. 1979. Estimating reproductive success in natural populations. The American Naturalist 114:221-231.

James, M.S., M.P. Stockwell, J. Clulow, S. Clulow, and M.J. Mahony. 2015. Investigating behaviour for conservation goals: Conspecific call playback can be used to alter amphibian distributions within ponds. Biological Conservation 192:287-293.

Jiang, F., G.Z. Fang, F. Xue, J.G. Cui, S.E. Brauth, and Y.Z. Tang. 2015. Male music frogs compete vocally on the basis of temporal sequence rather than spatial cues of rival calls. Asian Herpetological Research 6:305-316.

Kaiser, K., and J.L. Hammers. 2009. The effect of anthropogenic noise on male advertisement call rate in the neotropical treefrog, dendropsophus triangulum. Behaviour 146:1053-1069.

Knutson, M.G., J.R. Sauer, D.A. Olsen, M.J. Mossman, L.M. Hemesath, and M.J. Lannoo. 1999. Effects of landscape composition and wetland fragmentation on frog and toad abundance and species richness in iowa and wisconsin, USA. Conservation Biology 13:1437-1446.

Kruger, D.J.D., and L.H. Du Preez. 2016. The effect of airplane noise on frogs: A case study on the critically endangered pickersgill's reed frog (hyperolius pickersgilli). Ecological Research 31:393-405.

Kuczynski, M.C., T. Getty, and E. Gering. 2017. Larger females are choosier in the gray treefrog (hyla versicolor). Behavioural Processes 135:29-35.

Lengagne, T. 2008. Traffic noise affects communication behaviour in a breeding anuran, hyla arborea. Biological Conservation 141:2023-2031.

Luddecke, H. 2002. Male and female responses to call playback in the andean frog colostethus subpunctatus. Amphibia-Reptilia 23:141-150.

Martinez-Rivera, C.C., and H.C. Gerhardt. 2008. Advertisement-call modification, male competition, and female preference in the bird-voiced treefrog hyla avivoca. Behavioral Ecology and Sociobiology 63:195-208.

Mazerolle, M.J. 2005. Drainage ditches facilitate frog movements in a hostile landscape. Landscape Ecology 20:579-590.

Mazerolle, M.J., A. Perez, and J. Brisson. 2014. Common reed (phragmites australis) invasion and amphibian distribution in freshwater wetlands. Wetlands Ecology and Management 22:325-340.

McCallum, M.L. 2007. Amphibian decline or extinction? Current declines dwarf background extinction rate. Journal of Herpetology 41:483-491.

McCauley, S.J., S.S. Bouchard, B.J. Farina, K. Isvaran, S. Quader, D.M. Wood, and C.M. St Mary. 2000. Energetic dynamics and anuran breeding phenology: Insights from a dynamic game. Behavioral Ecology 11:429-436.

Murray, R.G., V.D. Popescu, W.J. Palen, and P. Govindarajulu. 2015. Relative performance of ecological niche and occupancy models for predicting invasions by patchily-distributed species. Biological Invasions 17:2691-2706.

Parris, K.M., M. Velik-Lord, and J.M.A. North. 2009. Frogs call at a higher pitch in traffic noise. Ecology and Society 14.

Passos, L.F., G. Garcia, and R.J. Young. 2017. Neglecting the call of the wild: Captive frogs like the sound of their own voice. Plos One 12.

Phillips, C.A., R.A. Brandon, and E.O. Moll. 1999. Field guide to amphibians and reptiles of illinois. Illinois Natural History Survey.

Phillips, K. 1990. Where have all the frogs and toads gone. Bioscience 40:422-424.

Pijanowski, B.C., L.J. Villanueva-Rivera, S.L. Dumyahn, A. Farina, B.L. Krause, B.M. Napoletano, S.H. Gage, and N. Pieretti. 2011. Soundscape ecology: The science of sound in the landscape. Bioscience 61:203-216.

Prestwich, K.N. 1994. The energetics of acoustic signaling in anurans and insects. American Zoologist 34:625-643.

Ptacek, M.B., H.C. Gerhardt, and R.D. Sage. 1994. Speciation by polyploidy in treefrogs: Multiple origins of the tetraploid, hyla versicolor. Evolution 48:898-908.

Reethof, G., and G.M. Heisler. 1976. Trees and forest for noise abatement and visual screening. USDA Forest Service General Technical Report, NE 22:39-48.

Roelants, K., D.J. Gower, M. Wilkinson, S.P. Loader, S.D. Biju, K. Guillaume, L. Moriau, and F. Bossuyt. 2007. Global patterns of diversification in the history of modern amphibians. Proceedings of the National Academy of Sciences of the United States of America 104:887-892.

Rossing, T. 2007. Springer handbook of acoustics. Springer Science & Business Media.

Ryan, M.J. 1980. The reproductive behavior of the bullfrog (rana catesbeiana). Copeia:108-114.

Ryan, M.J., S.A. Perrill, and W. Wilczynski. 1992. Auditory tuning and call frequency predict population-based mating preferences in the cricket frog, acris-crepitans. American Naturalist 139:1370-1383.

Sacerdote, A.B. 2009. Reintroduction of extirpated flatwoods amphibians into restored forested wetlands of northern illinois: Feasibility assessment, implementation, habitat restoration, and conservation implications. Northern Illinois University.

Sacerdote, A.B., and R.B. King. 2014. Direct effects of an invasive european buckthorn metabolite on embryo survival and development in xenopus laevis and pseudacris triseriata. Journal of Herpetology 48:51-58.

Schrode, K.M., J.L. Ward, A. Velez, and M.A. Bee. 2012. Female preferences for spectral call properties in the western genetic lineage of cope's gray treefrog (hyla chrysoscelis). Behavioral Ecology and Sociobiology 66:1595-1606.

Schwartz, J.J., B.W. Buchanan, and H.C. Gerhardt. 2001. Female mate choice in the gray treefrog (hyla versicolor) in three experimental environments. Behavioral Ecology and Sociobiology 49:443-455.

Shen, J.X., and Z.M. Xu. 2016. The lombard effect in male ultrasonic frogs: Regulating antiphonal signal frequency and amplitude in noise. Scientific Reports 6.

Soomets, E., A. Lohmus, and R. Rannap. 2017. Brushwood removal from ditch banks attracts breeding frogs in drained forests. Forest Ecology and Management 384:1-5.

Steelman, C.K., and M.E. Dorcas. 2010. Anuran calling survey optimization: Developing and testing predictive models of anuran calling activity. Journal of Herpetology 44:61-68.

Stratman, K.D., and G. Höbel. 2019. Opportunity for female choice in the wild is frequently curtailed by low male trait variation in hyla versicolor. Behavioral Ecology and Sociobiology 73:59.

Stuart, S.N., J.S. Chanson, N.A. Cox, B.E. Young, A.S. Rodrigues, D.L. Fischman, and R.W. Waller. 2004. Status and trends of amphibian declines and extinctions worldwide. Science 306:1783-1786.

Sun, J.W.C., and P.A. Narins. 2005. Anthropogenic sounds differentially affect amphibian call rate. Biological Conservation 121:419-427.

Vargas-Salinas, F., and A. Amezquita. 2013. Traffic noise correlates with calling time but not spatial distribution in the threatened poison frog andinobates bombetes. Behaviour 150:569-584.

Vargas-Salinas, F., G.M. Cunnington, A. Amezquita, and L. Fahrig. 2014. Does traffic noise alter calling time in frogs and toads? A case study of anurans in eastern ontario, canada. Urban Ecosystems 17:945-953.

Villanueva-Rivera, L.J. 2014. Eleutherodactylus frogs show frequency but no temporal partitioning: Implications for the acoustic niche hypothesis. Peerj 2.

Wasserstein, R.L., A.L. Schirm, and N.A. Lazar. 2019. Moving to a world beyond "p< 0.05". Taylor & Francis.

Wells, K.D. 1976. Multiple egg clutches in the green frog (rana clamitans). Herpetologica:85-87.

Wells, K. D., and C.R. Bevier. 1997. Contrasting patterns of energy substrate use in two species of frogs that breed in cold weather. Herpetologica, 70-80.

Wong, S., H. Parada, and P.M. Narins. 2009. Heterospecific acoustic interference: Effects on calling in the frog oophaga pumilio in nicaragua. Biotropica 41:74-80.

Yi, Y.Z., and J.A. Sheridan. 2019. Effects of traffic noise on vocalisations of the rhacophorid tree frog kurixalus chaseni (anura: Rhacophoridae) in borneo. Raffles Bulletin of Zoology 67:77-82.

Zeileis A., C. Kleiber, and S. Jackman. 2008. Regression models for count data in R. Journal of Statistical Software 27:8. http://www.jstatsoft.org/v27/i08/.

Zeileis A., and C. Kleiber. 2018. countreg: Count data regression. R package version 0.2-1. http://R-Forge.R-project.org/projects/countreg/.

Zhu, B.C., J.C. Wang, Z.X. Sun, Y. Yang, T.L. Wang, S.E. Brauth, Y.Z. Tang, and J.G. Cui. 2017. Competitive pressures affect sexual signal complexity in kurixalus odontotarsus: Insights into the evolution of compound calls. Biology Open 6:1913-1918.

2.7 TABLES

Table 1. Publications on the effects of anthropogenic/white noise on male anuran calling behavior. Locations include field (F), mesocosm (M), or lab (L). Types include experimental (E) or observational (O). Blue arrows, red arrows and zeros represent increases, decreases, and no change respectively for each call attribute.

Publication	Location	Type	Noise	No. Spp	Spp. in Common	Call Rate ↑	↓	Ø	Dominant Frequency ↑	↓	Ø	Call Length ↑	↓	Ø	Pulse Rate ↑	↓	Ø	Power ↑	↓	Ø
Bielinski Ch 2	F	E	Cars, airplanes, & trains	9	NA	0	8	1	1	0	8	1	2	6	1	1	3	-	-	-
Bielinski Ch 3	F	E	Cars	2	NA	0	1	1	0	1	1	0	2	0	-	-	-	1	0	1
Bleach et al. 2015	F	E	Lawn mower	2	-	1	0	1	0	0	2	-	-	-	-	-	-	-	-	-
Engbrecht et al. 2015	F	E	Airplanes	1	-	1	0	0	-	-	-	-	-	-	-	-	-	-	-	-
Hanna 2014	F	E	White noise ‡§	1	1*	0	0	1	0	1‡	0	0	1§	0	-	-	-	-	-	-
Kaiser & Hammers 2009	F	E	Motorcycle	1	-	1	0	0	-	-	-	-	-	-	-	-	-	-	-	-
Sun & Narins	F	E	Motorcycle	4	-	1	3	0	-	-	-	-	-	-	-	-	-	-	-	-
Brooke et al. 2000	F	E	White noise	1	-	0	0	1	-	-	-	-	-	-	-	-	-	-	-	-
Yi & Sheridan 2019	F	E	Cars	1	-	0	0	1	0	0	1	-	-	-	-	-	-	1	0	0
Kruger & Du Preez 2016	F	E	Airplanes	1	-	1	0	0	1	0	0	-	-	-	-	-	-	-	-	-
Forti et al. 2015	F	E	White noise	2	-	0	2	1	-	-	-	-	-	-	-	-	-	-	-	-
Lengagne 2008	M	E	Cars	1	-	0	1	0	0	0	1	0	0	1	-	-	-	-	-	-
Halfwerk et al. 2016	L/M	E	White noise‡	1	-	1	0	0	-	-	-	-	-	-	-	-	-	1	0	0
Cunnington & Fahrig 2010	F	E	Cars	4	3†	0	3	1	2	0	2	-	-	-	-	-	-	0	2	2
Cunnington & Fahrig 2010	F	O	Cars	4	3†	3	0	1	1	0	3	-	-	-	-	-	-	0	1	3
Hoskin & Goosem 2010	F	O	Cars	1	-	1	0	1	0	0	1	0	0	1	0	0	1	-	-	-
Grenat et al. 2019	F	O	Cars	1	-	-	-	-	1	0	0	0	0	1	1	0	0	-	-	-
Vargas-Salinas & Amezquita 2013	F	O	Cars	1	-	0	1	0	-	-	-	-	-	-	-	-	-	-	-	-
Parris et al. 2009	F	O	Cars	2	-	-	-	-	2	0	0	-	-	-	-	-	-	-	-	-

*Spring Peeper; †Green frog, eastern gray treefrog, and American toad; ‡ high frequency; §low frequency

Table 2. Species grouping by natural history. Species were grouped by different categories of interest for the multivariate analysis. For breeding season, species were grouped by early (E), middle (M), and late (L). For breeding season length, species were grouped by short (S), intermediate (I) or long (L). For habitat preference, species were grouped by temporary hydrology (T), temporary and permanent hydrology (B), or permanent hydrology (P).

Common name	Genus	Season	S.Length	Habitat
Bullfrog	*L.*	L	L	P
Green frog	*L.*	L	L	P
Wood Frog	*L.*	E	S	T
Copes gray t.f.	*H.*	L	I	B
Eastern gray t.f.	*H.*	L	I	B
Spring peeper	*P.*	E	I	T
Western chorus	*P.*	E	L	T
Cricket frog	*Ac.*	L	L	P
American toad	*An.*	M	I	B

Table 3. Effect sizes and model statistics of call attributes for all noise treatments pooled (car, train, and airplane noise) *vs.* control treatments. The split components of the hurdle models include the zero model (HZ) and the count model (HC). Other models include linear models (LM) and linear mixed-effect models (LMEM). HZ effect sizes are the percent change between a 1:1 ratio of zeros and the control : treatment ratio of zeros. Dominant frequency effect sizes are listed as the change in hertz from the control. The remaining effect sizes are listed as percent change from the control. Only results with P < 0.10 are listed. Effect sizes from all comparisons are located in the Appendices.

Pooled Noise *vs.* Control

Response Variable	Species	Model	Effect Size	F-Statistic or Hurdle Model df	P-Value
Call Rate	American bullfrog	HZ	-38%	5	0.057
	Green frog	HZ	-65%	5	< 0.001
	Wood frog	HZ	-50%	5	0.008
	Wood frog	HC	-57%	5	0.024
	Western chorus frog	HC	-15%	5	0. 097
	Eastern gray treefrog	LMEM	-13%	$F_{1,5} = 6.7$	0.012
	Cope's gray treefrog	LM	-70%	$F_{1,10} = 24.9$	< 0.001
	Northern cricket frog	HC	-32%	5	0.008
	American toad	HZ	-73%	5	0.095
	American toad	HC	-32%	5	0.065
Dominant Frequency	Spring peeper	LM	+91 Hz	$F_{1,275} = 4.5$	0.035
Call Length	Western chorus frog	LMEM	-8%	$F_{1,552} = 8.2$	0.004
	Cope's gray treefrog	LM	+19%	$F_{1,23} = 8.2$	0.009
	Northern cricket frog	LM	-18%	$F_{1,145} = 5.8$	0.018
Pulse Rate	Cope's gray treefrog	LMEM	+12%	$F_{1,22} = 6.9$	0.016
	American toad	LMEM	-7%	$F_{1,60} = 8.6$	0.005

Table 4. Effect sizes and model statistics of call attributes for the car (C) *vs.* airplane (P) *vs.* train (T) treatments. Comparisons were only made if effects were seen in the pooled noise *vs* control comparison or the Noise Power Experiment. The split components of the hurdle models include the zero model (HZ) and the count model (HC). Other models include linear models (LM) and linear mixed-effect models (LMEM). HZ effect sizes are the percent change between a 1:1 ratio of zeros and the control : treatment ratio of zeros. Dominant frequency effect sizes are listed as the change in hertz from the control. The remaining effect sizes are listed as percent change from the control. Only results with P < 0.10 are listed.

| | | | Car *vs.* Plane *vs.* Train | | | | |
Response Variable	Species	Model	P *vs* C Effect Size	P *vs* T Effect Size	C *vs* T Effect Size	F-Statistic or Hurdle Model df	P-Value(s)
Call Rate	American bullfrog	HZ	-	-63%	-	7	0.086
	American bullfrog	HC	-64%	-	-	7	0.035
	American toad	LM	+2.5x	+5.1x	-	$F_{2,15} = 6.4$	0.033, 0.003
Dominant Frequency	Spring peeper	LM	-128 Hz	-	-	$F_{2,178} = 2.1$	0.123
Call Length	Western chorus frog	LMEM	+9.5%	+9.4%	-	$F_{2,304} = 8.6$	< 0.001
	Cope's gray treefrog	LM	-19.70%	-	-	$F_{2,9} = 3.9$	0.024
	Northern cricket frog	LM	-24.50%	-	-	$F_{2,66} = 2.2$	0.068
	American toad	LM	-	-50%	-	$F_{2,17} = 8.5$	0.030
Pulse Rate	Cope's gray treefrog	LM	-8.25%	-	+14%	$F_{2,9} = 5.7$	0.011

Table 5. Effect sizes and model statistics of call attributes for the car noise treatments *vs* control treatments. Noise treatments include car noise at 70 dB (C70), 80 dB (C80), and 90 dB (C90). The split components of the hurdle models include the zero model (HZ) and the count model (HC). Other models include linear models (LM) and linear mixed-effect models (LMEM). HZ effect sizes are the percent change between a 1:1 ratio of zeros and the control : treatment ratio of zeros. Dominant frequency effect sizes are listed as the change in hertz from the control. The remaining effect sizes are listed as percent change from the control. Only results with P < 0.10 are listed. Effect sizes from all comparisons are located in the Appendices.

| Response Variable | Species | Model | Car at 70, 80 & 90 dB *vs.* Control | | | | |
			C70 vs Control Effect Size	C80 vs Control Effect Size	C90 vs Control Effect Size	F-Statistic or Hurdle Model df	P-Value(s)
Call Rate	American bullfrog	HC	-	-	-89%	9	0.090
	Wood frog	HC	-	-	-28%	9	0.006
	Western chorus frog	HC	-	-	-70%	9	0.011
	Cope's gray treefrog	LMEM	-	-64%	-91.40%	$F_{3,4} = 31.1$	0.005, 0.001
	American toad	HZ	-	-76%	-81%	8	0.093, 0.052
Dominant frequency	Western chorus frog	LM	-55 Hz	-	-	$F_{3,276} = 1.3$	0.095
	Spring peeper	LM	-	-	+207 Hz	$F_{3,99} = 5.3$	0.009
	Eastern gray treefrog	LMEM	-56 Hz	-	-	$F_{3,168} = 2.1$	0.021
	American toad	LM	-	-67 Hz	-	$F_{3,65} = 1.5$	0.046
Call Length	Green frog	LM	-13%	-	-	$F_{3,117} = 1.1$	0.092
	Wood frog	LM	-9.8%	-	-	$F_{3,158} = 2.9$	0.060
	Western chorus frog	LMEM	-	-13%	-15%	$F_{3,264} = 12.4$	< 0.001, < 0.001
	Cope's gray treefrog	LM	+41.8%	-	+71%	$F_{3,15} = 8.1$	0.002, 0.04
	Northern cricket frog	LM	-	-	-36%	$F_{3,110} = 4.0$	0.035
	American toad	LM	-	+30%	-	$F_{3,63} = 2.2$	0.017
Pulse Rate	Western chorus frog	LMEM	-	+7%	+5%	$F_{3,262} = 4.2$	< 0.001, 0.017

Cope's gray treefrog	LM	+52%	-	-	$F_{3,15} = 8.5$	< 0.001
Northern cricket frog	LM	-	+22%	22%	$F_{3,110} = 3.7$	0.006, 0.009
American toad	LMEM	-	-	-8%	$F_{3,59} = 2.3$	0.023

Table 5 (continued)
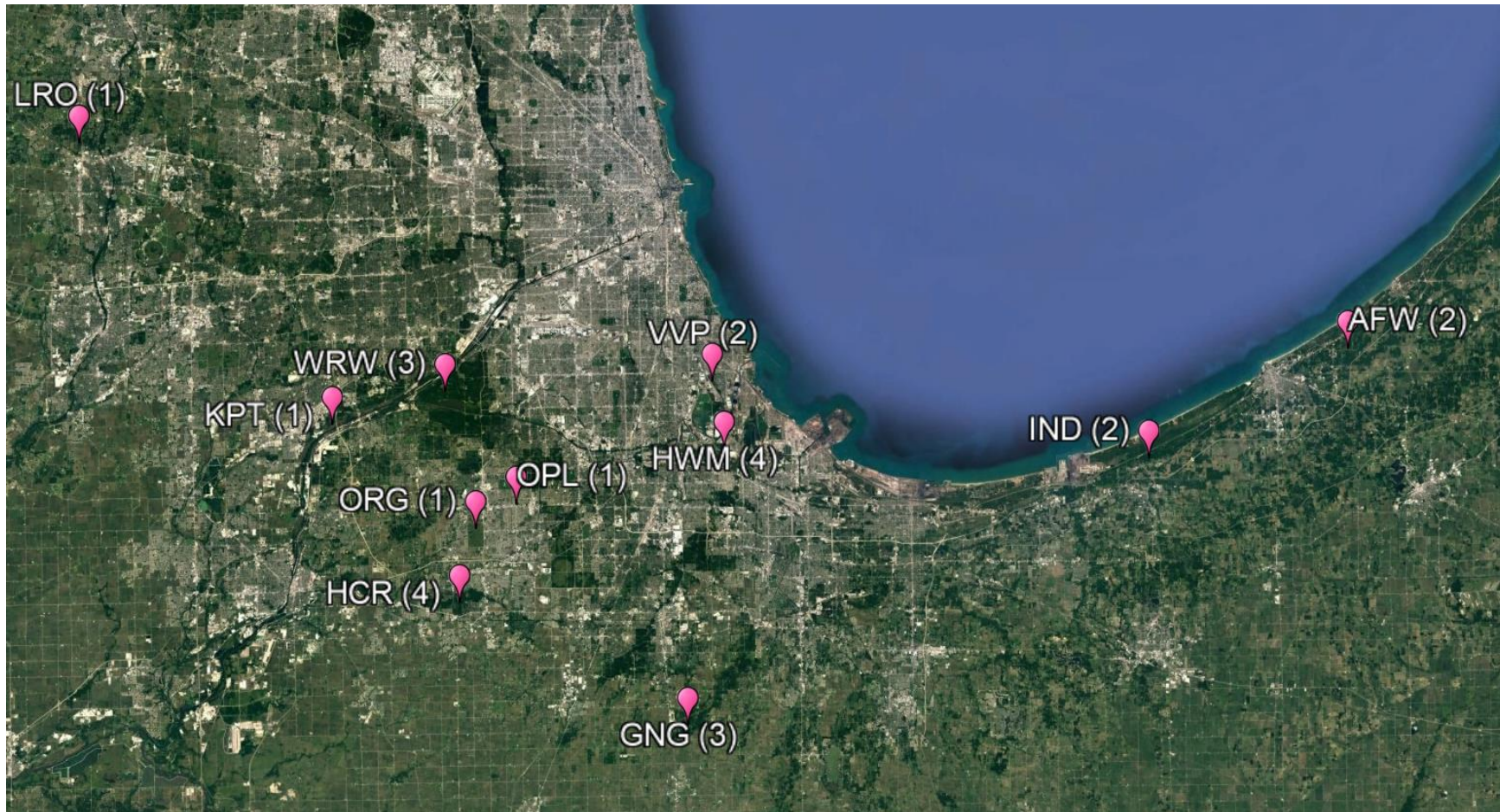
Figure 1. Study sites. Pink markers indicate different locations, and numbers represent the number of sites (breeding ponds) within each location where we conducted playback experiments. Locations include: LeRoy Oakes (LRO; Kane County Forest Preserve), Keepataw Preserve (KPT; Will County Forest Preserve), Wolf Road Woods (WRW; Cook County Forest Preserve), Hickory Creek Preserve (HCR; Will County Forest Preserve), Orland Grasslands (ORG; Cook County Forest Preserve), unnamed land (OPL; Cook County Forest Preserve); Goodenow Grove (GNG; Will County Forest Preserve), Van Vlissingen Prairie (Chicago Park), Hegewisch Marsh (Chicago Park), Indiana Dunes (IND; National Park), and Ambler Flatwoods (AFW; Shirley Heinze Land Trust Preserve).
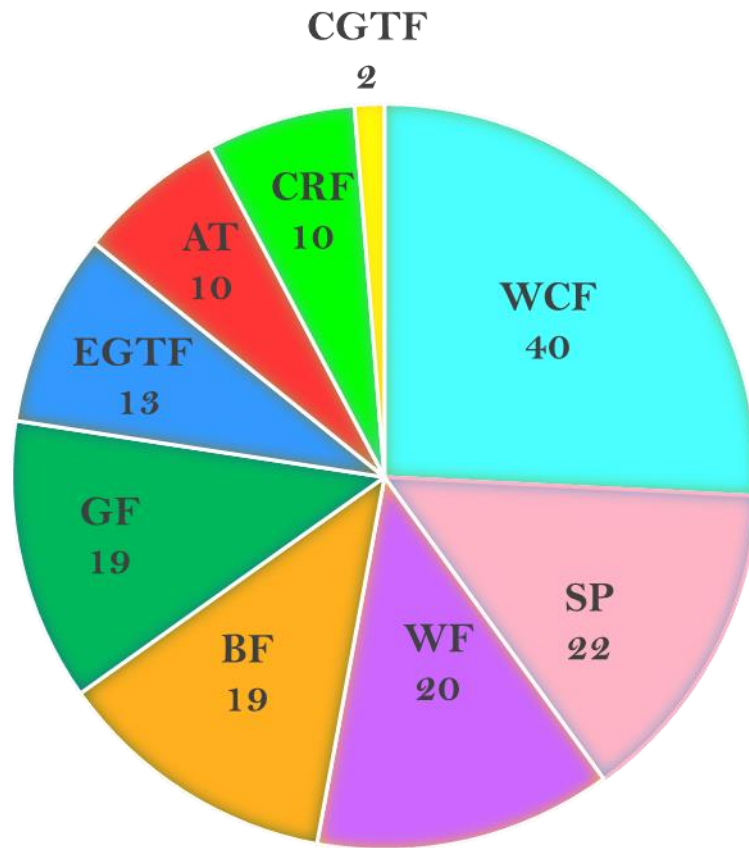
Figure 2. Number of playback experiments by species. Species include the Cope's gray tree frog (CGTF), the western chorus frog (WCF), the spring peeper (SP), the wood frog (WF), the American bullfrog (BF), the green frog (GF), the eastern gray tree frog (EGTF), the American toad (AT), and the northern cricket frog (CRF).
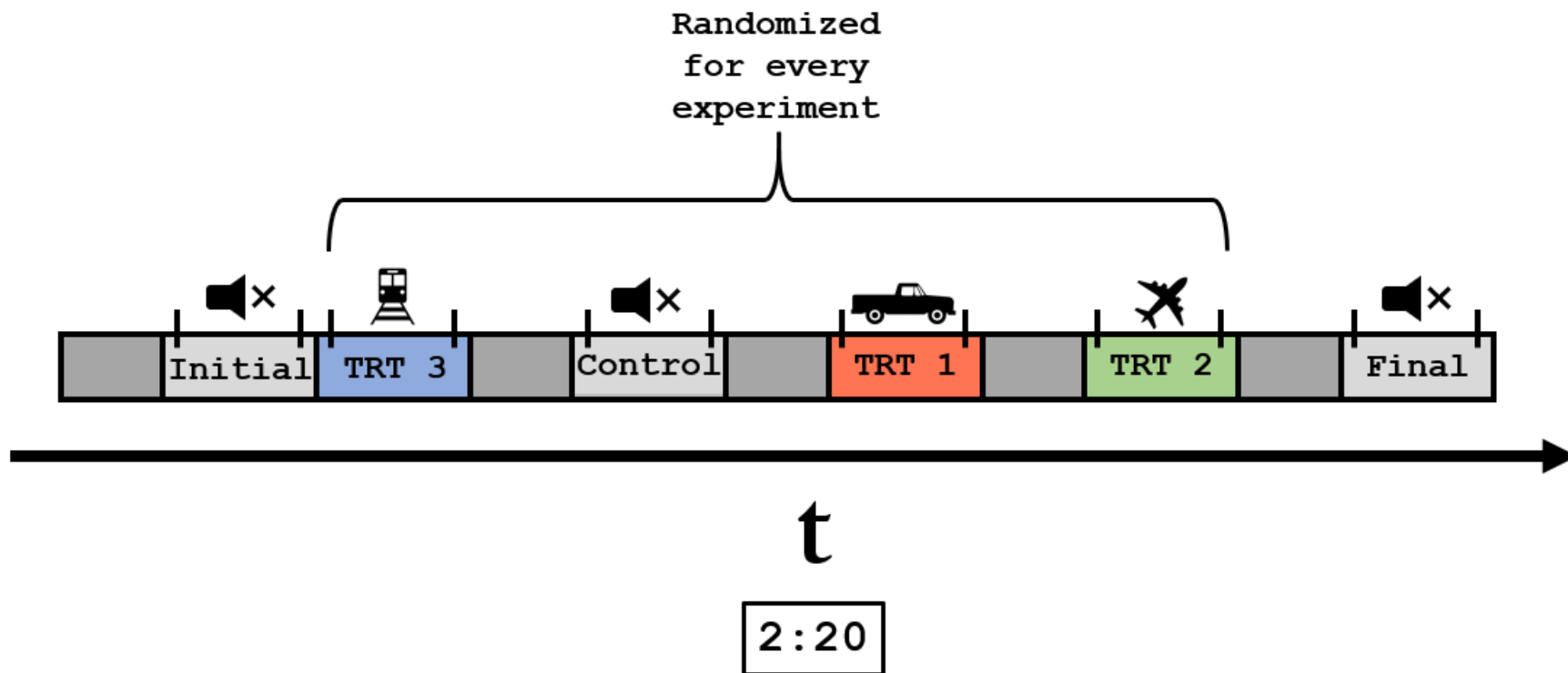
Figure 3. Playback experimental structure. At each pond, we used a crossover design of three noise treatments and one silent control treatment. Larger blocks represent 15-minute periods, smaller dark grey boxes represent 10-minute periods when no data was collected. Each experiment lasted 2 h 20 min.
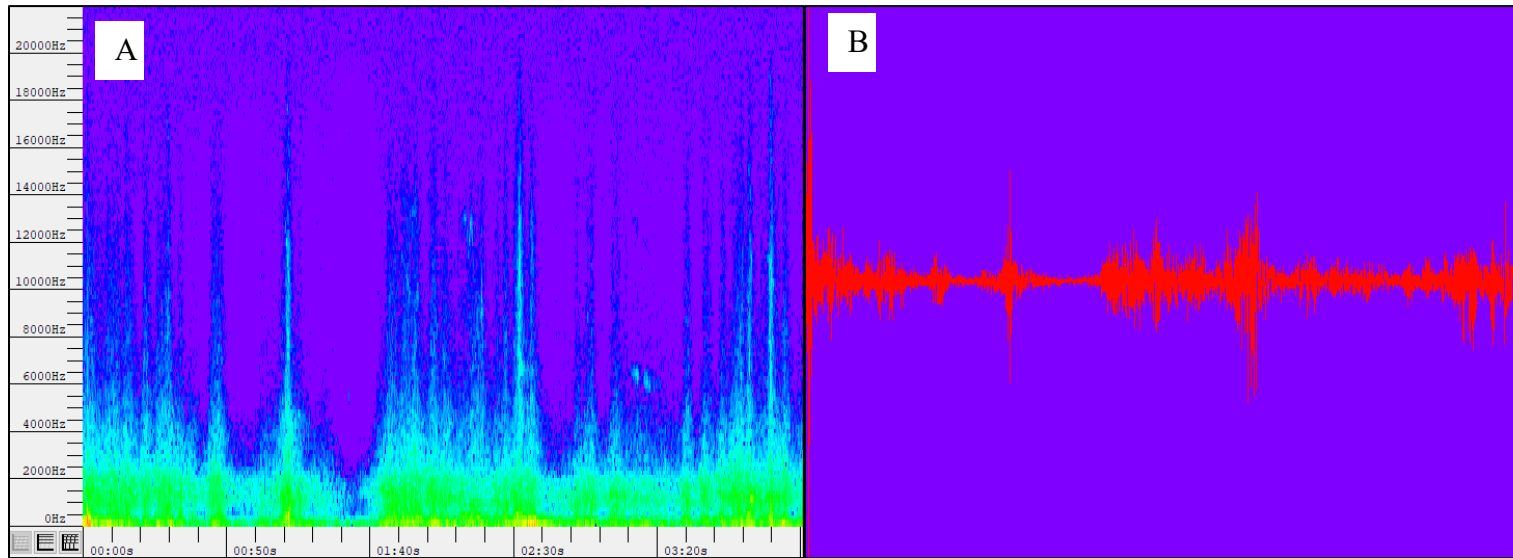
Figure 4. Car noise treatment. A) A spectrogram of a portion of the car noise treatment. The X-axis is time (mm:ss) and the Y-axis is frequency (Hz). Colors represent the amplitude (SPL in dB) of the noise, with red/yellow being the loudest, green being intermediate, and purple representing a lack of sound. Spikes in frequency and amplitude correspond to cars driving by the recorder and have a random pattern. B) A waveform view of the car noise treatment, showing amplitude as the Y-axis.
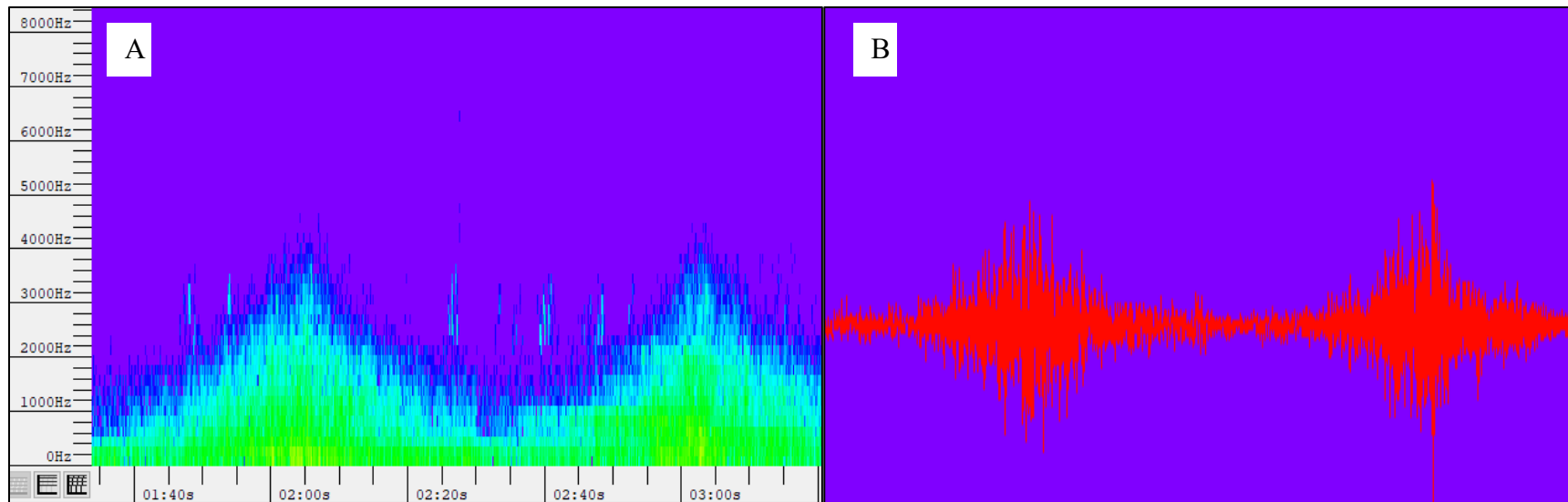
Figure 5. Airplane noise treatment. A) A spectrogram of a portion of the airplane noise treatment. The X-axis is time (mm:ss) and the Y-axis is frequency (Hz). Colors represent the amplitude (SPL in dB) of the noise, with red/yellow being the loudest, green being intermediate, and purple representing a lack of sound. Two large spikes can be seen representing two different planes flying overhead. B) A waveform view of the airplane noise treatment, showing amplitude as the Y-axis. Both amplitude and frequency show a predictable pattern with a quieter period in between flights.
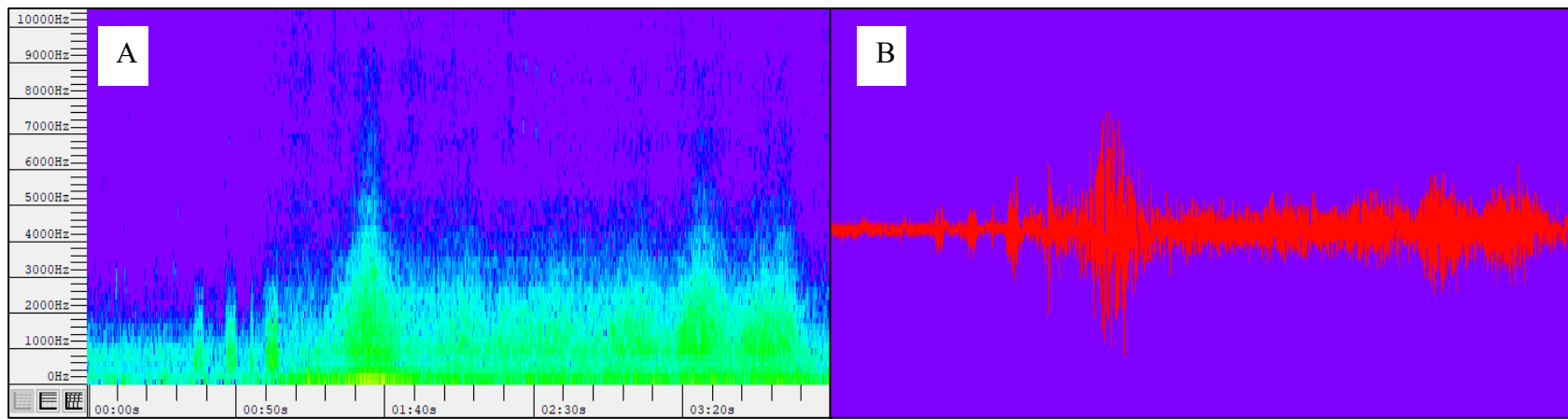
Figure 6. Train noise treatment. A) A spectrogram of a portion of the train noise treatment. The X-axis is time (mm:ss) and the Y-axis is frequency (Hz). Colors represent the amplitude (SPL in dB) of the noise, with red/yellow being the loudest, green being intermediate, and purple representing a lack of sound. Initial spikes are train horns, followed by a wall of noise of the train traveling by the recorded. B) A waveform view of the train noise treatment, showing amplitude as the Y-axis. Both amplitude and frequency show a random pattern as the train passes by, but in between trains the noise subsides.
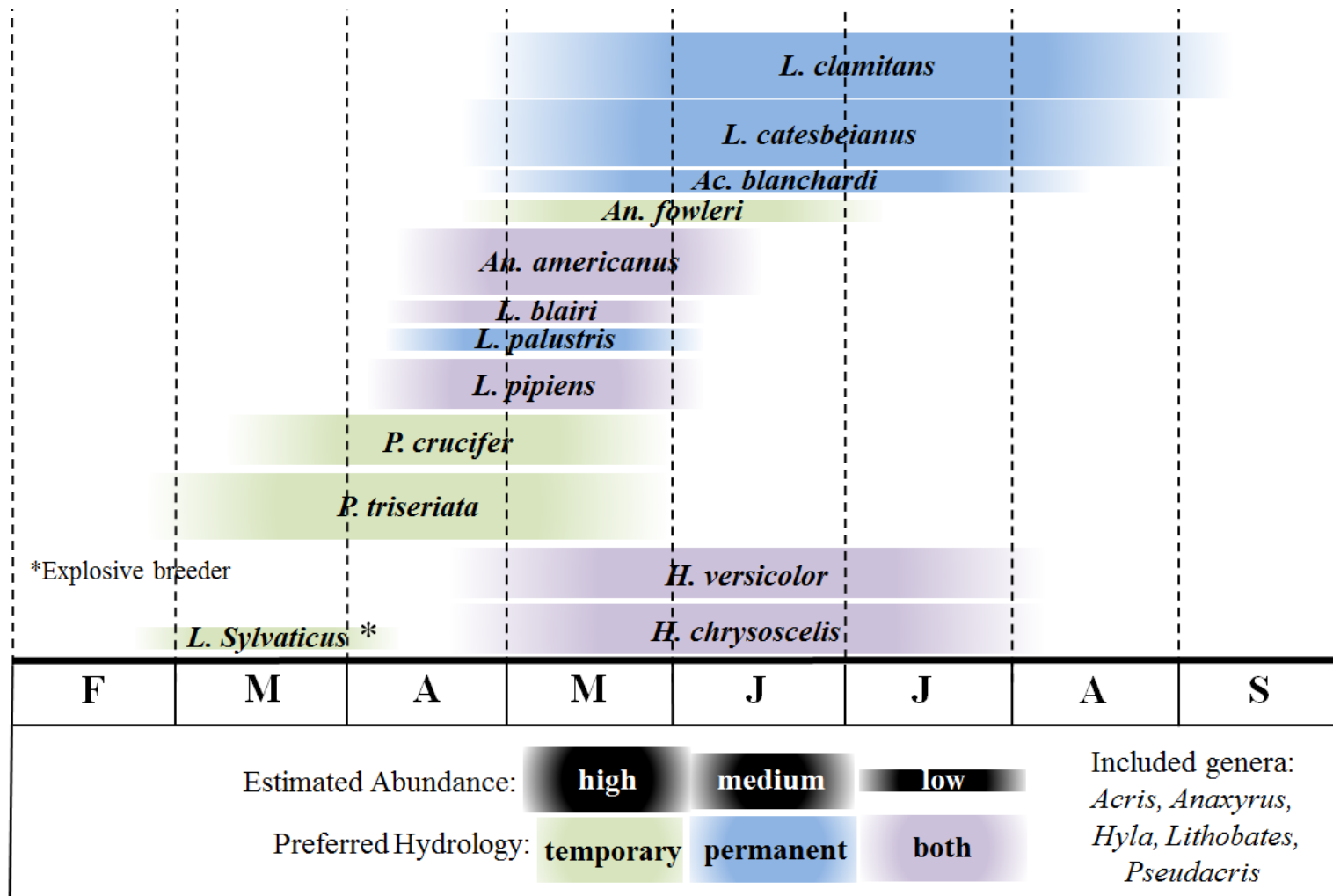
Figure 7. The calling phenology, abundance, and breeding pond hydrological preferences for the anuran community of the Chicago region. Phenologies and hydrology are from personal observation and Phillips *et al*. 1999. Abundance was estimated from the Calling Frog Survey 2014-2017 dataset (Peggy Notebaert Nature Museum).
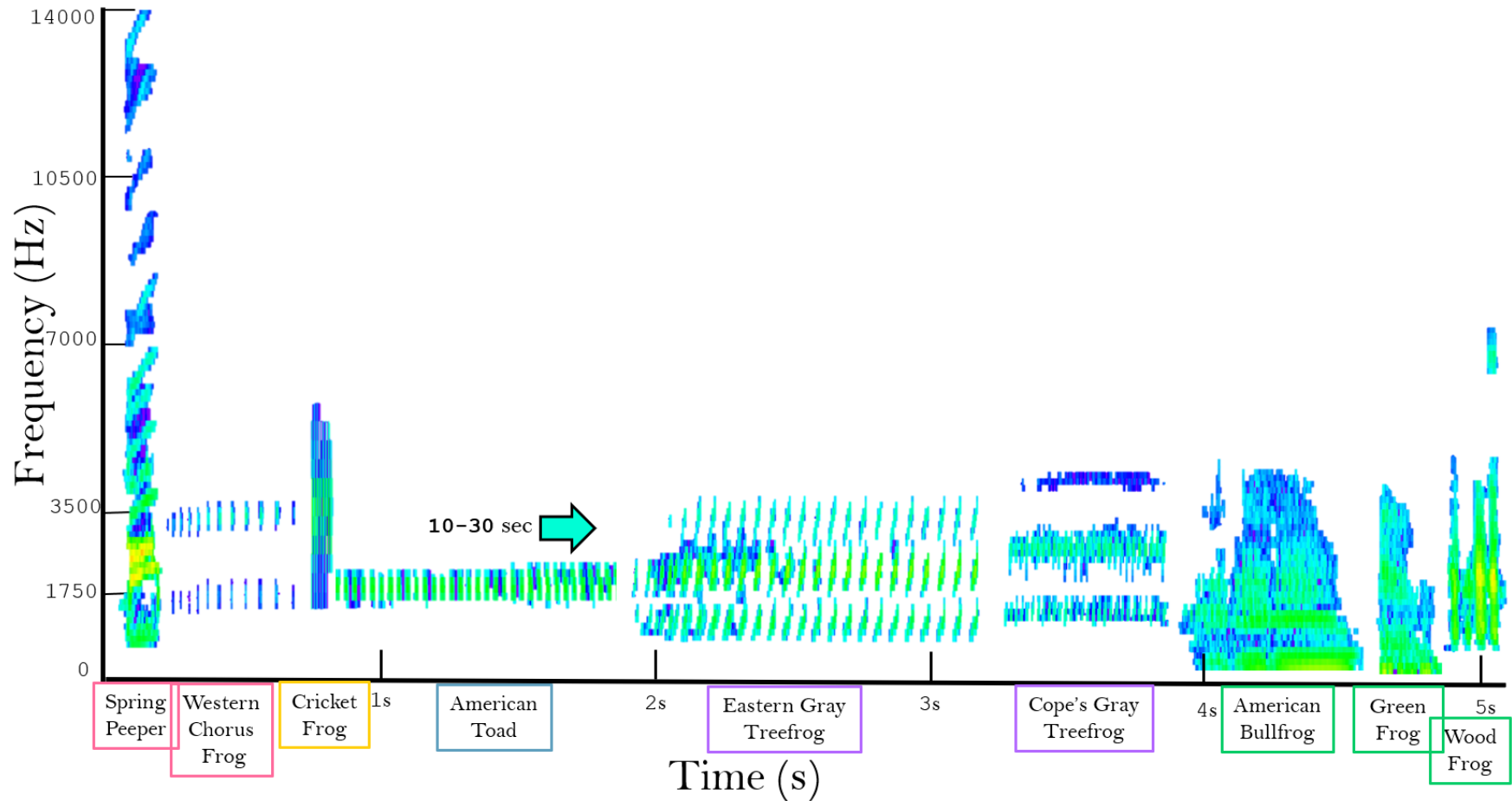
Figure 8. Spectrograms of calls from all target species. Calls were recorded during control treatments. Striations along the X-axis shows pulses in a call. Multiple lines along the Y-axis indicate formants (prominent bands of frequency resulting from resonance). Colored boxes represent membership to genera. The call for the American toad can last up 30 seconds or longer, so the green arrow is an indication that this displayed is only a small portion of the call.

Modeling Decision tree
Response variables per species, across treatments

Histogram → Unimodal?

Y / N

Transform sqrt, log

lm (# calls ~ trt)

lm (# calls ~ trt + offset (initial calls))

lmer (# calls ~ trt + (1|experiment))

glm (# calls ~ trt + offset (initial calls), poisson)

Hurdle(# calls ~ trt, negbin)

Hurdle(# calls ~ trt + offset (initial calls), negbin)

Hurdle(# calls ~ trt, poisson)

Hurdle(# calls ~ trt + offset (initial calls), poisson)

Choose Best Model
AIC, QQ, residuals
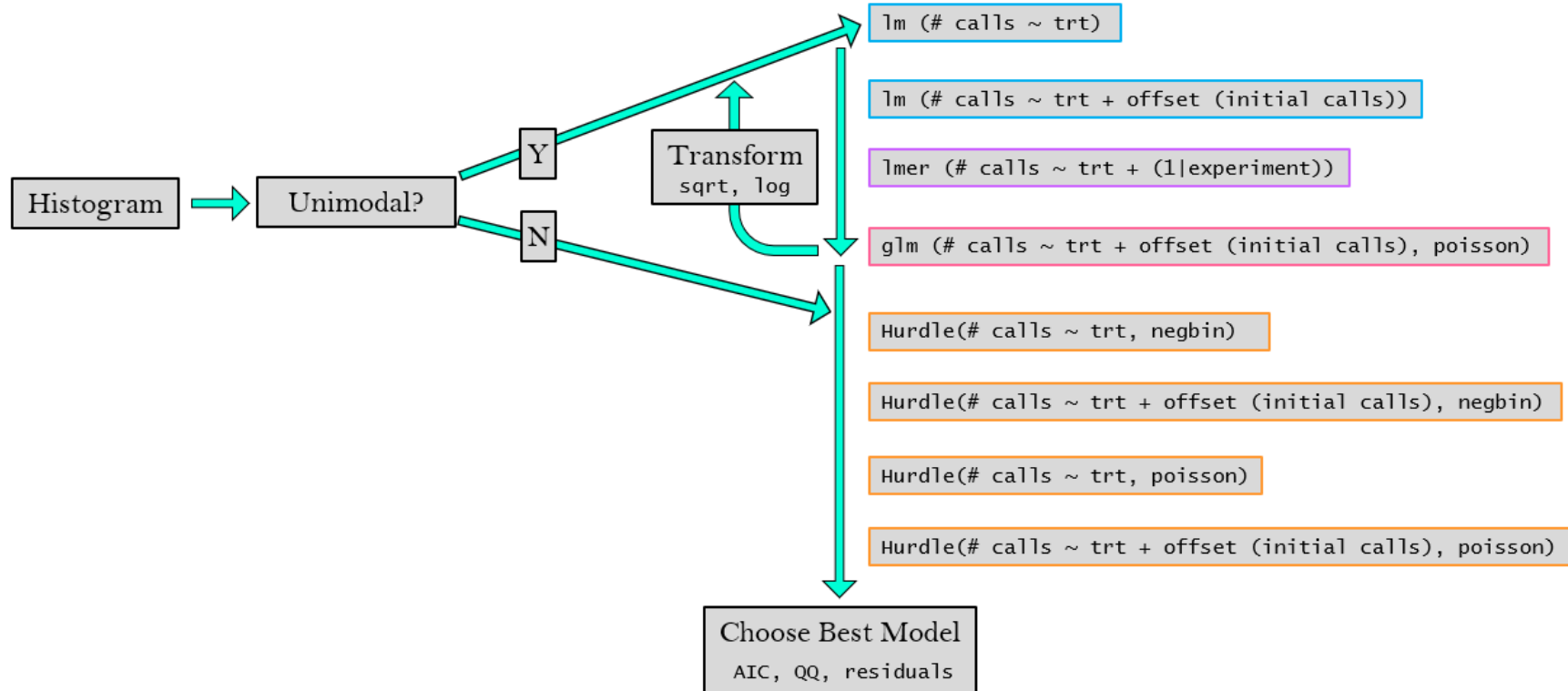
Figure 9. Fitting univariate response variables to a statistical model. This diagram shows the logical approach we used to find the correct model in R for each species × response variable combination. This was a necessary process as some distributions were normal, skewed, or non-normal due to zero inflation.
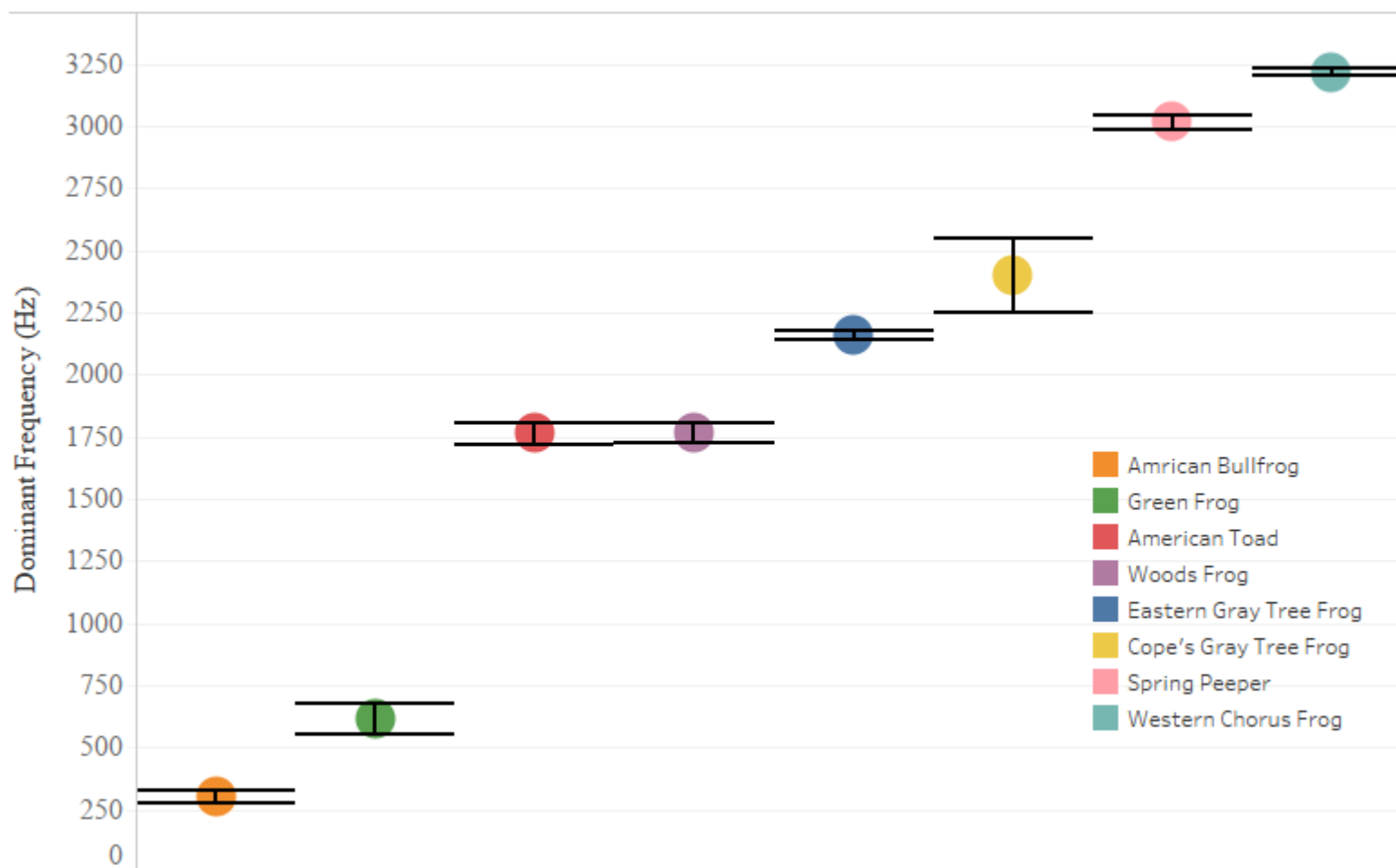
Figure 10. Dominant frequency of each species. Data is from dominant frequencies (from the loudest formant) collected across all treatments. There is a large variation in dominant frequency across the community, which highlights the value of this dataset.
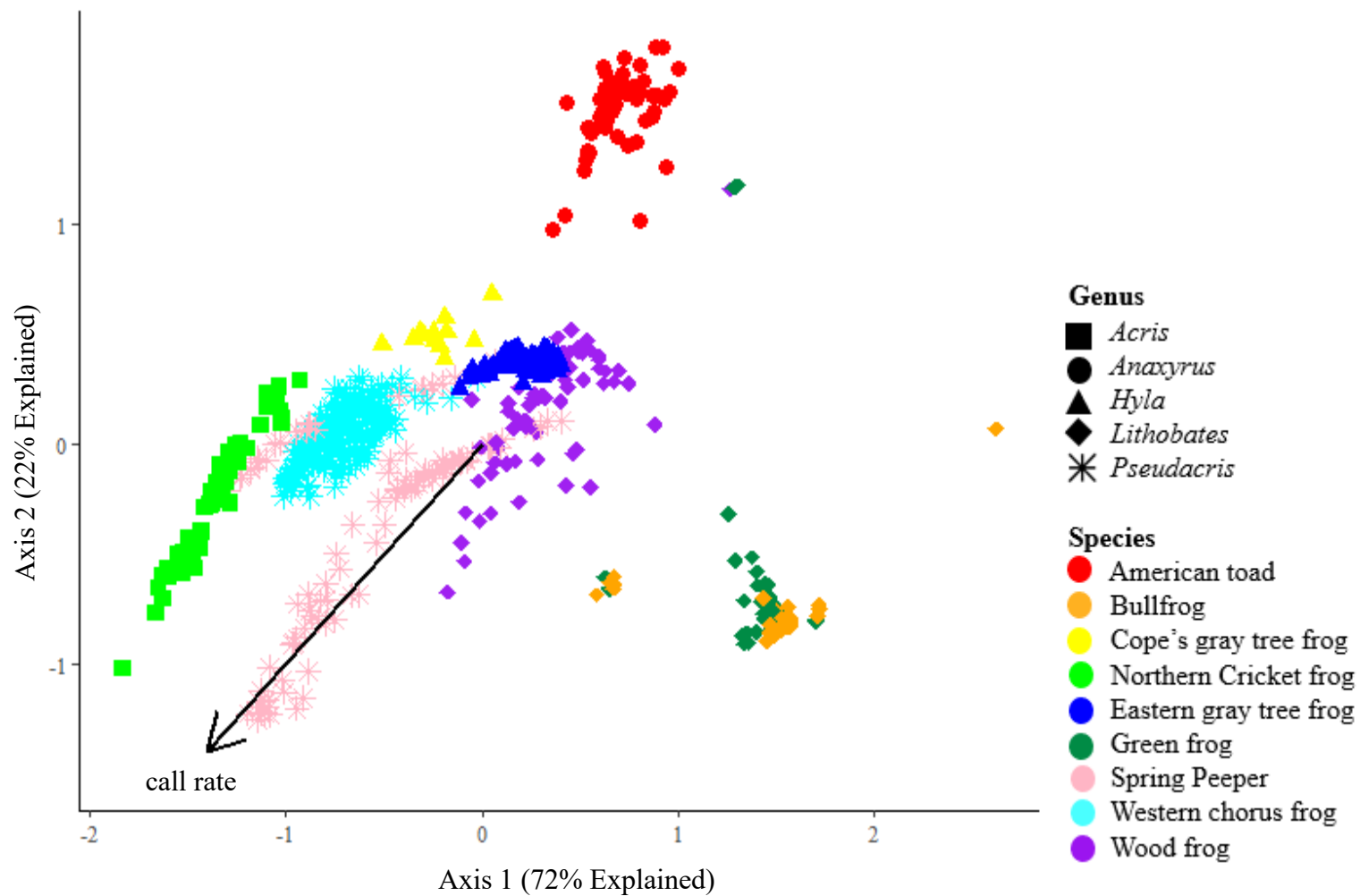
Figure 11. PCoA with grouping by genus and species. The call attributes in multivariate space clearly separate between species and genus. The only response variable that contributes substantially to this ordination is the call rate of each sample, represented by the arrow. The northern cricket frog and spring peeper seem to have the greatest variation in call rate.

99
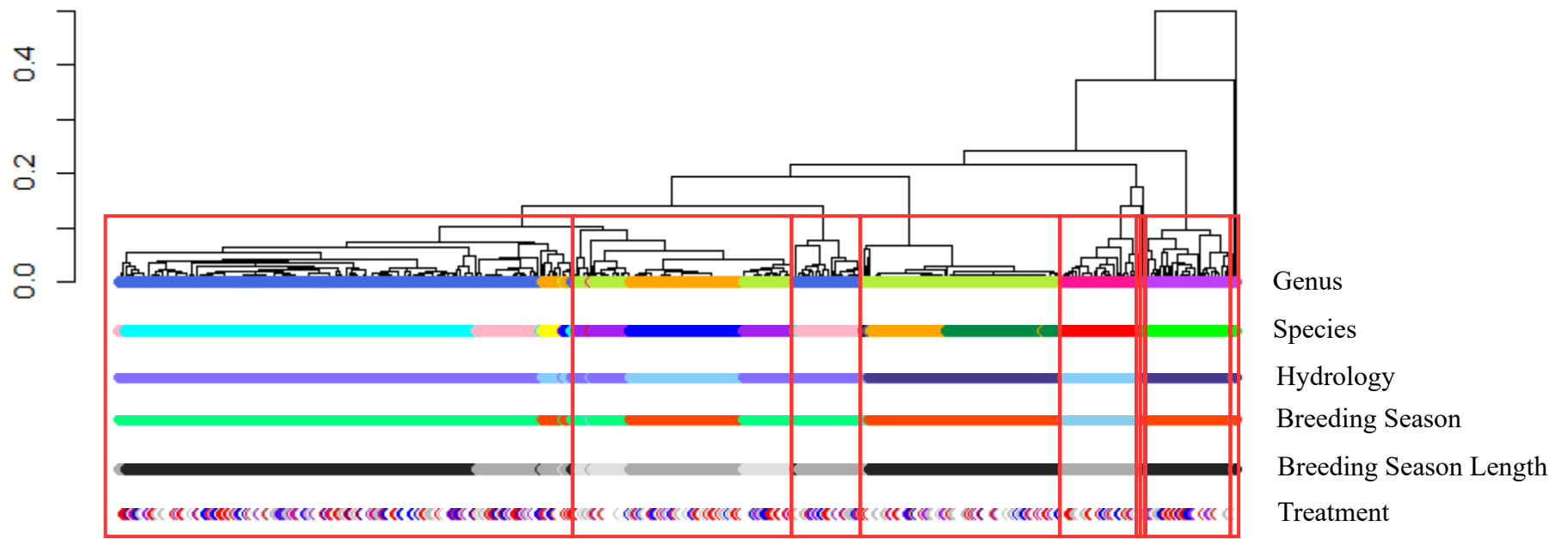
Figure 12. Cluster analysis. The UPGMA clustering method was used to create a dendrogram of the multivariate data of both species combined. Red boxes represent suggested number of groups (k =10), as indicated by the silhouette widths that measure the intensity of the link between objects and their groups. Color of the dendrogram leaves represent affiliation with the listed categories. There is clear separation by species, some clustering by genus, and no clustering by the other factors.
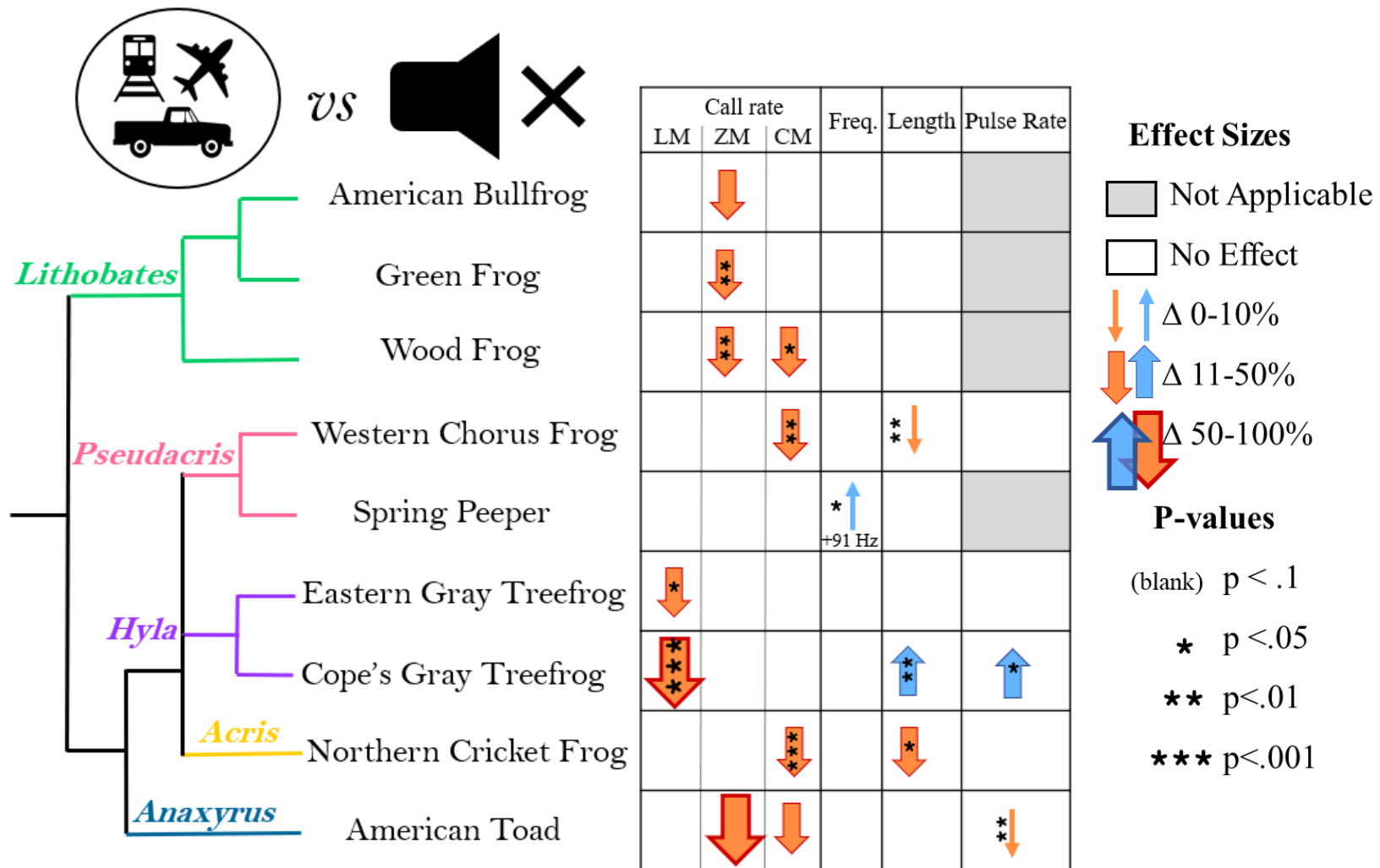
Figure 13. Summary of effects for pooled noise versus control. Results from all univariate analyses of species × response variable combinations are shown. Arrow size represents the strength of the effect size estimated from the model. Directions and color of the arrows represent the direction of the change in the attribute. Stars represent P-values. For call rate, when linear models were used, results were placed in the "LM" column. When hurdle models were used, results were placed in the "ZM" and "CM" columns (for the zero and count values of the hurdle model, respectively).
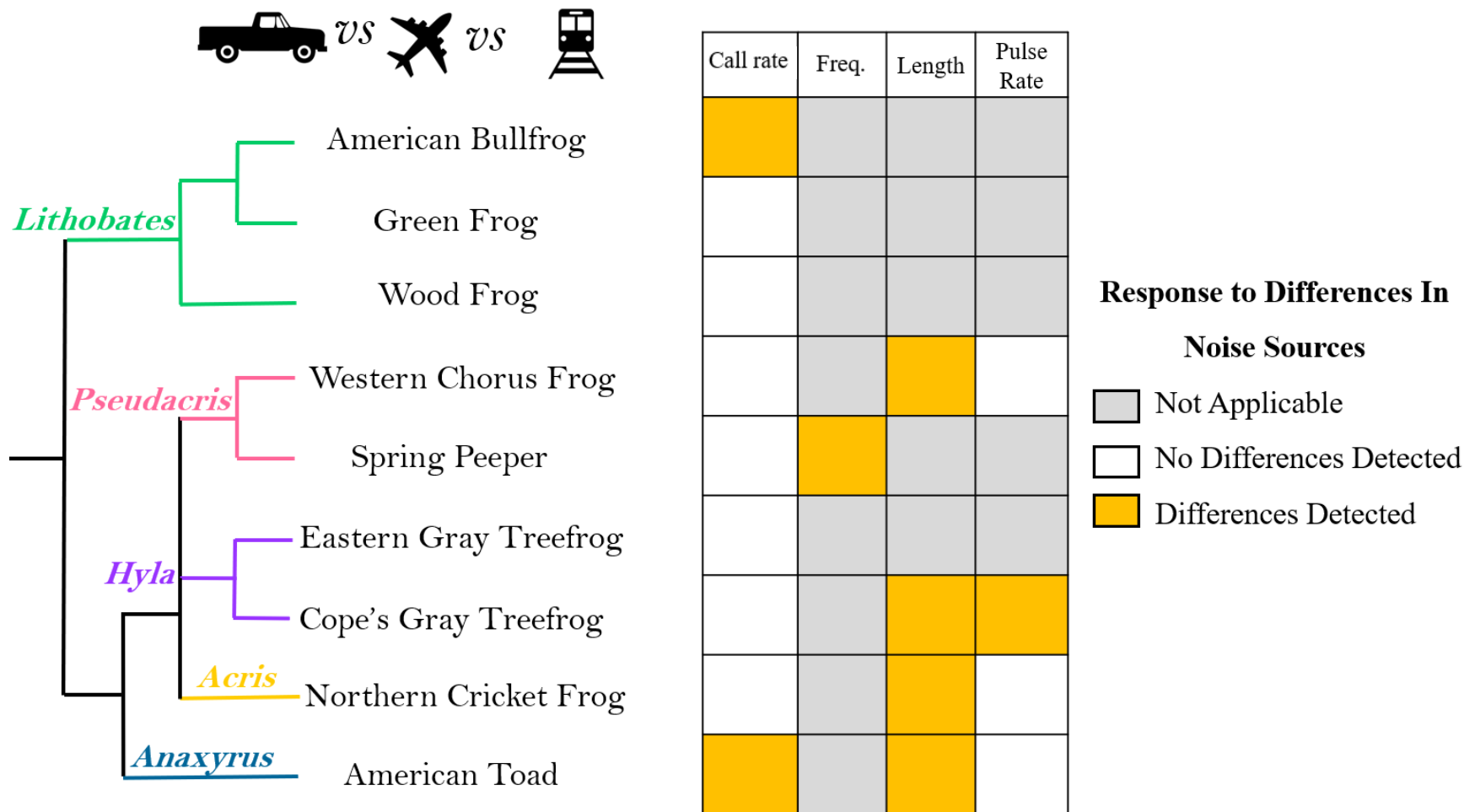
Figure 14. Differential response to car, airplane, and train noise treatments. Cells are highlighted if there were changes detected between any of the three combinations car *vs.* airplane, car *vs.* train, and airplane *vs.* train.
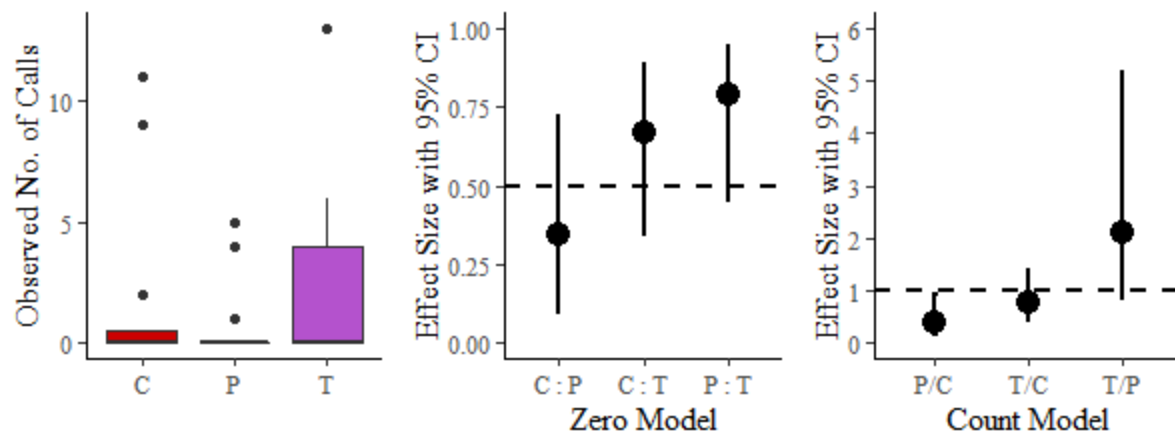
Figure 15. Impact of car (C), plane (P), and train (T) noise at 90 dB on call rate for the American bullfrog. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call number ~ treatment). For more detail of effect size estimates, refer to Appendix A.
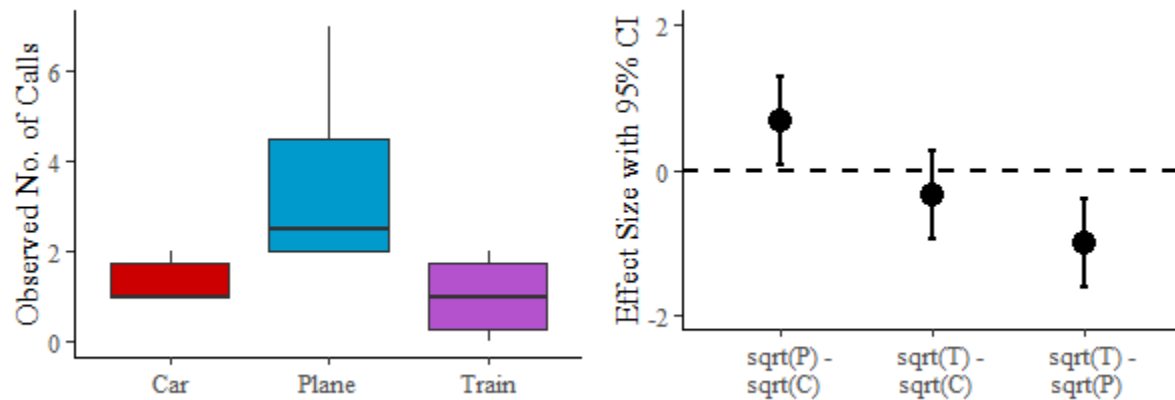
Figure 16. Impact of car (C), plane (P), and train (T) noise at 90 dB on call rate for the American toad. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (sqrt(call number) ~ treatment). For more detail of effect size estimates, refer to Appendix A.
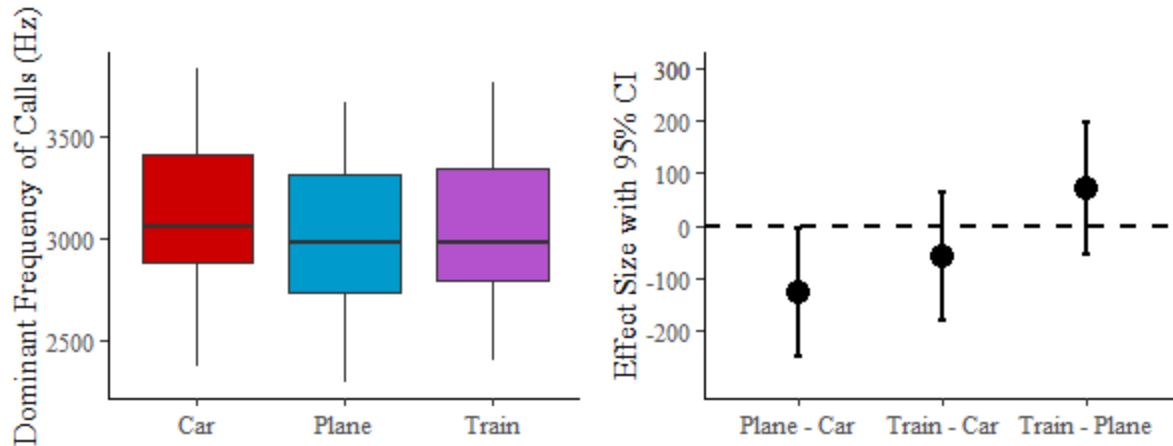
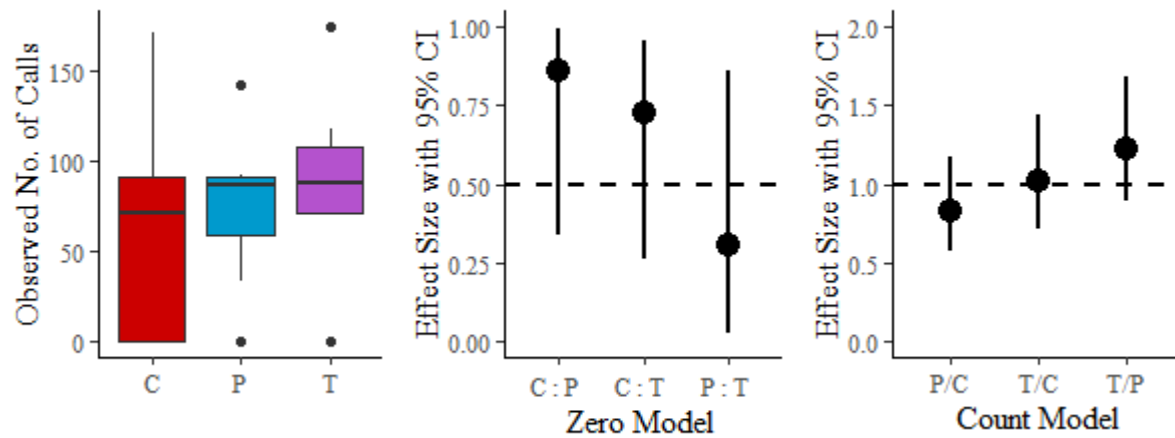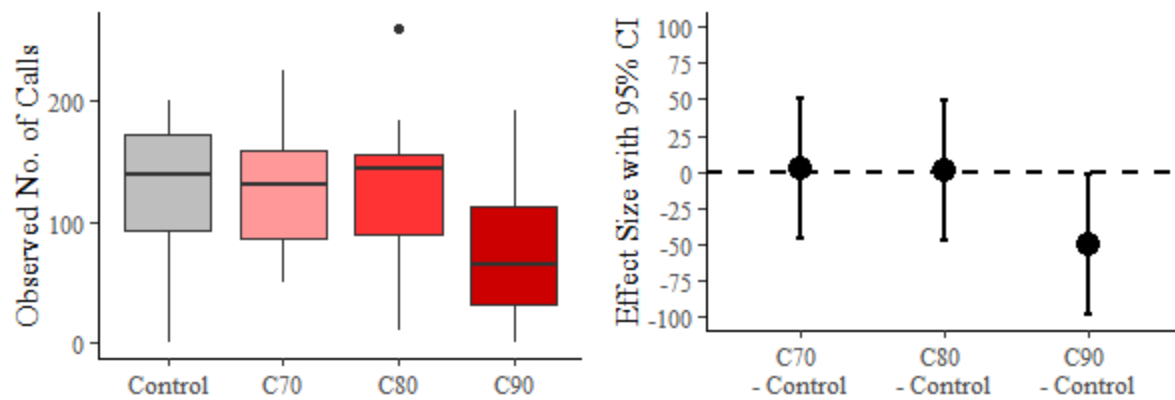Figure 17. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the spring peeper. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
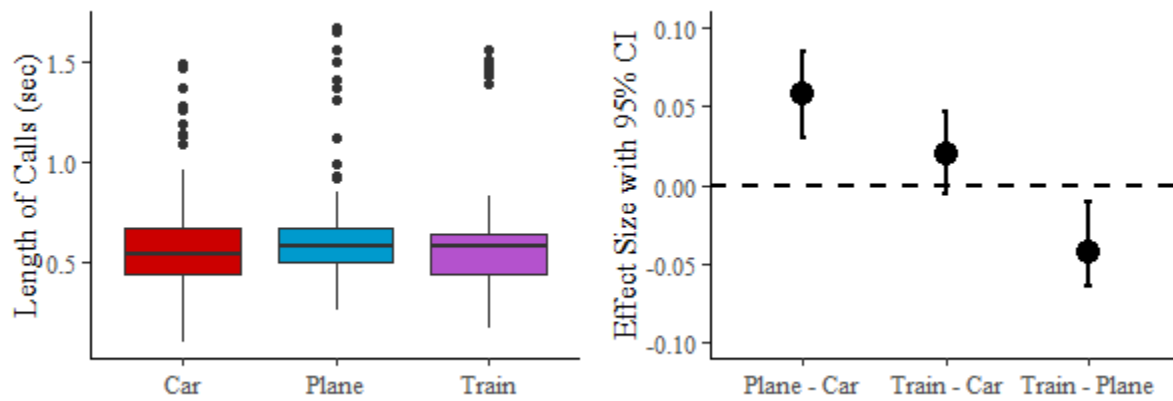
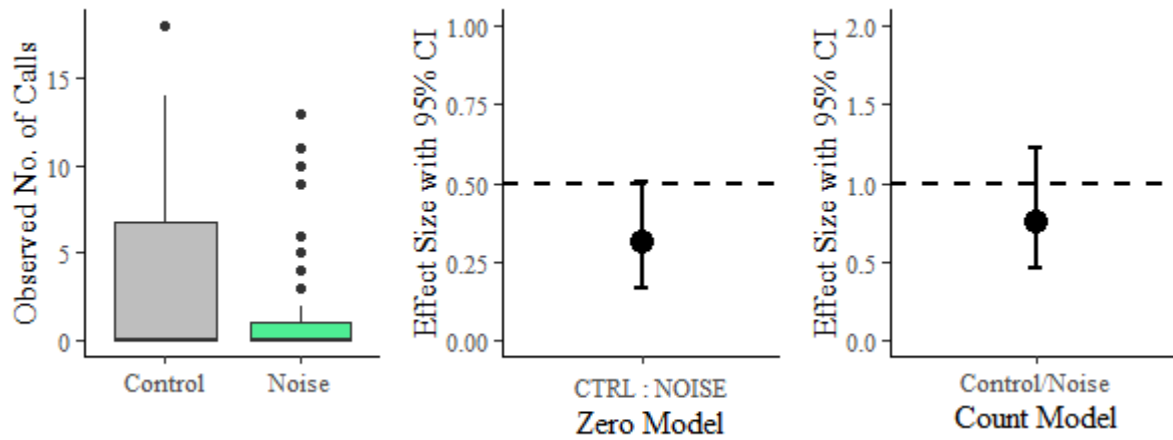Figure 18. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the western chorus frog. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model (call length ~ treatment + (1|block)). For more detail of effect size estimates, refer to Appendix A.
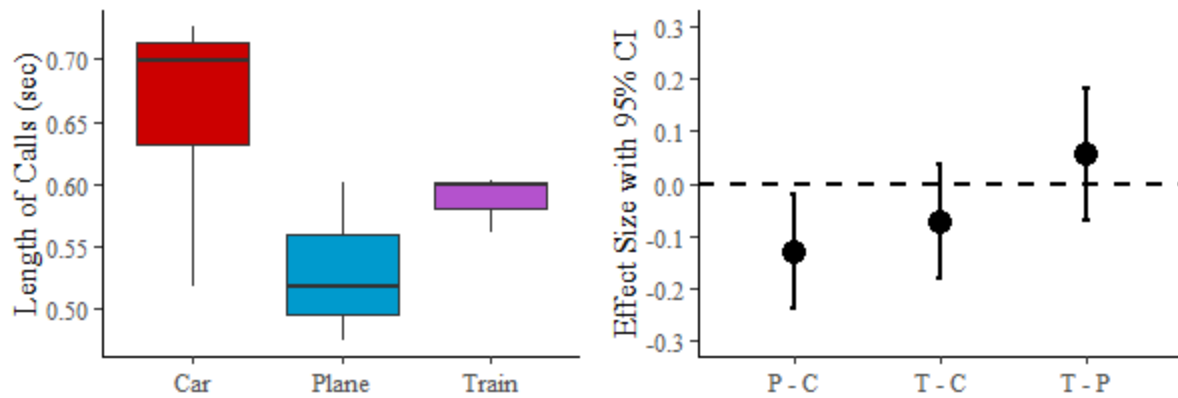
Figure 19. Impact of car (C), plane (P), and train (T) noise at 90 dB on call length for the Cope's gray treefrog. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
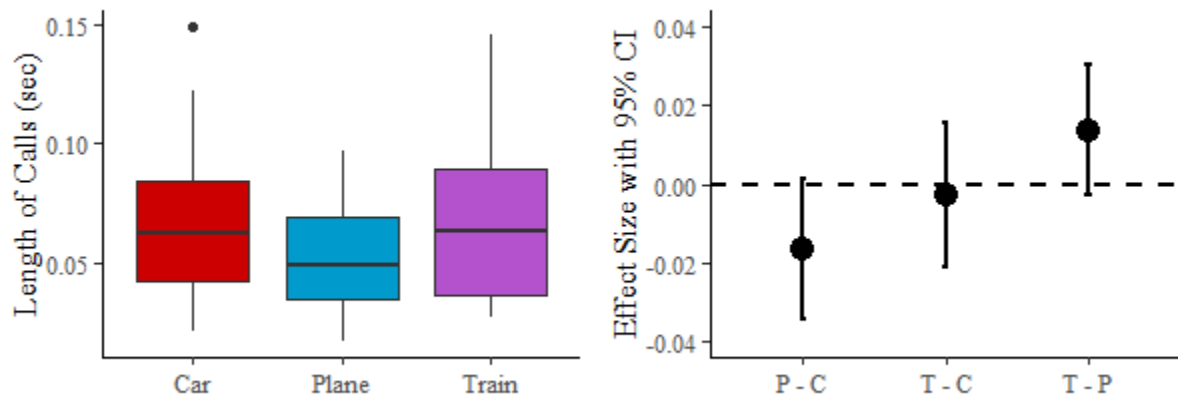
Figure 20. Impact of car (C), plane (P), and train (T) noise at 90 dB on call length for the northern cricket frog. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
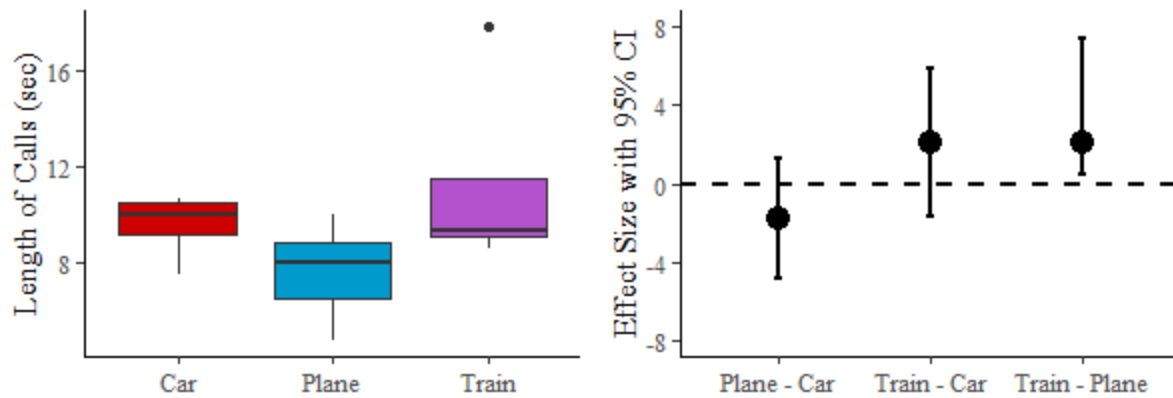
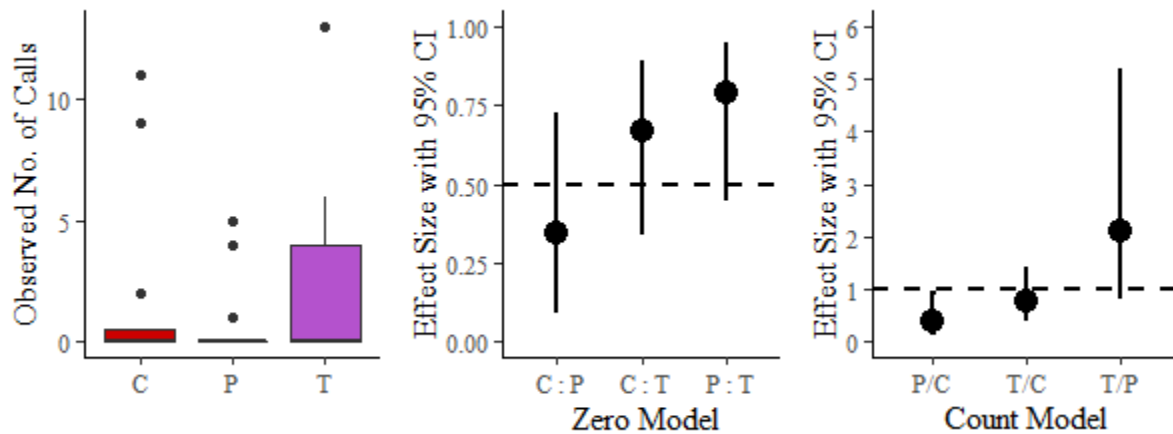Figure 21. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the American toad. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (call length ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (call length ~ treatment + offset(initial conditions). For more detail of effect size estimates, refer to Appendix A.
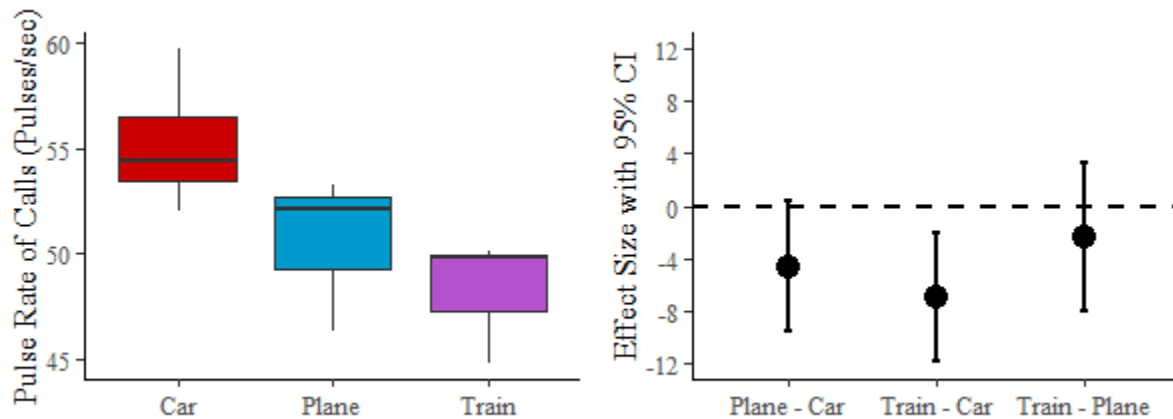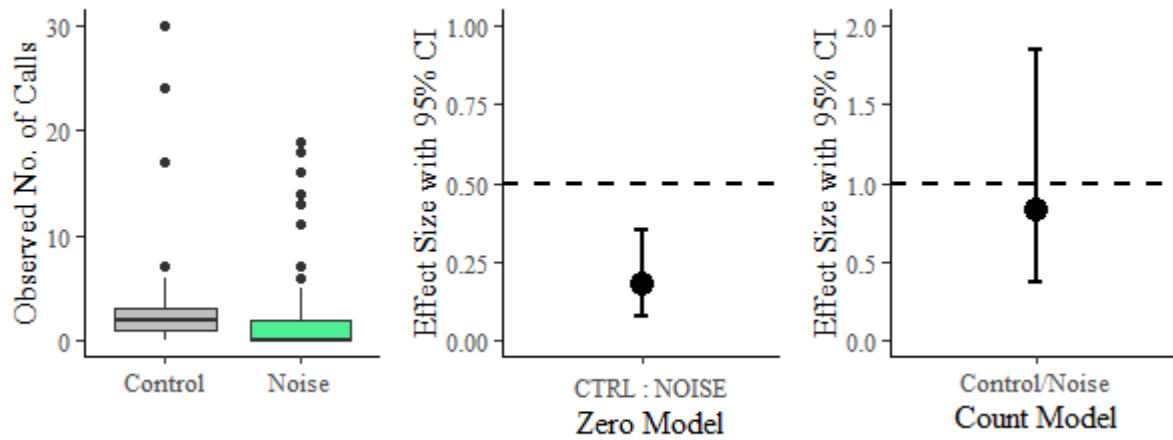

Figure 22. Impact of three types of noise (car, plane, and train) at 90 dB on pulse rate for the Cope's gray treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (pulse rate~ treatment + (1|block)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (pulse rate ~ treatment).For more detail of effect size estimates, refer to Appendix A.
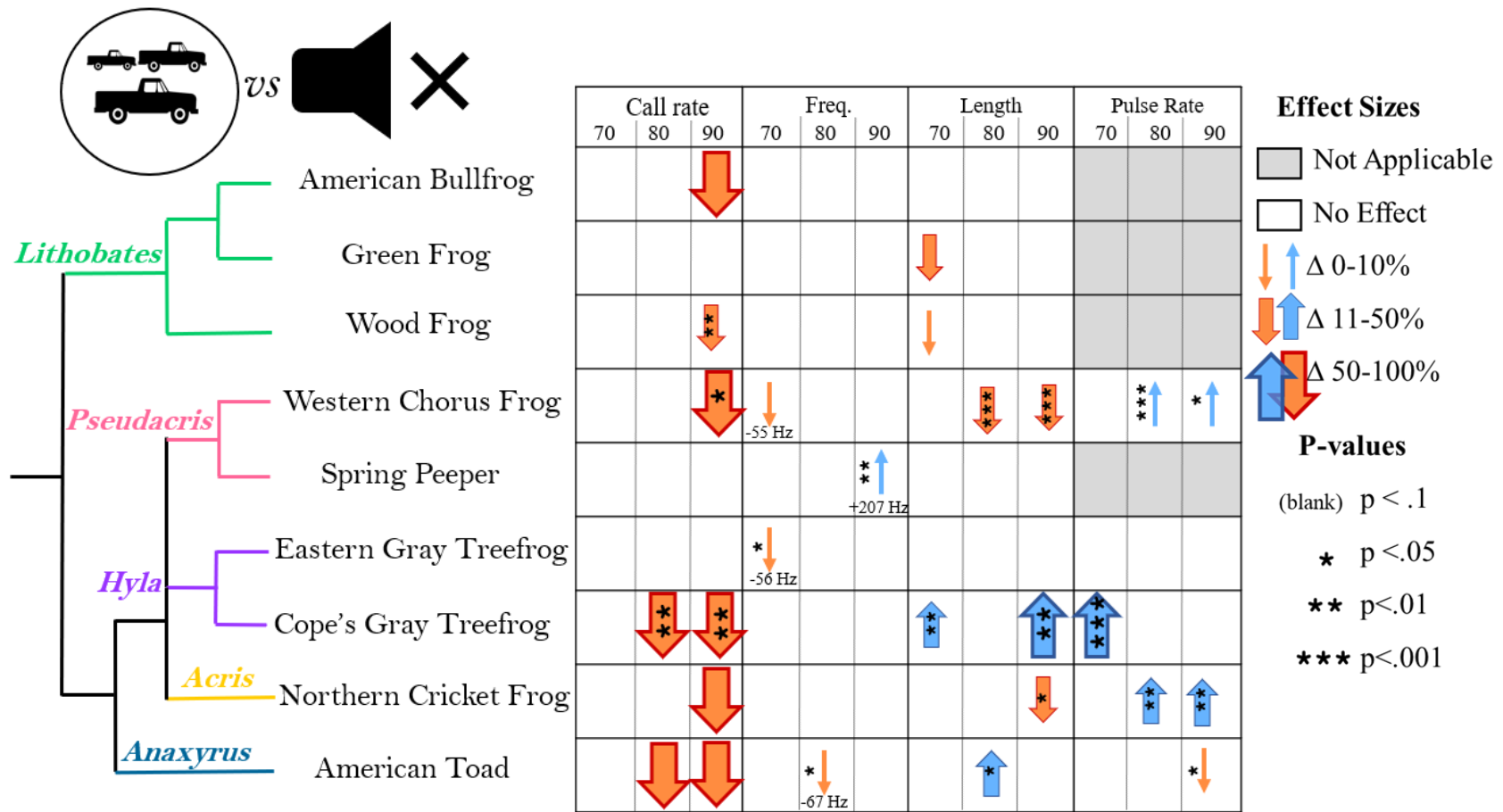
Figure 23. Summary of effects for car 70 dB, car 80 dB and car 90 dB *vs.* control. Results from all univariate analyses of species × response variable combinations are shown. Arrow size represents the strength of the effect size estimated from the model. Directions and color of the arrows represent the direction of the change in the response variable. Stars represent P-values.

2.9 Appendices

Appendix A. Explanation of effect size comparisons for all model types.

How effect size was calculated depended on the model used for each response variable. In linear models and linear mixed-effect models, effect sizes were simply the difference between the control estimate and the treatment estimate. If these were square root transformed, then the square of the estimates was taken. For all these examples, a difference of zero would indicate no change between control and treatment. If the linear models were log transformed, then the exponent of the estimates was taken. This transformation creates a fraction of (treatment / control). Therefore, an effect size of 1 would indicate no change between treatments. For each hurdle model (which was a generalized linear model), each half of the model required separate transformations. The zero model's link function is the logit. Thus, we took the inverse logit of the model estimations to come up with the effect size. This transformation creates a ratio of

(# zeros in the data belonging to control) : (# of zeros in the data belonging to the treatment) scaled to 1.0 as the total. For these models, an effect size of 0.5 represents no change. The count model's link function is the log. We took the exponent of the estimates. This transformation creates a fraction of treatment/control. An effect size of 1 would indicate no change between treatments.

For ease of comparison, for all effect size plots, we set the "no difference" value at exactly the halfway point on the Y-axis that is represented by the dashed line. The value of this dashed line is either 0, 0.5, or 1, depending on the model. On rare occasions, this dashed line was not placed in the center because of remarkably high or low confidence intervals.

The confidence intervals were calculated from each model type following the same procedure as the effect sizes.
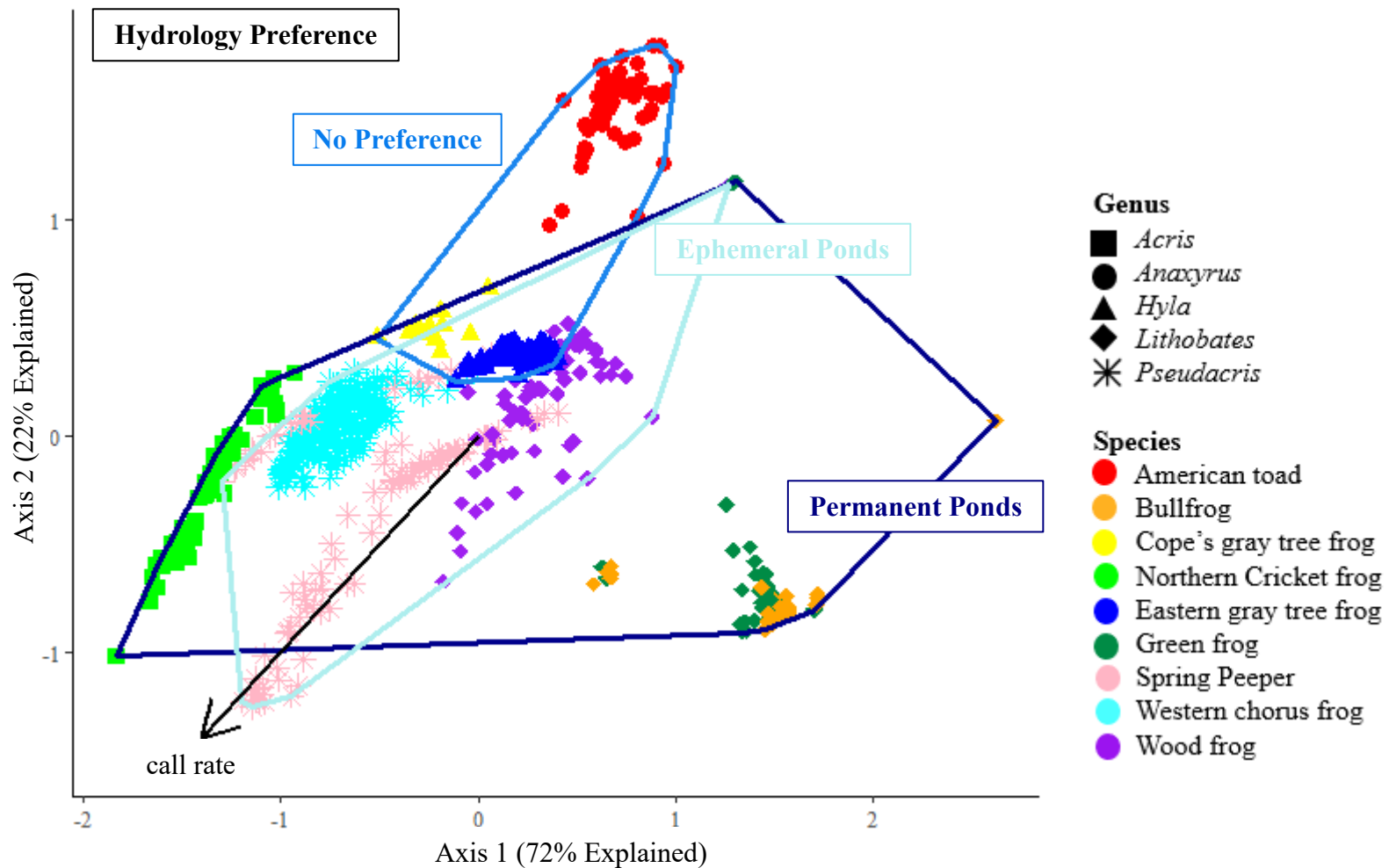
Appendix B. PCoA with grouping by genus, species, and hydrology. Polygons are drawn around species with the same hydrology preferences for breeding pools. There is considerable overlap for across all groups. The arrow represents the number of calls counted in each sample.
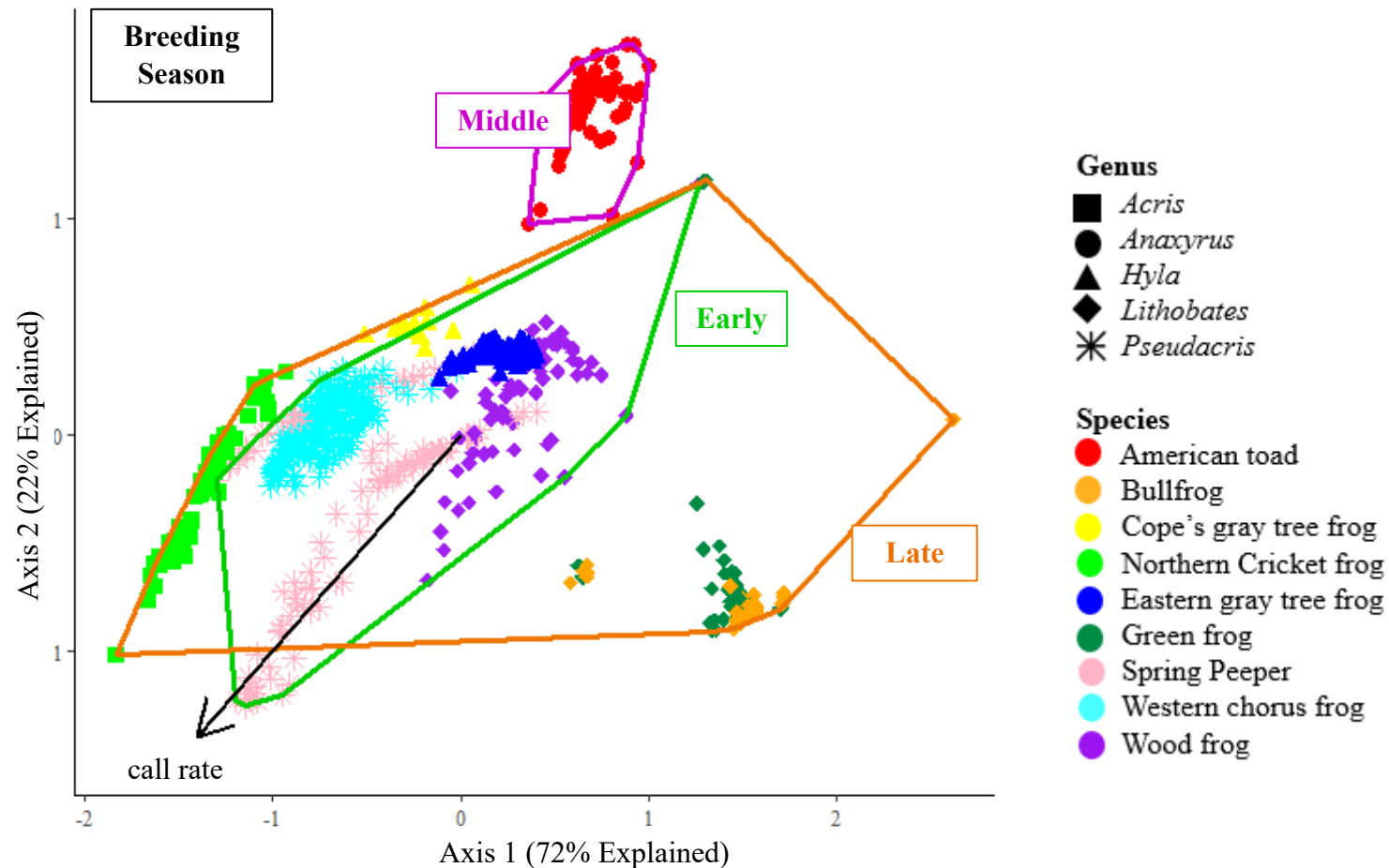
Appendix C. PCoA with grouping by genus, species, and breeding season timing. Polygons are drawn around species that breed during the same time. There is considerable overlap for the early and late groups. The noticeable separation for "middle" cannot be decoupled from the other unique traits of the American toad, so it is impossible to say if this separation is affected by breeding season. The arrow represents the number of calls
counted in each sample.

110
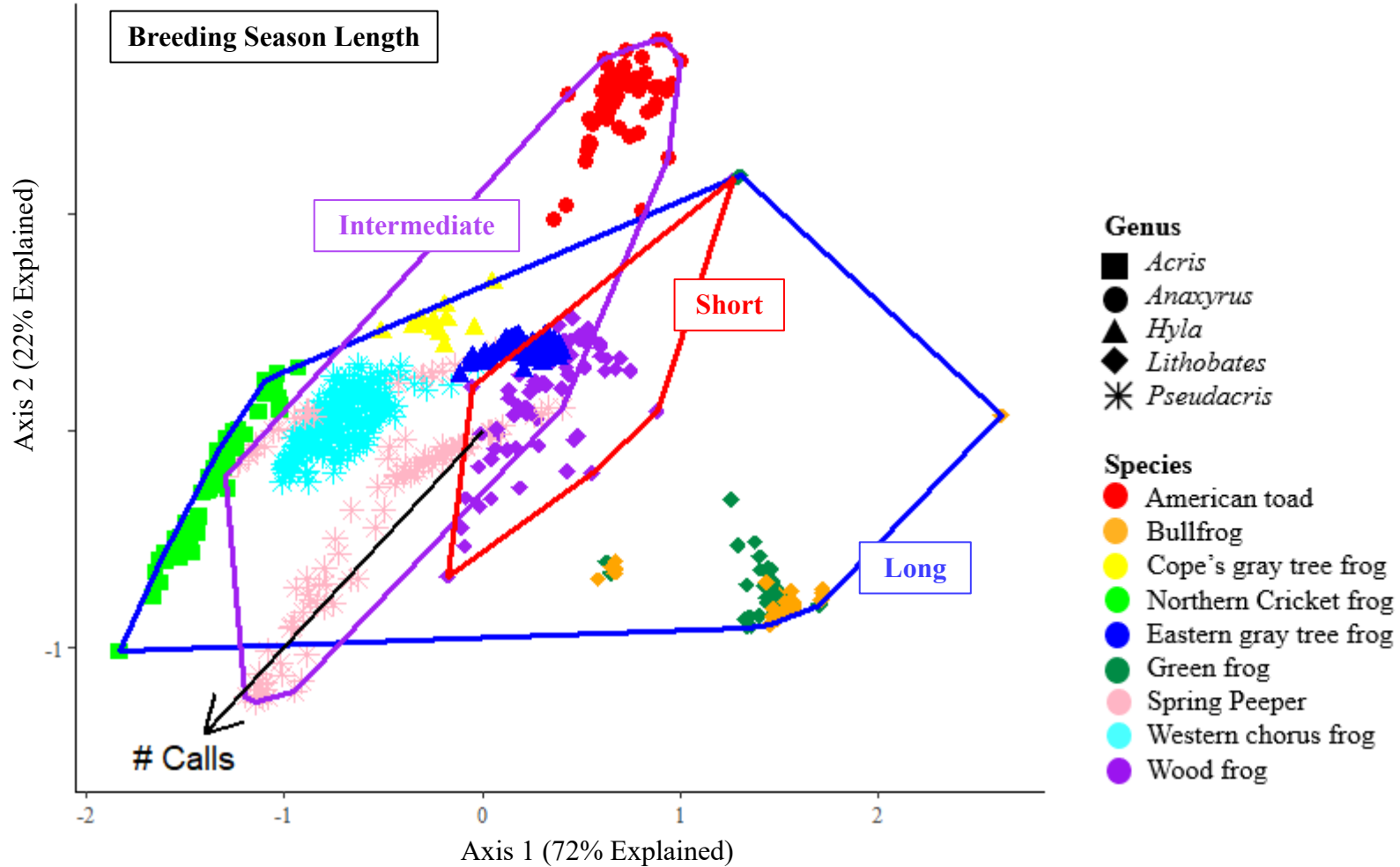
Appendix D. PCoA with grouping by genus, species, and breeding season length Polygons are drawn around species with same categorization for length of breeding season. There is considerable overlap for across all groups. The arrow represents the number of calls counted in each sample.
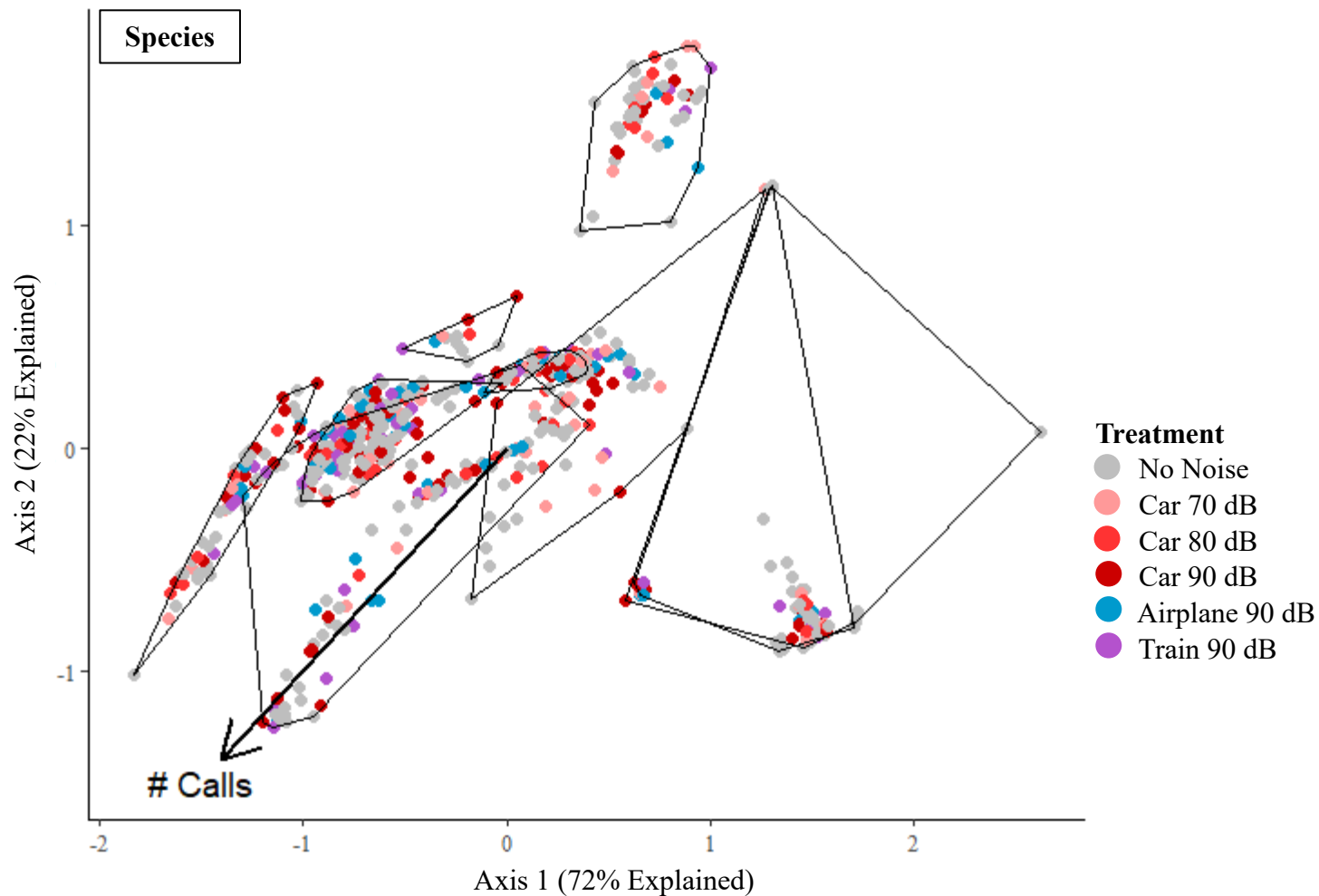
Appendix E. PCoA with grouping by species and treatment. Polygons are drawn around species. There is no organization by treatment, as call attributes are confined by species. Within each species there is no obvious separation between noise (colored dots) and no noise (gray dots). The arrow represents the number of calls counted in each sample.
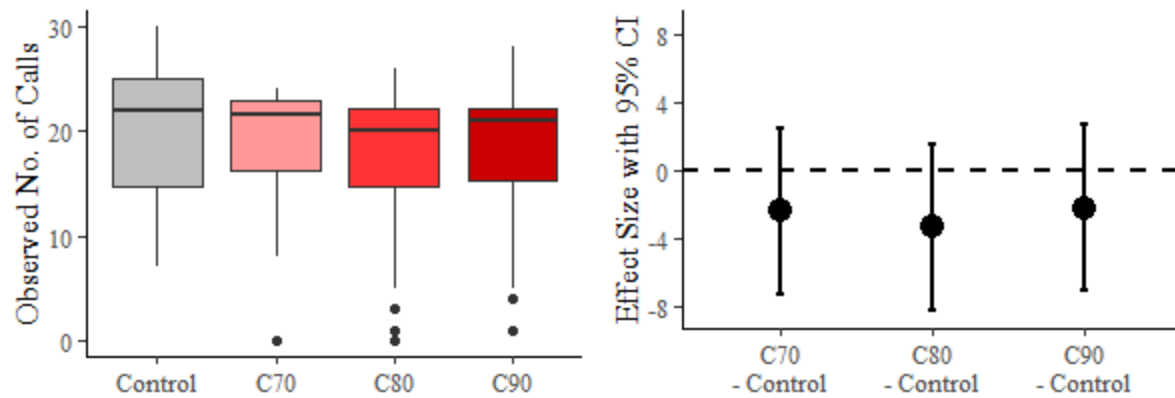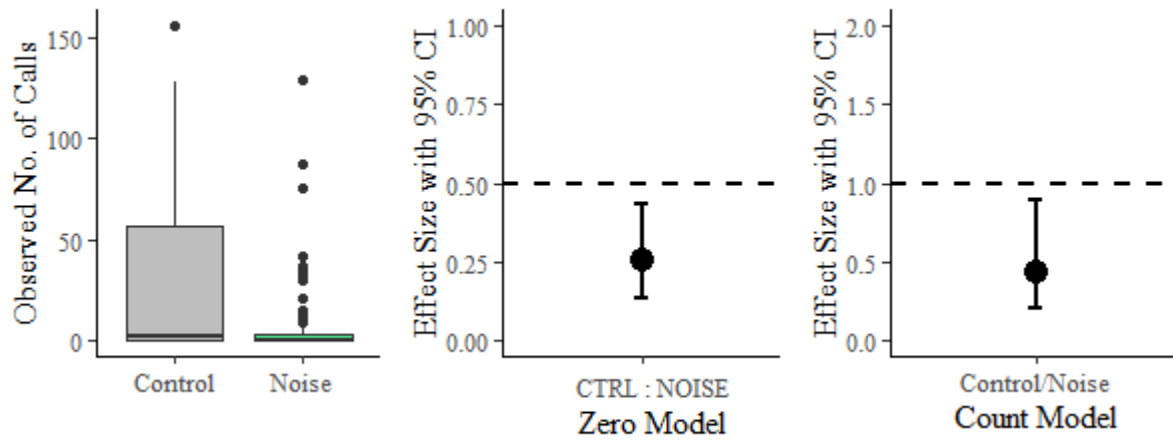
A) Control *vs.* noise
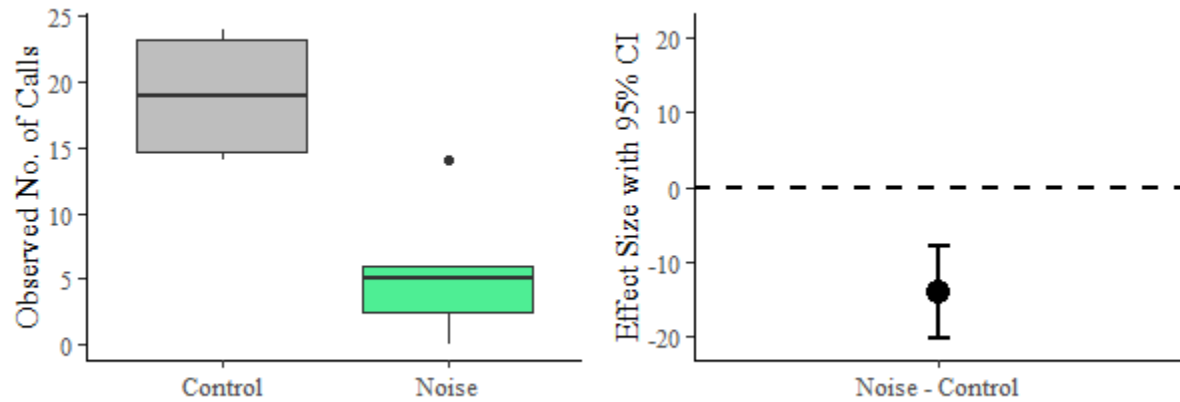
B) Car *vs.* plane *vs.* train
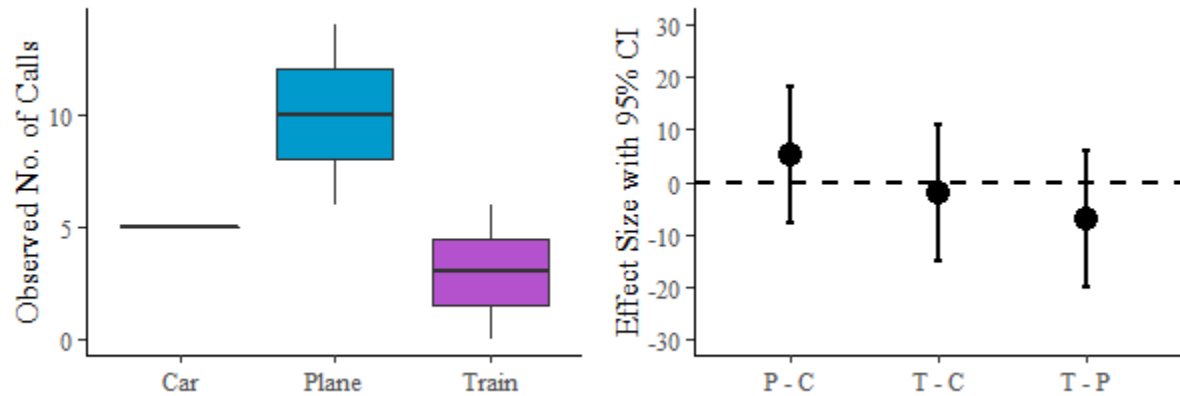
Appendix F. Impact of three types of noise (car, plane, and train) at 90 dB on call rate for the Cope's gray treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (call number ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (call number ~ treatment). For more detail of effect size estimates, refer to Appendix A.
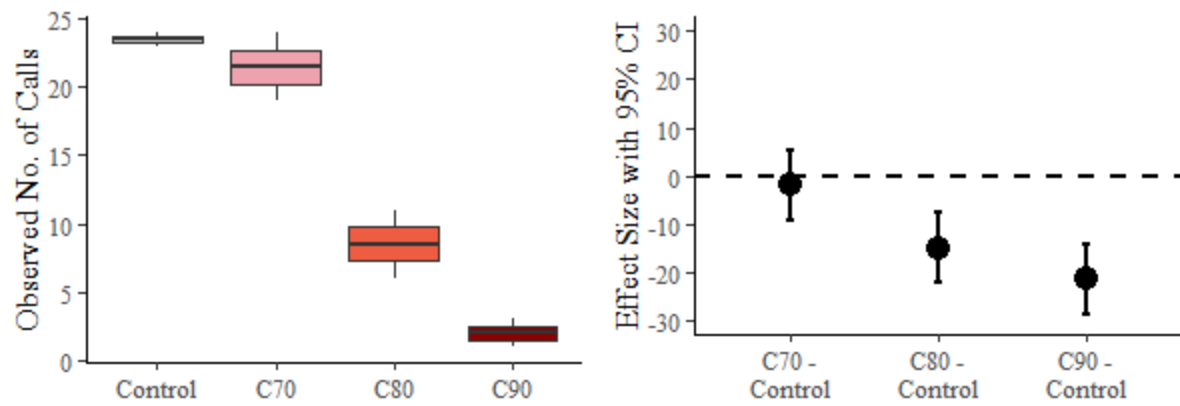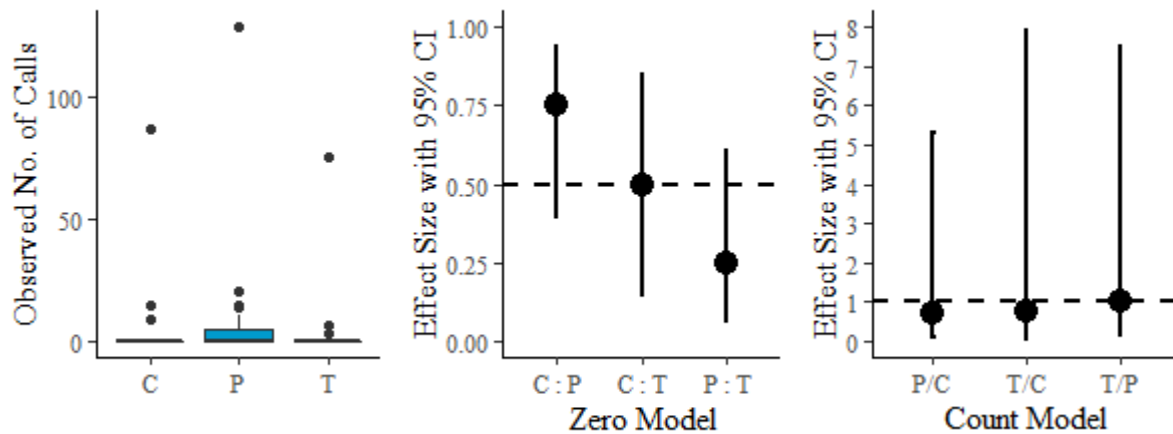
Appendix G. Effect of intensity of car noise on call rate for the Cope's gray treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call number ~ treatment). For more detail of effect size estimates, refer to Appendix A.
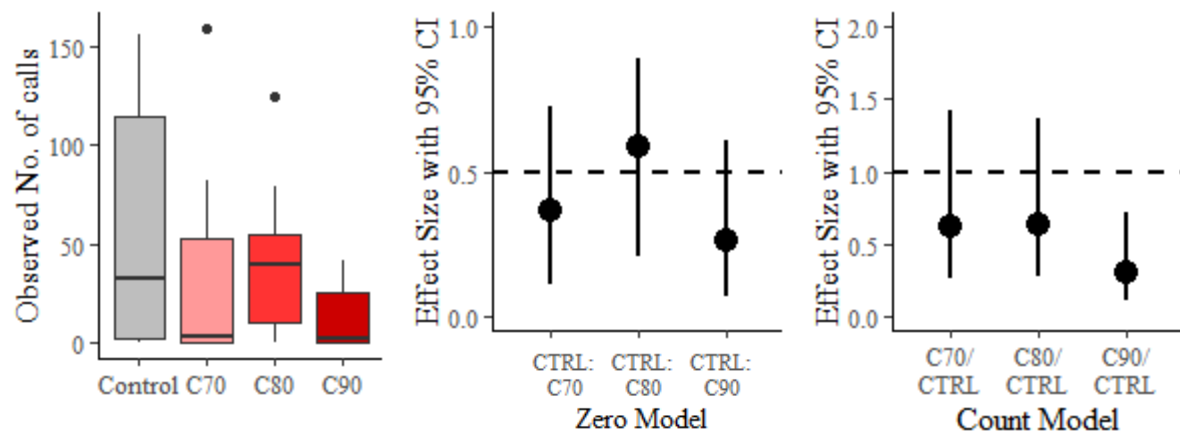
A) Control *vs.* noise
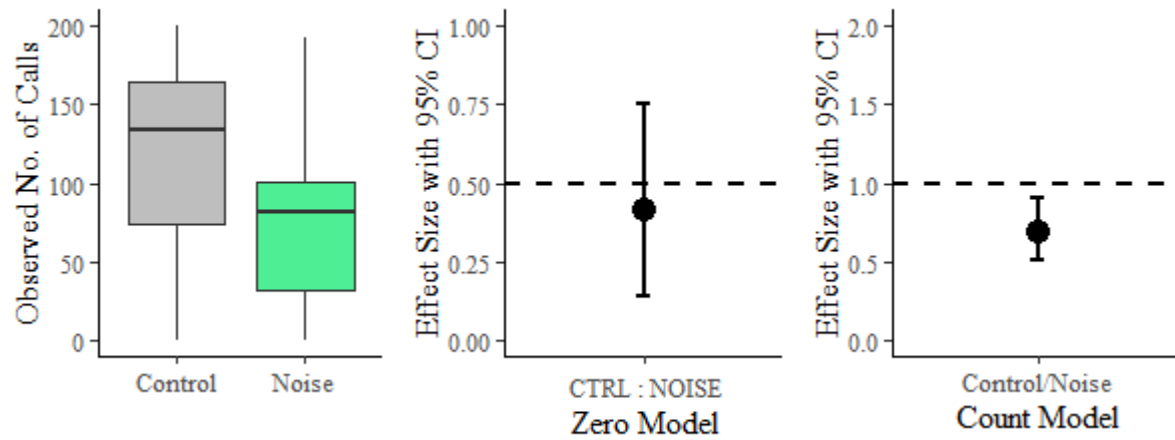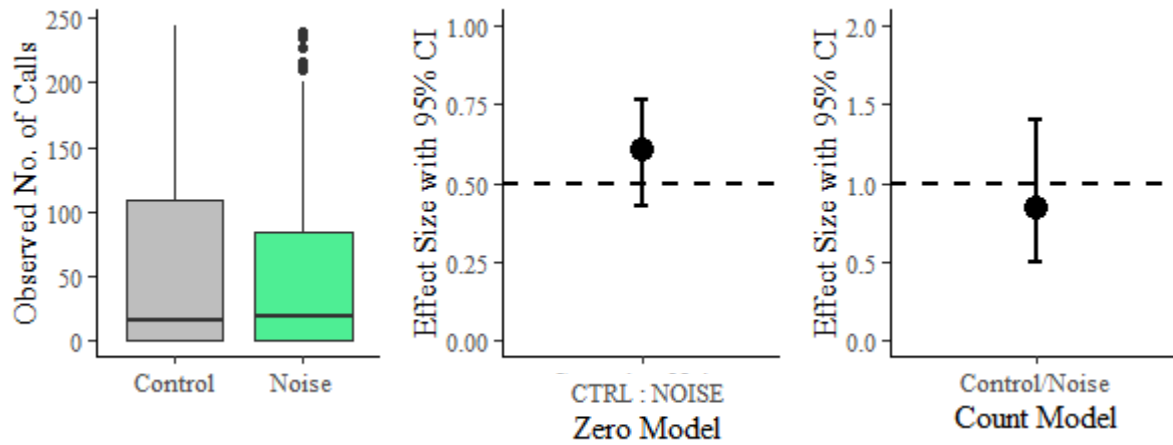

B) Car *vs.* plane *vs.* train

Appendix H. Impact of three types of noise (car, plane, and train) at 90 dB on call rate for the northern cricket frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (call number ~ treatment). For more detail of effect size estimates, refer to Appendix A.
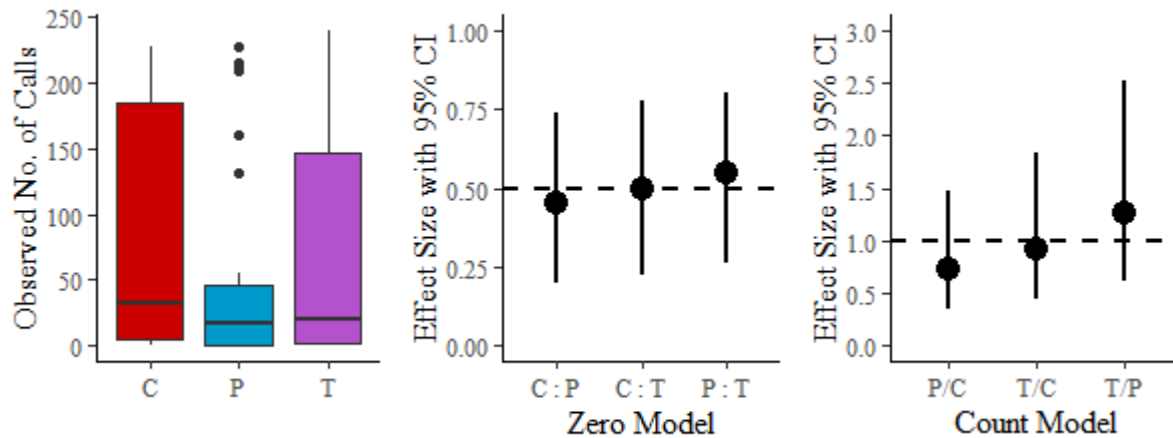
Appendix I. Effect of intensity of car noise on call rate for the northern cricket frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lmer (call number ~ treatment + 1|block). For more detail of effect size estimates, refer to Appendix A.

A) Control *vs.* noise

B) Car *vs.* plane *vs.* train

Appendix J. Impact of three types of noise (car, plane, and train) at 90 dB on call rate for the American bullfrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (call number ~ treatment). For more detail of effect size estimates, refer to Appendix A.

117
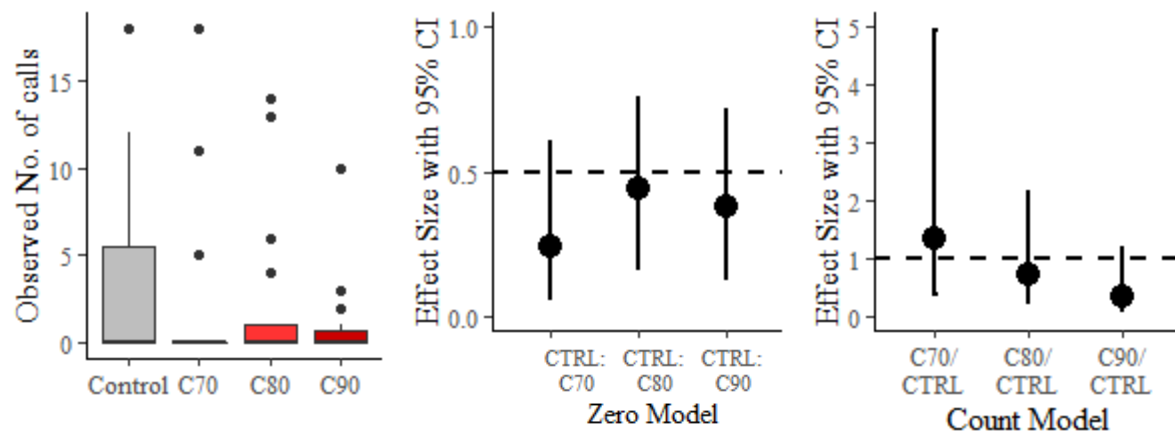
Appendix K. Effect of intensity of car noise on call rate for the American bullfrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.
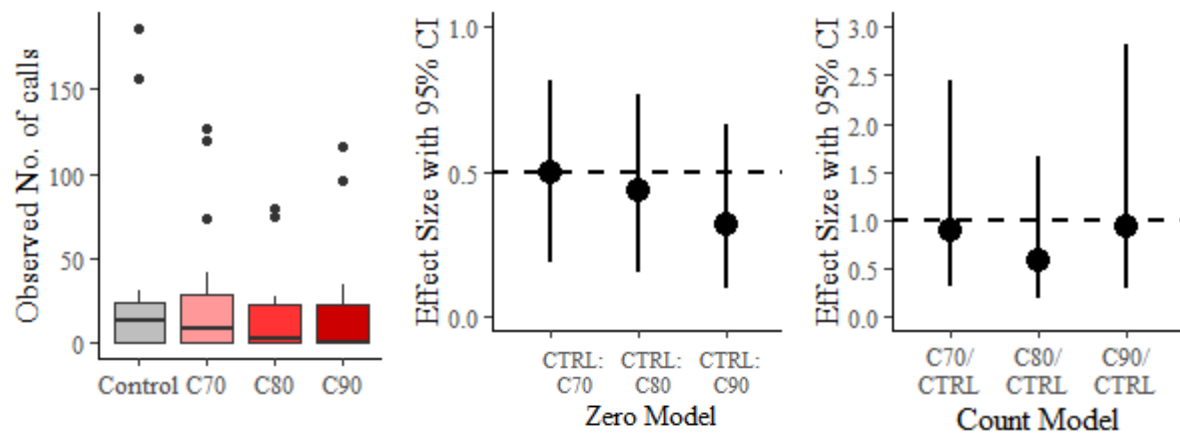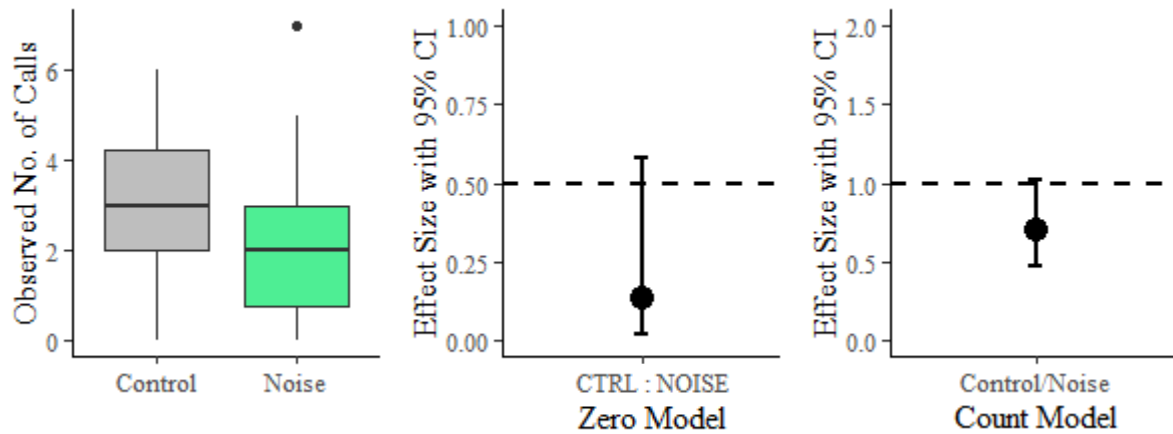
A) Control *vs.* noise

B) Car *vs.* Plane *vs.* train
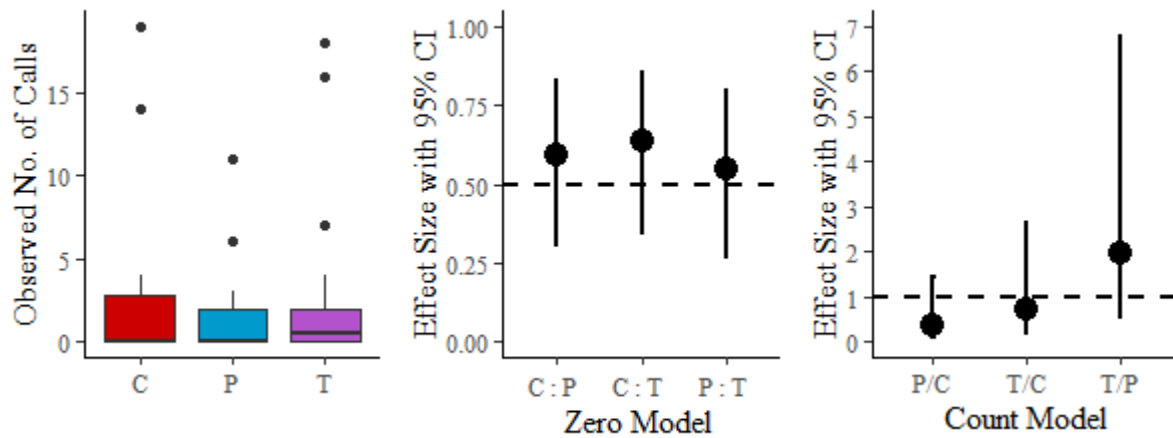
Appendix L. Impact of three types of noise (car, plane, and train) at 90 dB on call rate for the green frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.
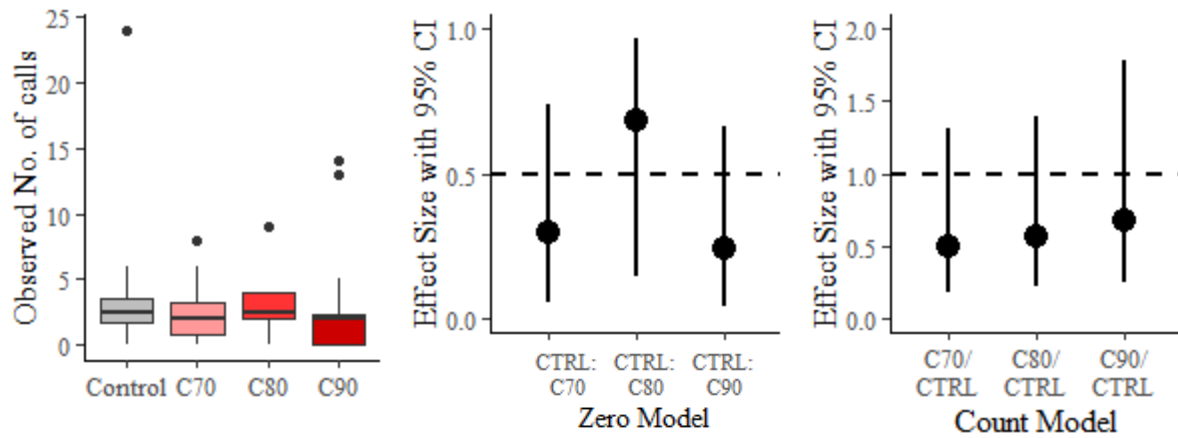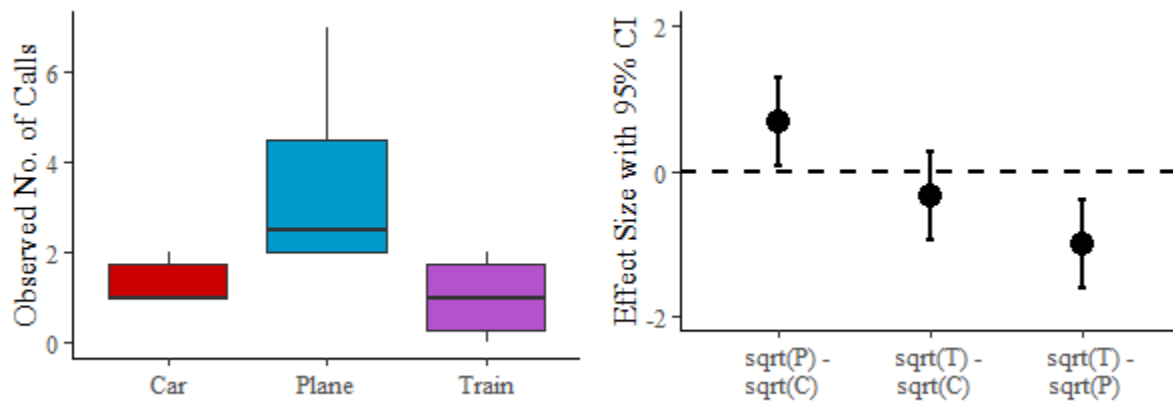
Appendix M. Effect of intensity of car noise on call rate for the green frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.

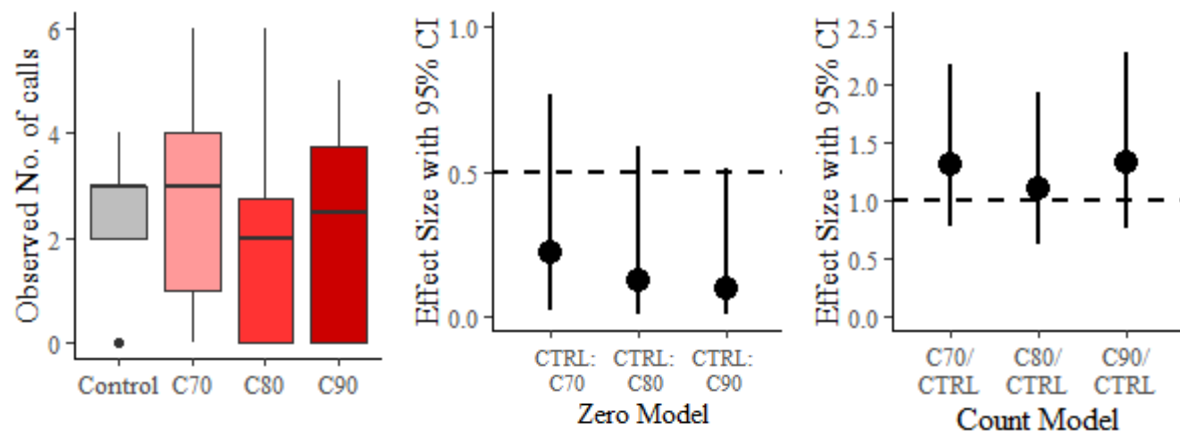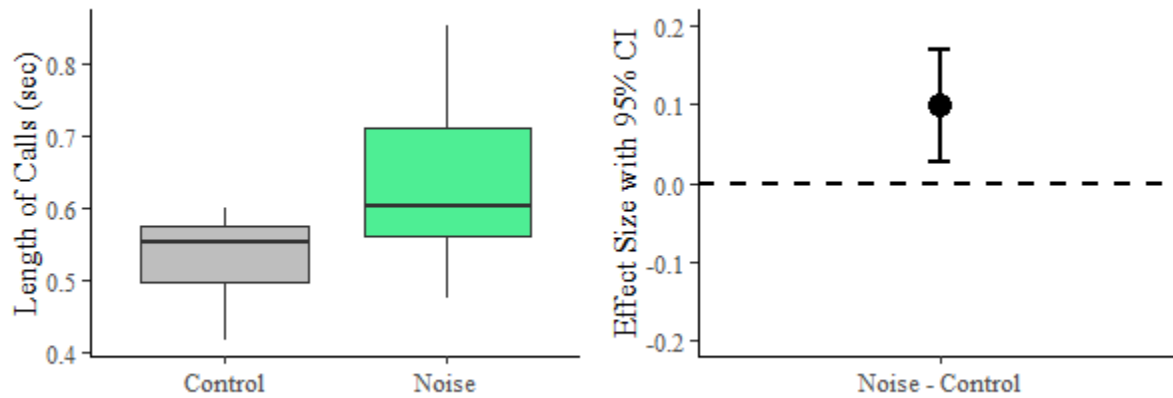A) Control *vs.* noise
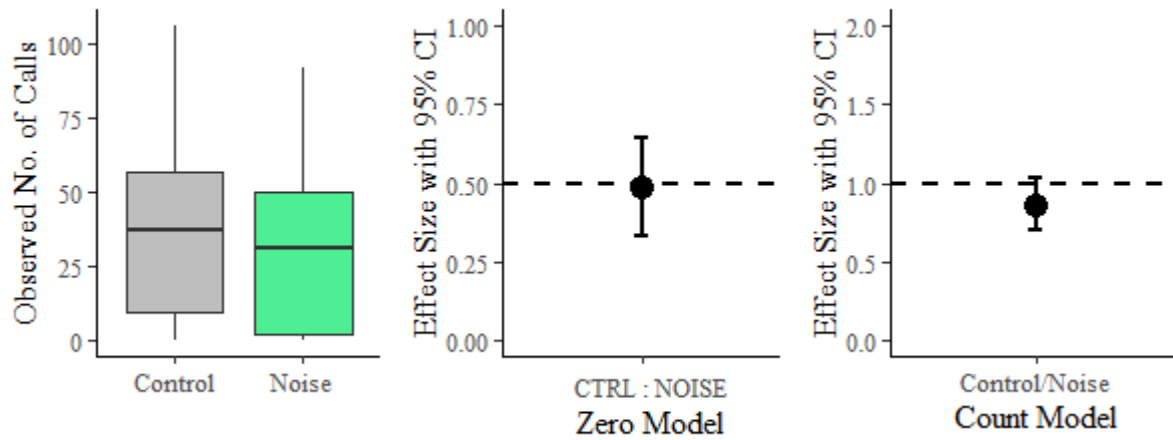
B) Car *vs.* plane *vs.* train
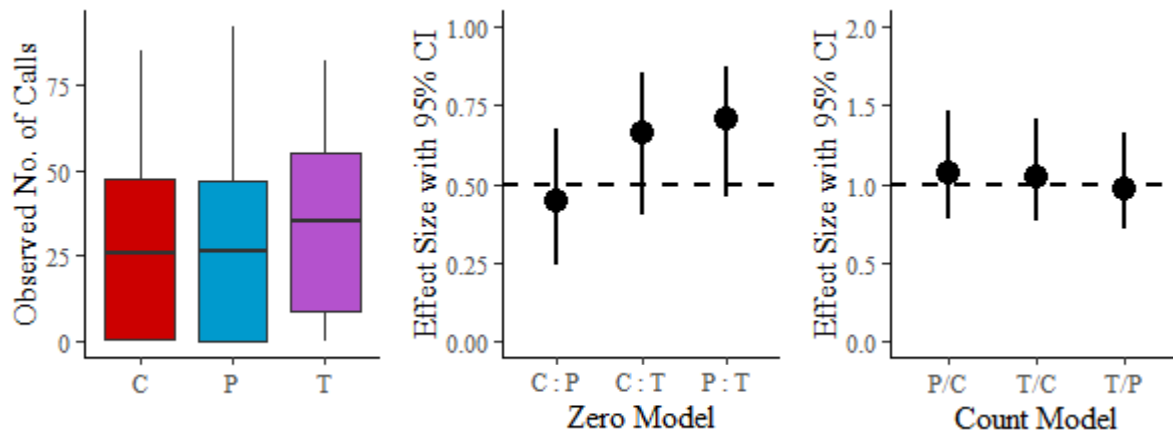
Appendix N. Impact of three types of noise (car, plane, and train) at 90 dB on call rate for the western chorus frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.
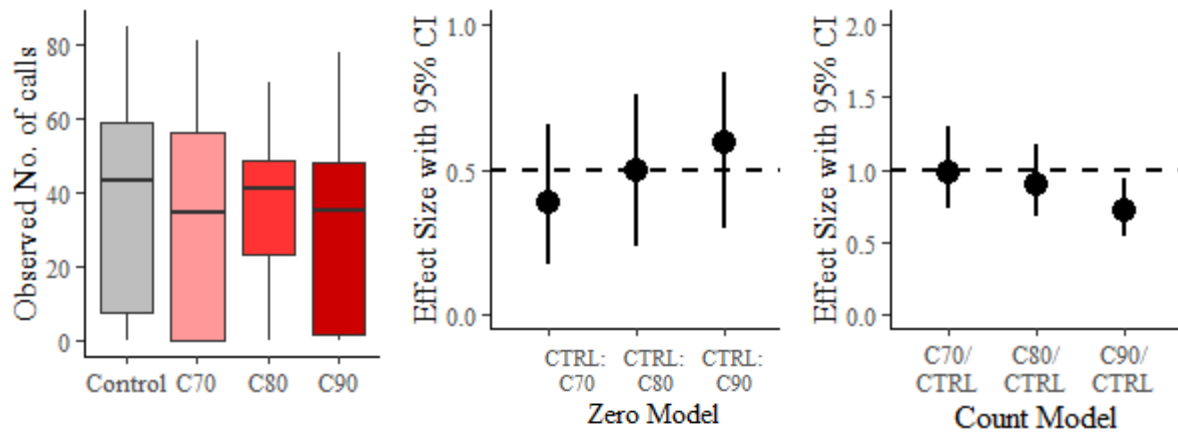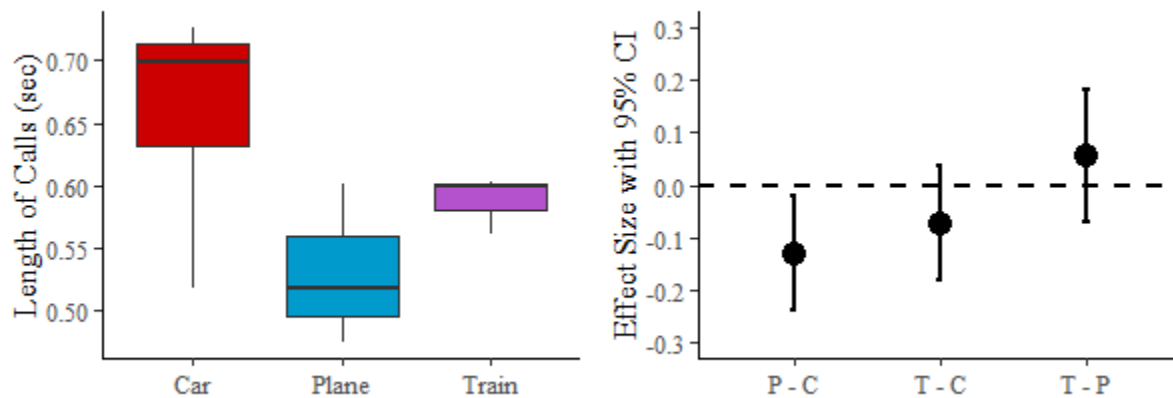
Appendix O. Effect of intensity of car noise on call rate for the western chorus frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.
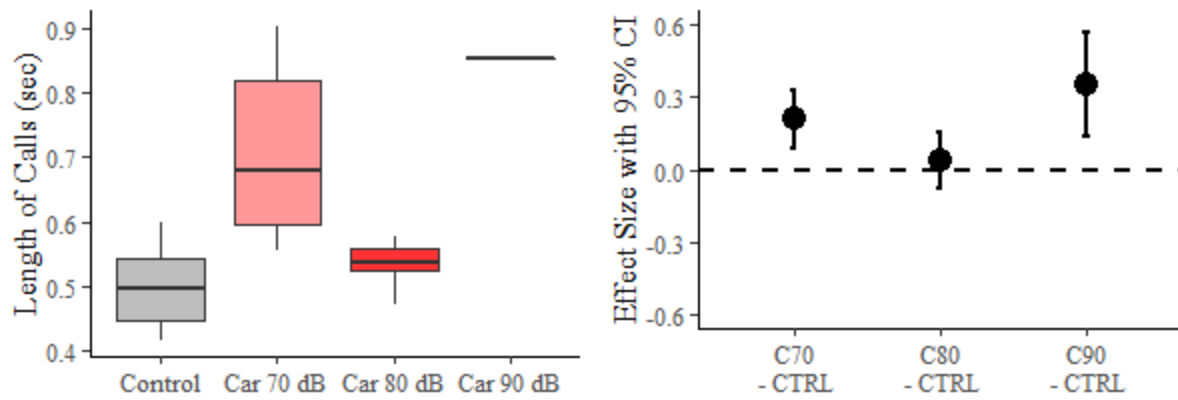
A) Control *vs.* Noise
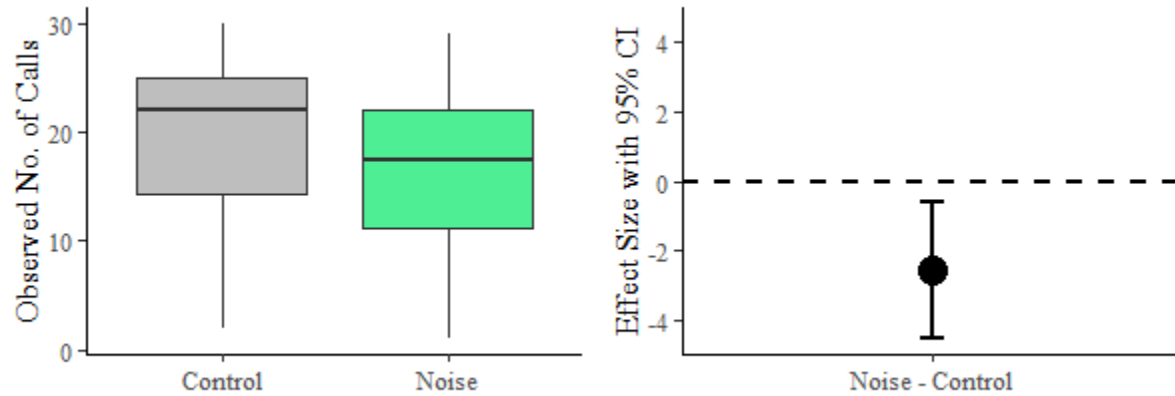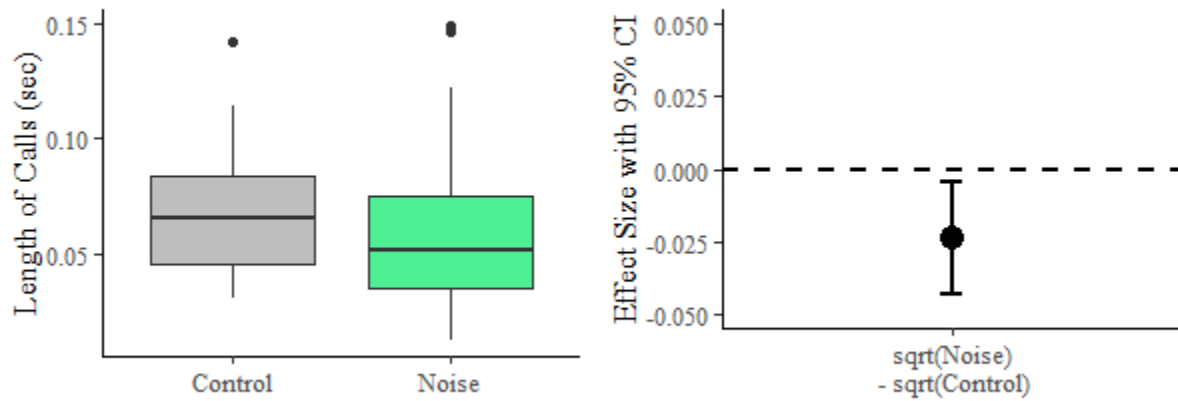

B) Car *vs.* Plane *vs.* Train
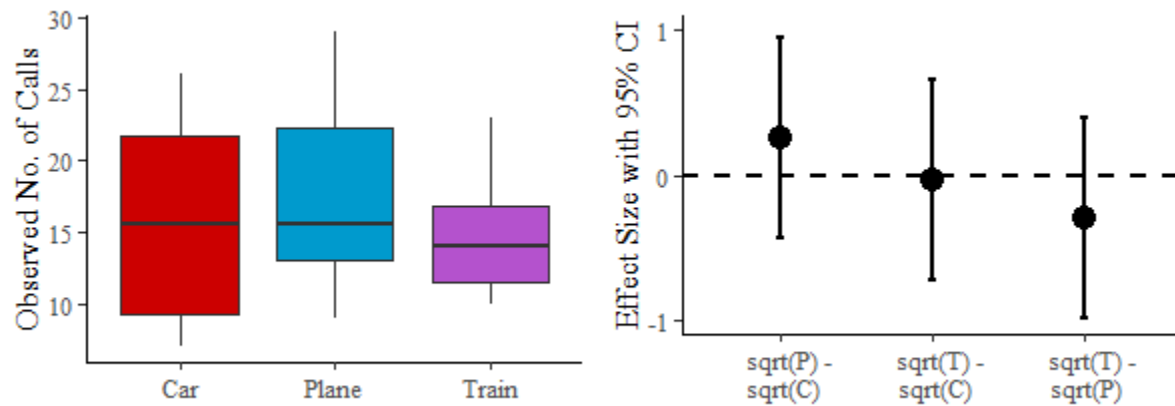
Appendix P. Impact of three types of noise (car, plane, and train) at 90 dB on call rate for the eastern gray tree frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lmer (call number ~ treatment + 1|block). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (sqrt (call number) ~ treatment). For more detail of effect size estimates, refer to Appendix A.
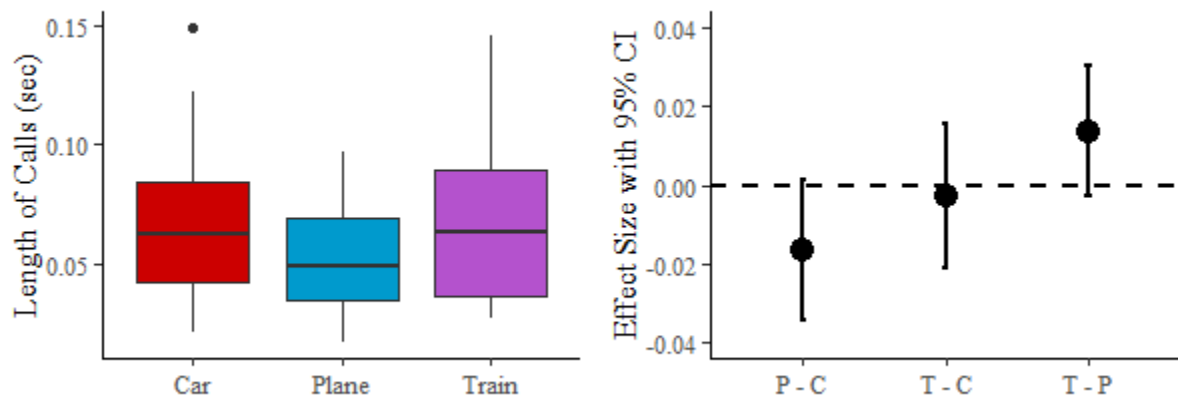
Appendix Q. Effect of intensity of car noise on number of calls for the eastern gray treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call number ~ treatment + offset (initial conditions)). For more detail of effect size estimates, refer to Appendix A.
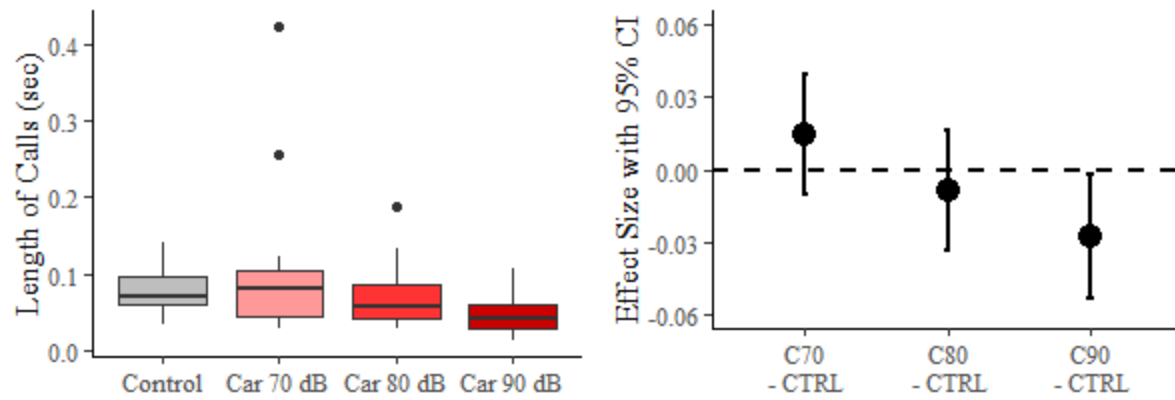
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix R. Impact of Three Types of Noise (car, plane, and train) at 90 dB call rate for the wood frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.

Appendix S. Effect of intensity of car noise on call rate for the wood frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.

A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train
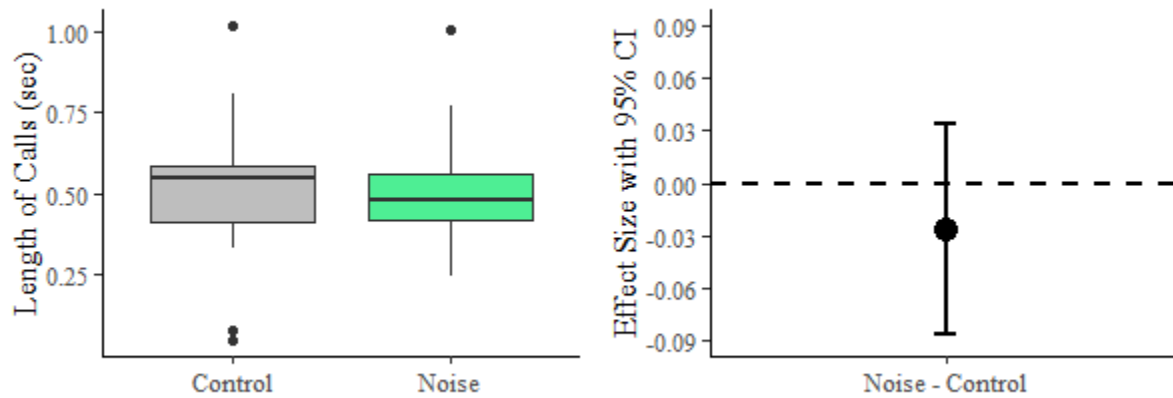

Appendix T. Impact of three types of noise (car, plane, and train) at 90 dB call rate for the spring peeper. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.
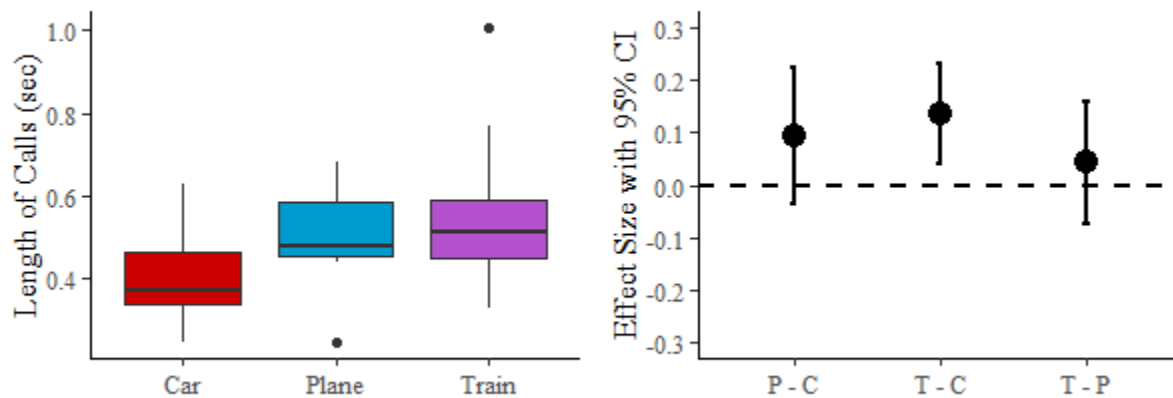
Appendix U. Effect of intensity of car noise on call rate for the spring peeper. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = negative binomial). For more detail of effect size estimates, refer to Appendix A.
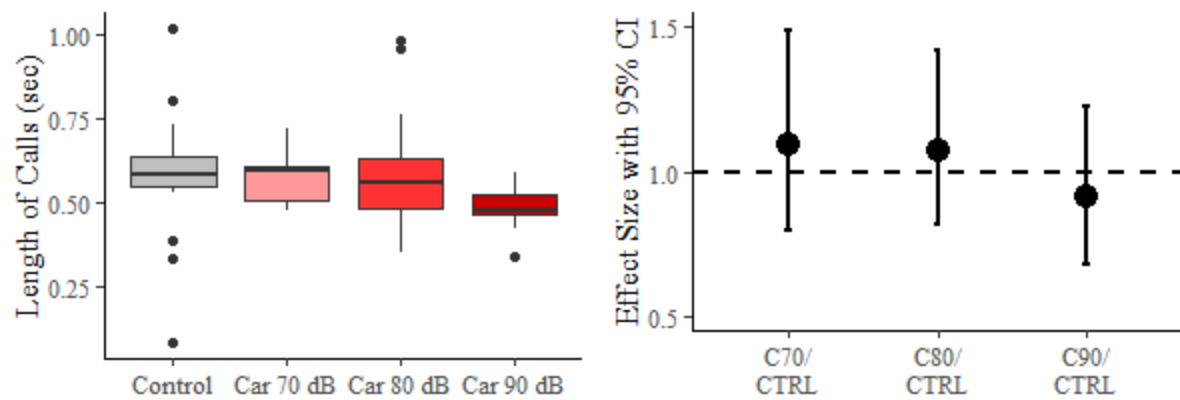
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

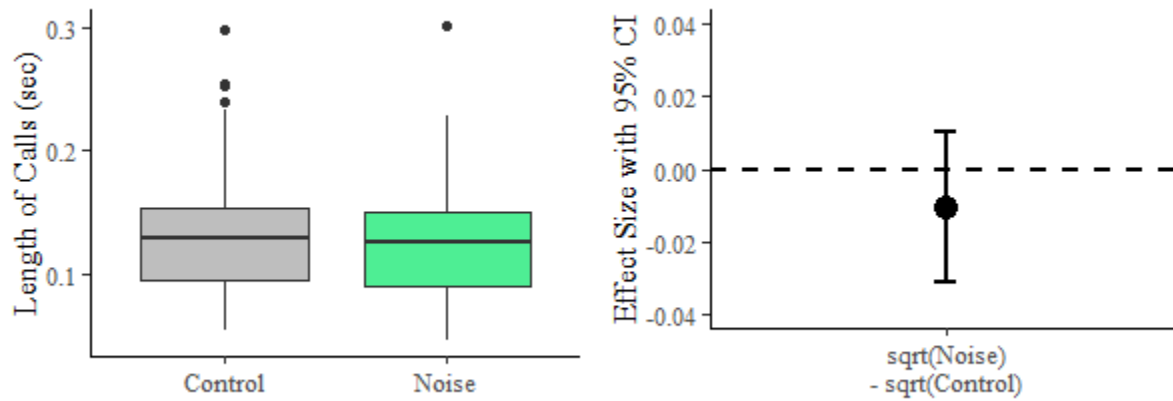
Appendix V. Impact of three types of noise (car, plane, and train) at 90 dB on call rate for the American toad. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = Poisson). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (sqrt(call number) ~ treatment). For more detail of effect size estimates, refer to Appendix A.

Appendix W. Effect of intensity of car noise on number of calls for the American Toad. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model hurdle (call number ~ treatment, distribution = Poisson). For more detail of effect size estimates, refer to Appendix A.
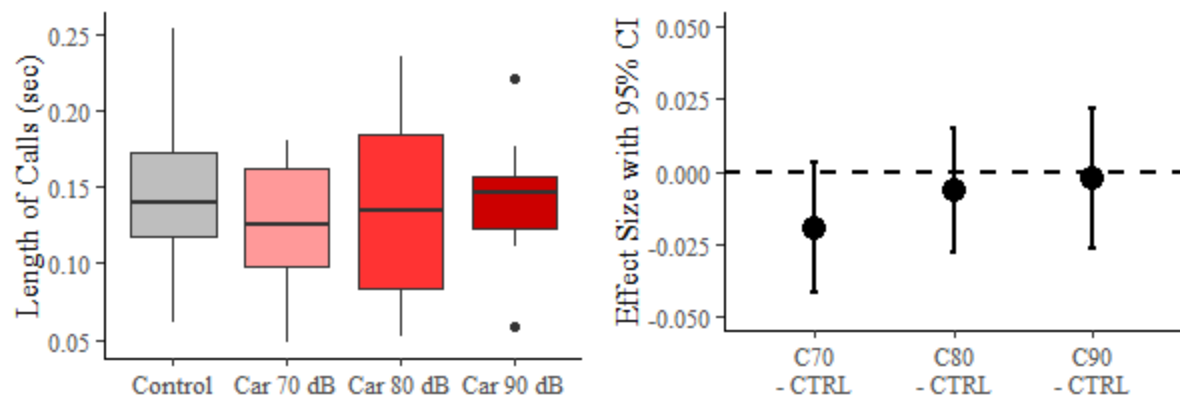
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix X. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the Cope's gray treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (call length ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.

Appendix Y. Effect of intensity of car noise on call length for the cope's gray treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
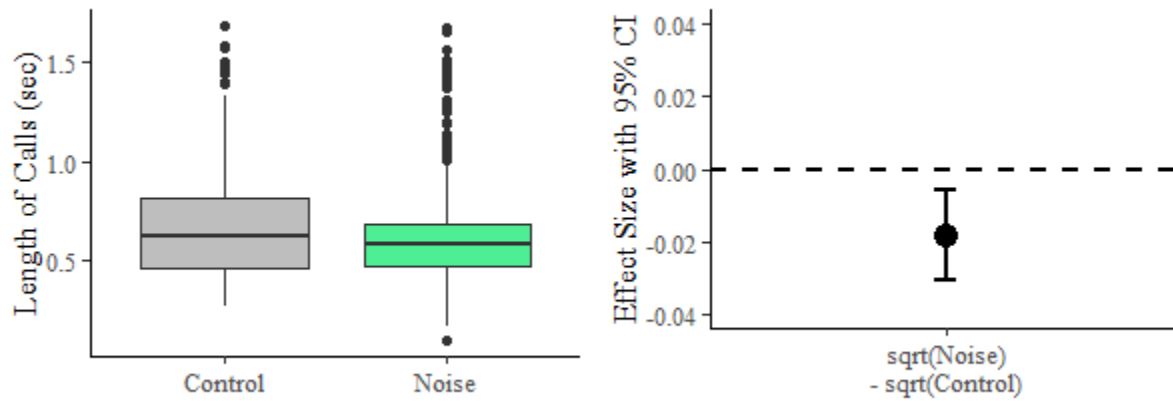
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix Z. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the northern cricket frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (sqrt(call length) ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
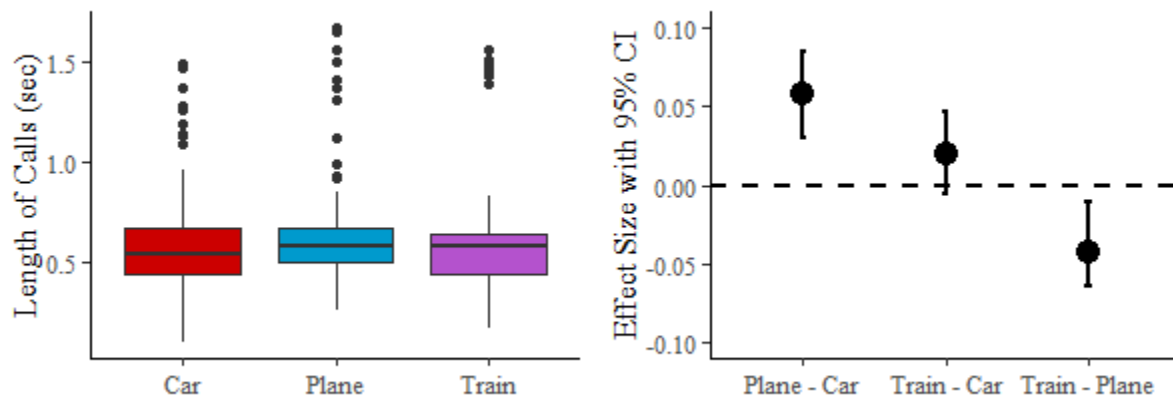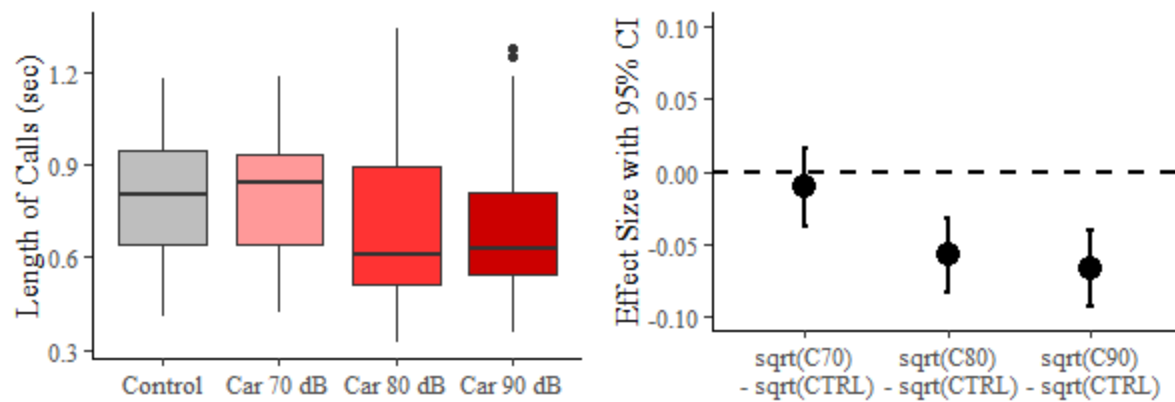
133

Appendix AA. Effect of intensity of car noise on call length for the northern cricket frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
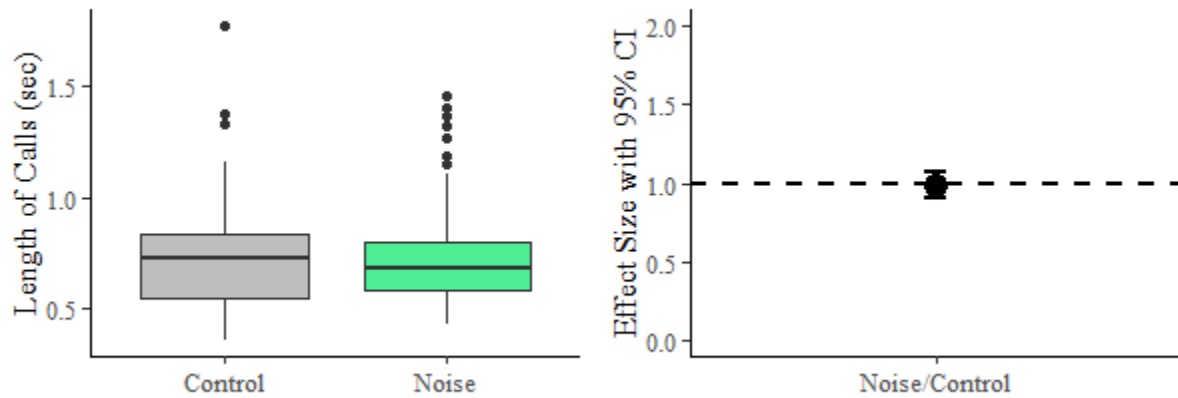
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix BB. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the American bullfrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (call length ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
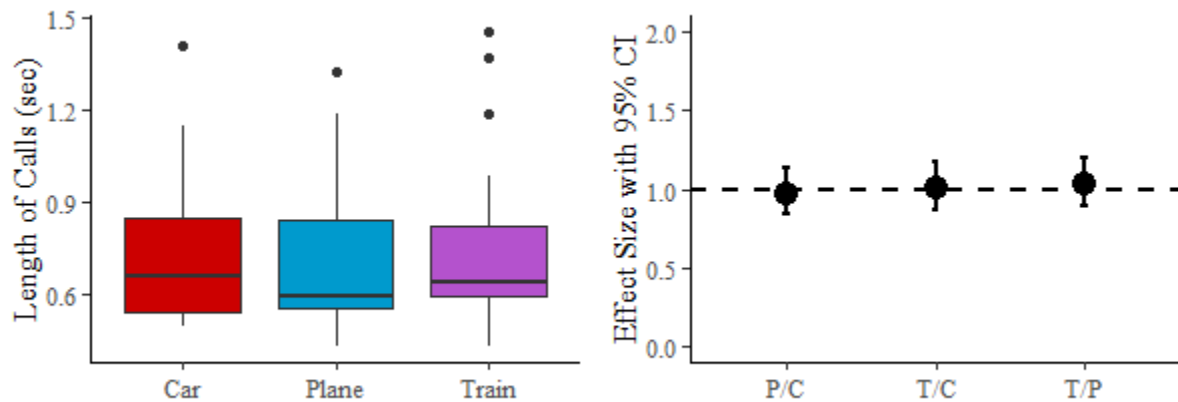
Appendix CC. Effect of intensity of car noise on call length for the American bullfrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
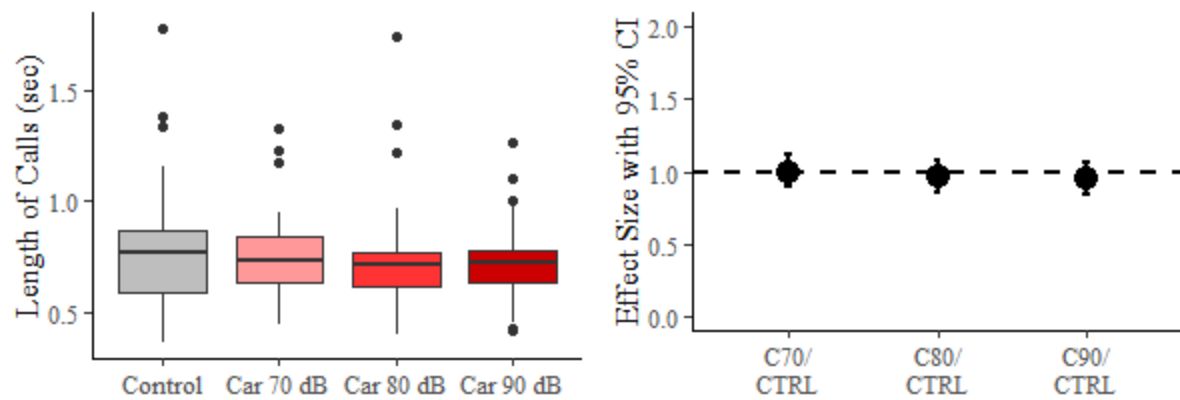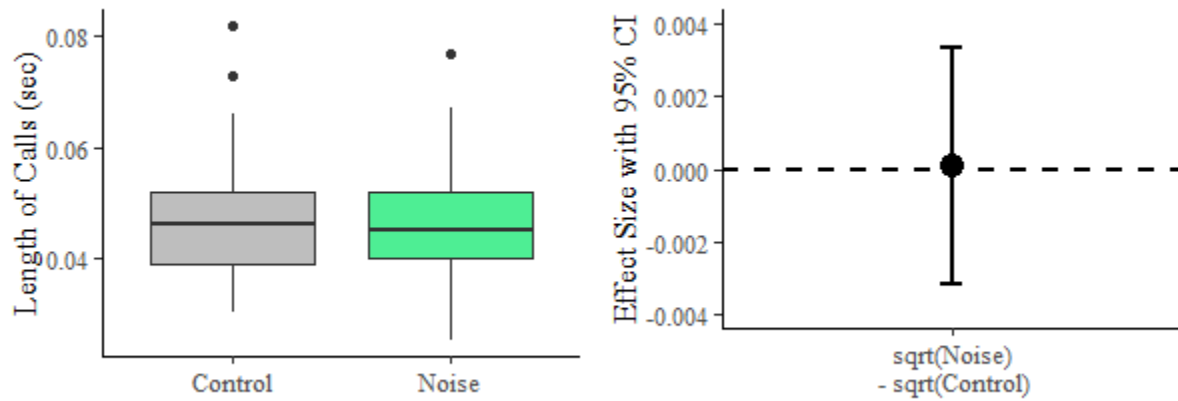
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train
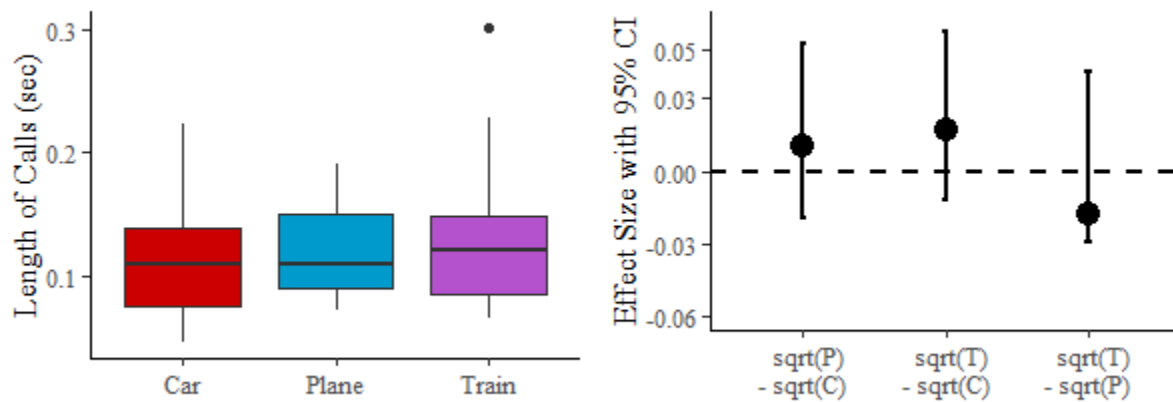
Appendix DD. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the Green Frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lmer (sqrt(call length) ~ treatment + (1|block)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model (sqrt(call length) ~ treatment + (1|block)). For more detail of effect size estimates, refer to Appendix A.
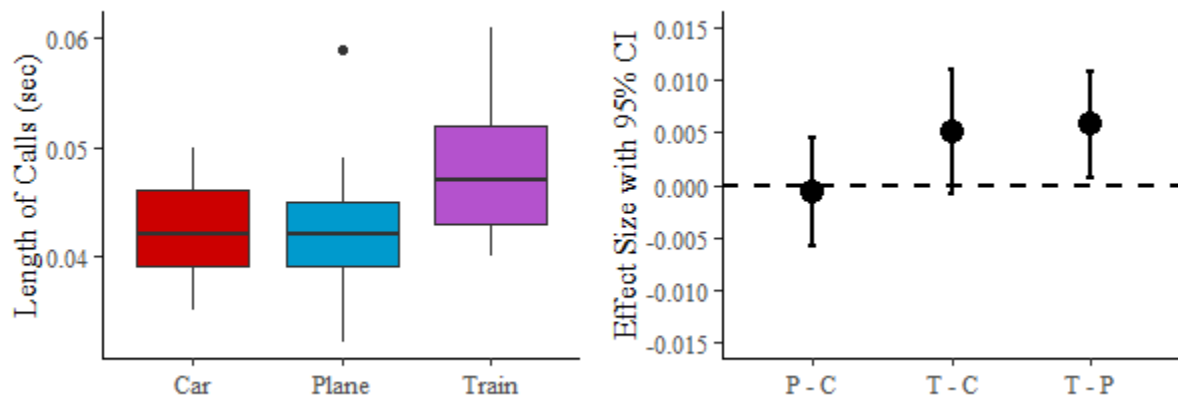
Appendix EE. Effect of intensity of car noise on call length for the green frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
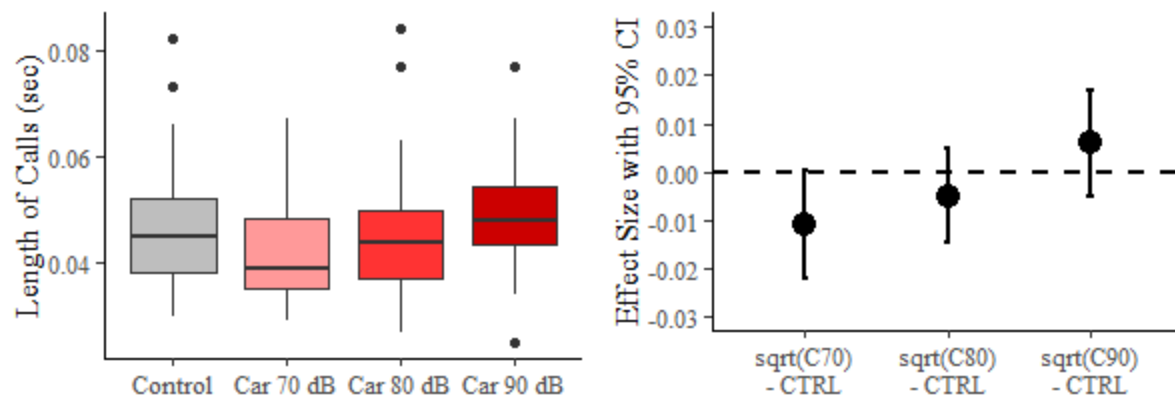
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix FF. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the Western Chorus Frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lmer (sqrt(call length) ~ treatment + (1|block)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the model (call length ~ treatment + (1|block)). For more detail of effect size estimates, refer to Appendix A.

Appendix GG. Effect of intensity of car noise on call length for the western chorus frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lmer (sqrt(call length) ~ treatment + (1|block)). For more detail of effect size estimates, refer to Appendix A.

A) Control *vs.* Noise

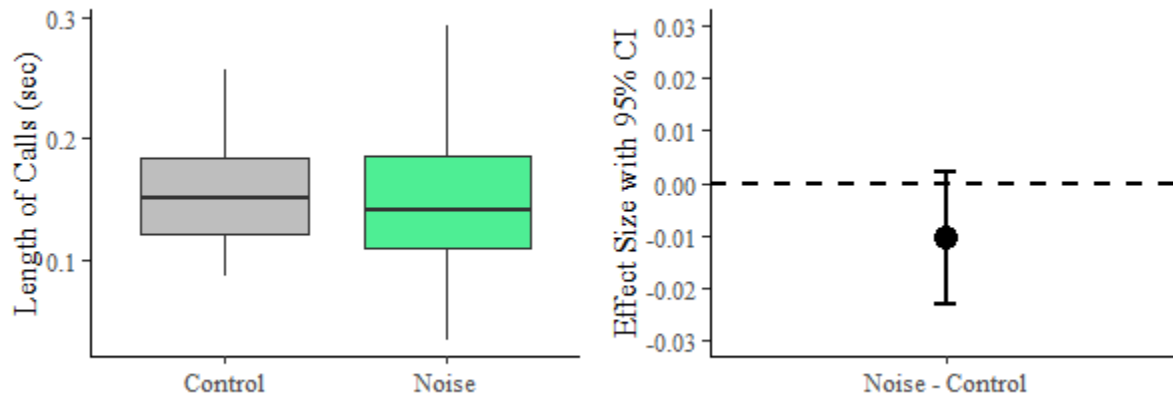B) Car *vs.* Plane *vs.* Train

Appendix HH. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the eastern gray treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (log(call length) ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (log(call length) ~ treatment). For more detail of effect size estimates, refer to Appendix A.
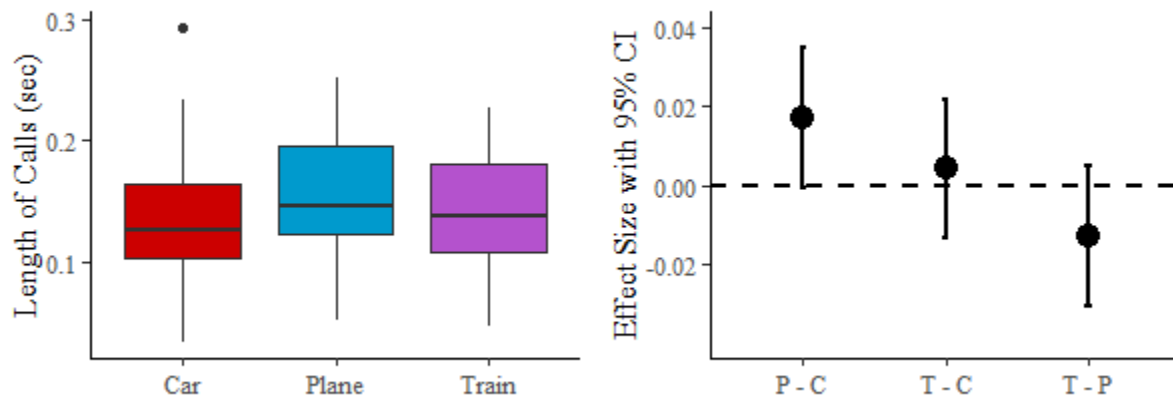
Appendix II. Effect of intensity of car noise on call length for the eastern gray Treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (log(call length) ~ treatment). For more detail of effect size estimates, refer to Appendix A.
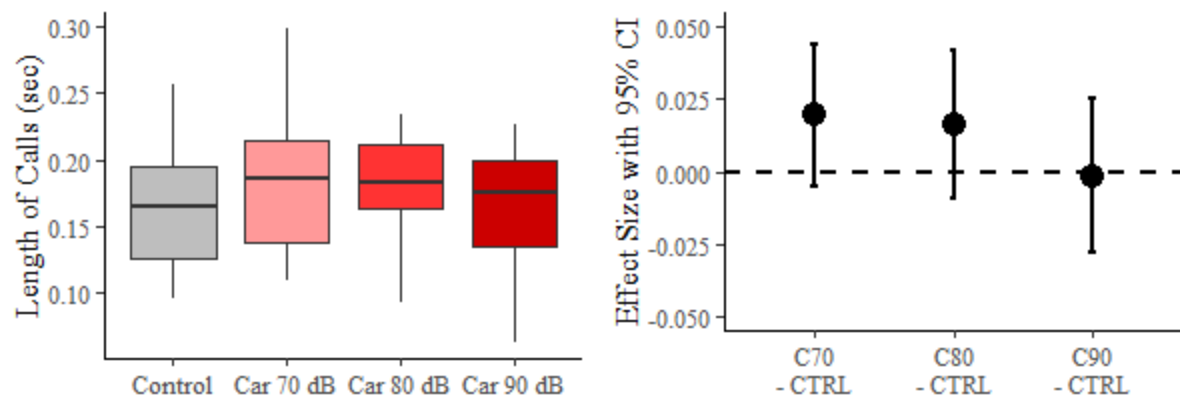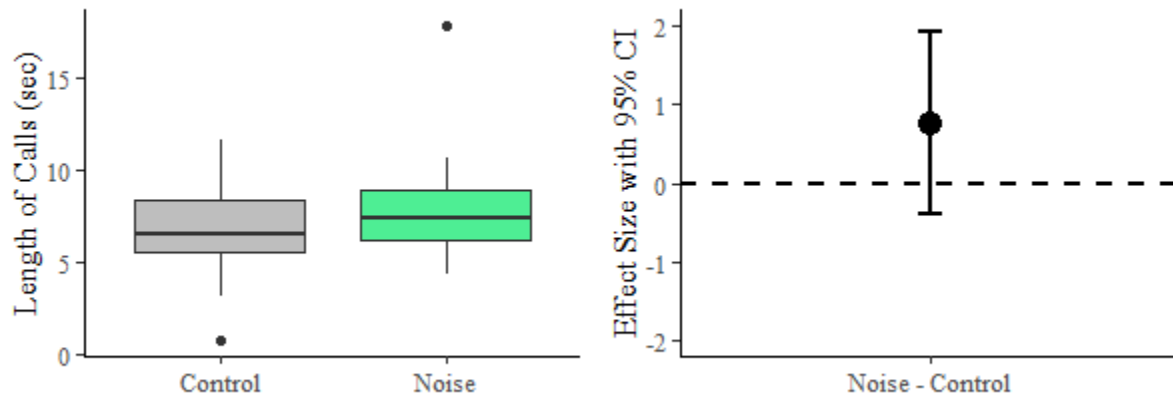
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix JJ. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the wood frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (sqrt(call length) ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.
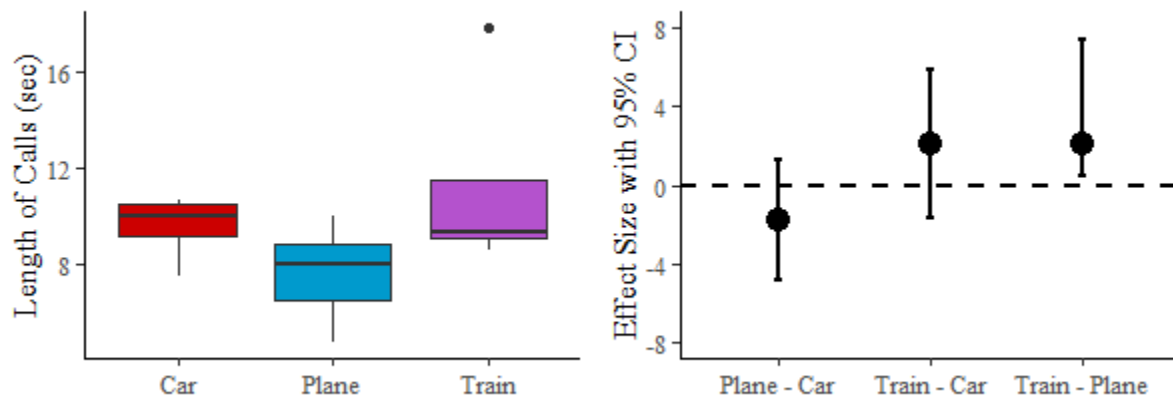
Appendix KK. Effect of intensity of car noise on call length for the wood frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (sqrt (call length) ~ treatment). For more detail of effect size estimates, refer to Appendix A.
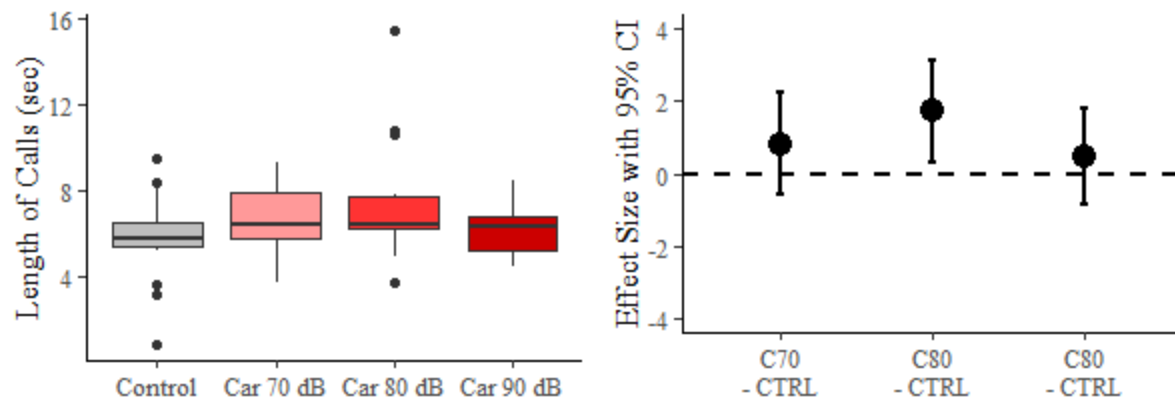
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix LL. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the Spring Peeper. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (call length ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (call length ~ treatment). For more detail of effect size estimates, refer to Appendix A.

Appendix MM. Effect of intensity of car noise on call length for the spring peeper. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (sqrt (call length) ~ treatment). For more detail of effect size estimates, refer to Appendix A.
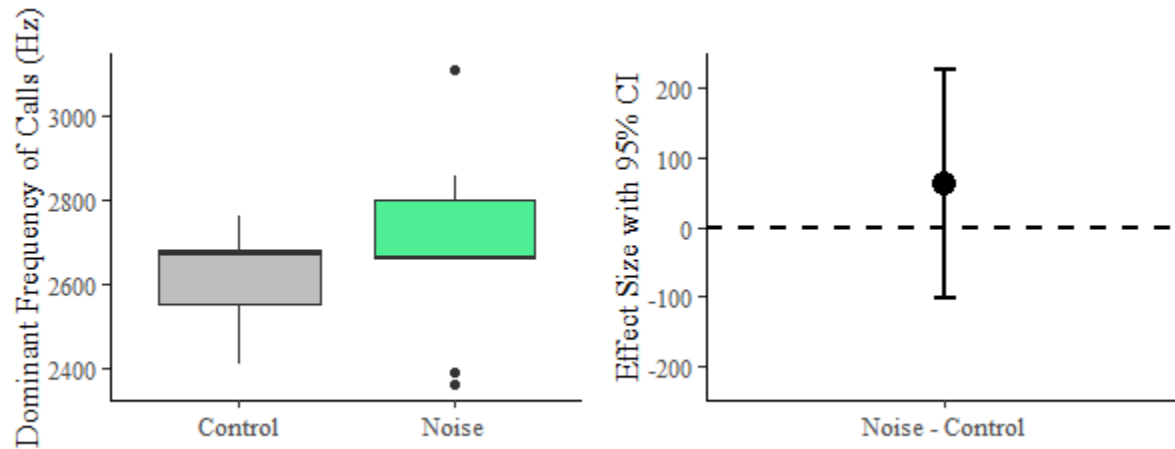
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix NN. Impact of three types of noise (car, plane, and train) at 90 dB on call length for the American toad. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (call length ~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (call length ~ treatment + offset(initial conditions). For more detail of effect size estimates, refer to Appendix A.

Appendix OO. Effect of intensity of car noise on call length for the American toad. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call length ~ treatment + offset(initial conditions). For more detail of effect size estimates, refer to Appendix A.
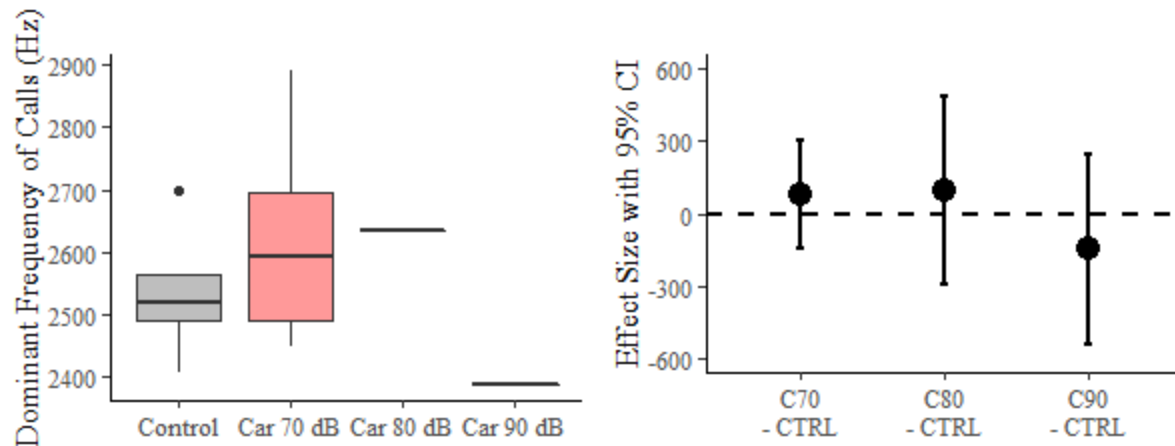
A) Control *vs.* Noise
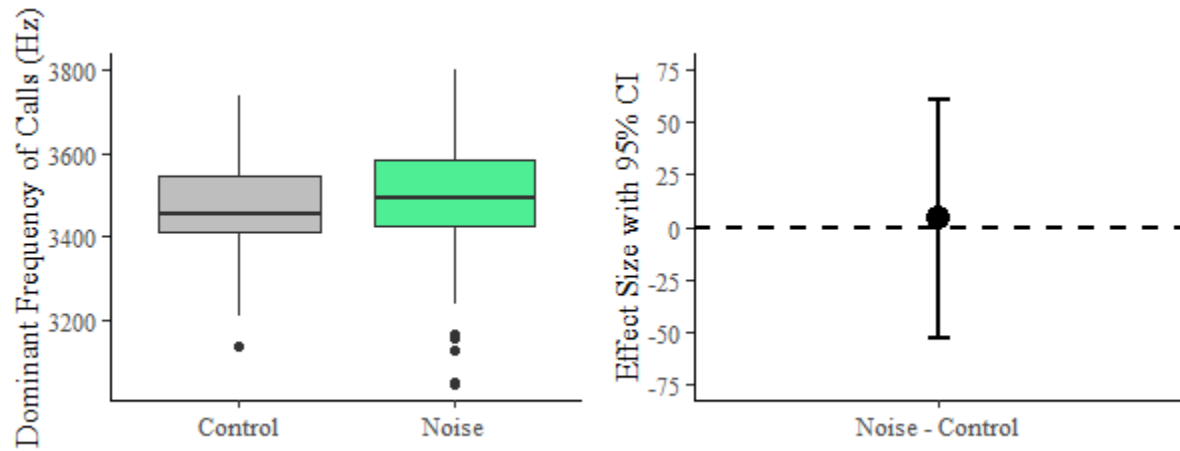

B) Car *vs.* Plane *vs.* Train
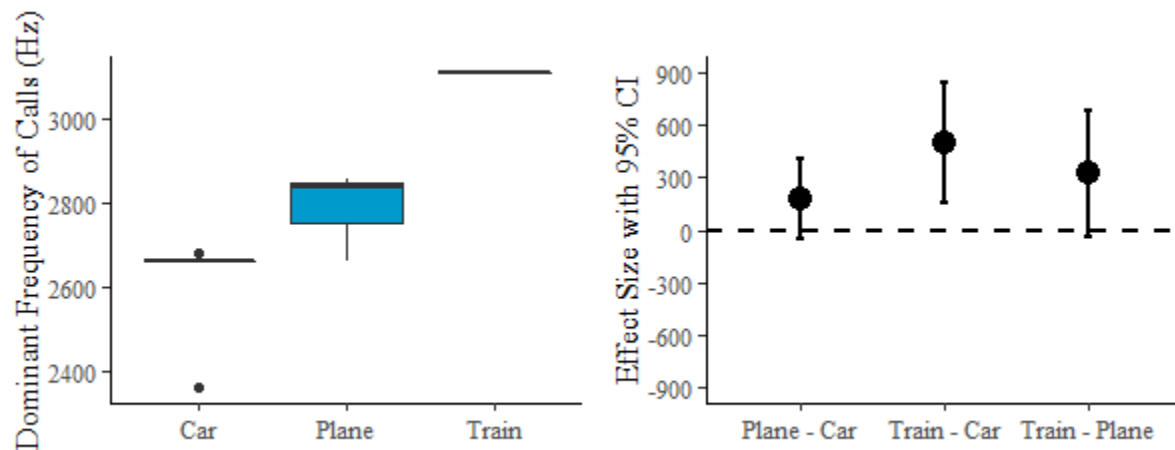
Appendix PP. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the Cope's Gray Treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
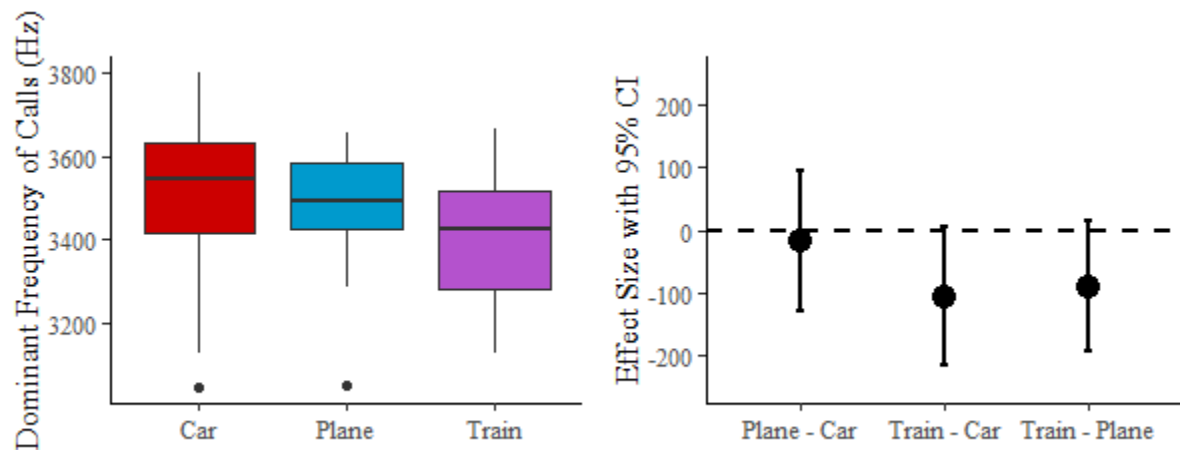
Appendix QQ. Effect of intensity of car noise on dominant frequency for the Cope's gray treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
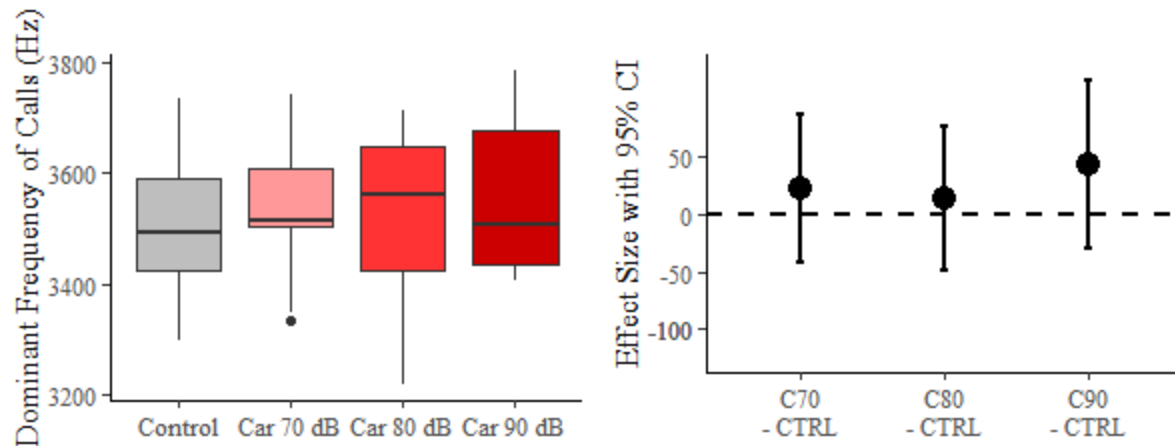
A) Control *vs.* Noise
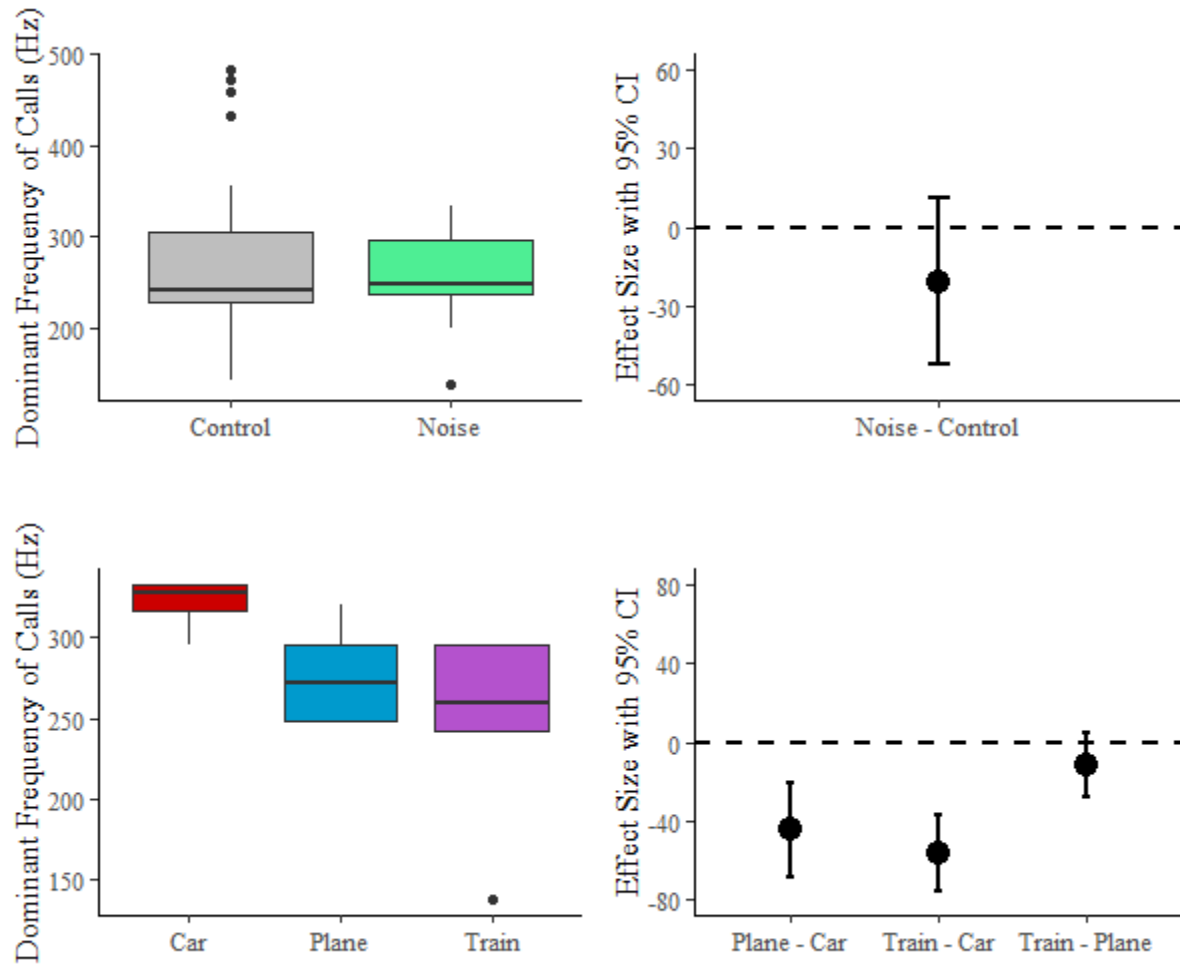

B) Car *vs.* Plane *vs.* Train

Appendix RR. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the northern cricket frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
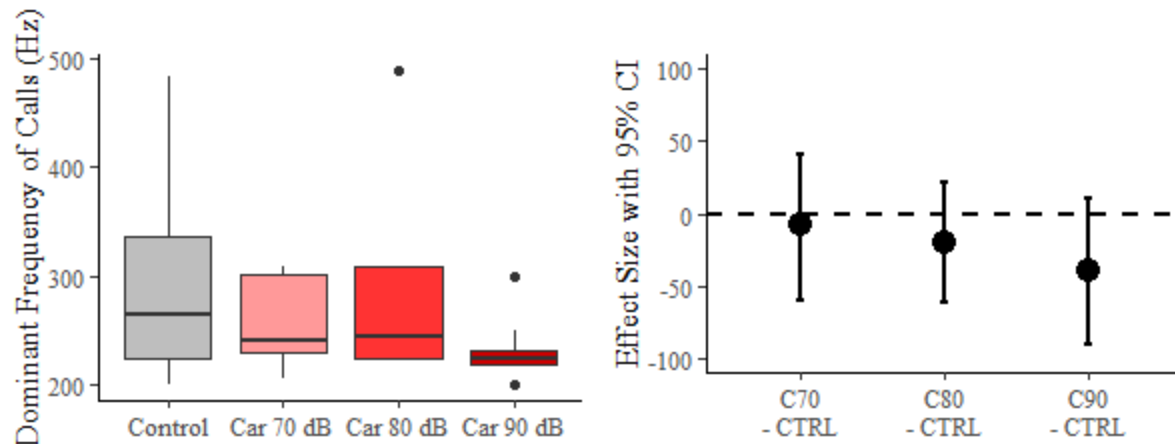
Appendix SS. Effect of intensity of car noise on dominant frequency for the northern cricket frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lmer (dominant frequency ~ treatment + (1|initial conditions)). For more detail of effect size estimates, refer to Appendix OO.

A) Control *vs.* Noise

Appendix TT. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the American bullfrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lmer (dominant frequency ~ treatment + (1|initial conditions)). For more detail of effect size estimates, refer to Appendix A.

Appendix UU. Effect of intensity of car noise on dominant frequency for the American bullfrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lmer (dominant frequency ~ treatment + (1|initial conditions)). For more detail of effect size estimates, refer to Appendix A.
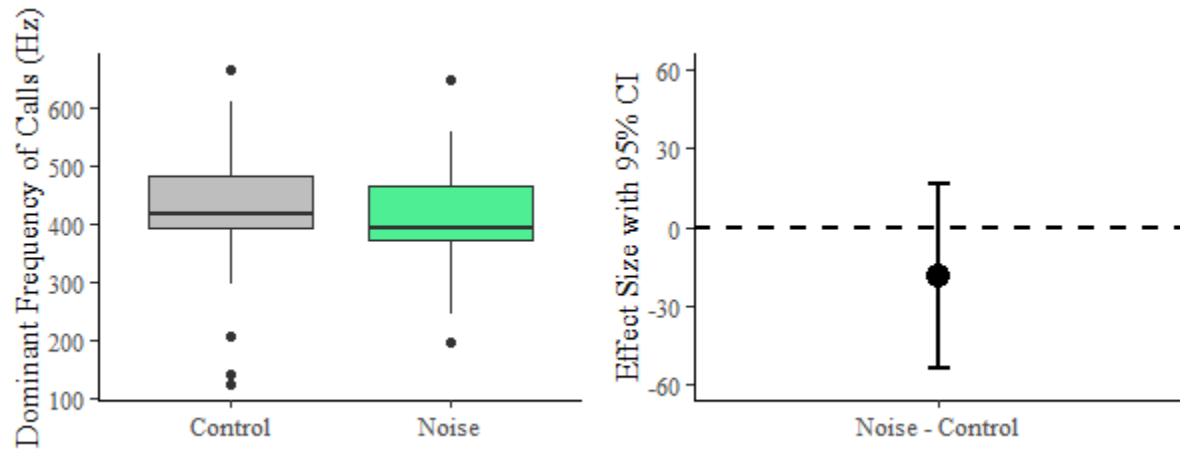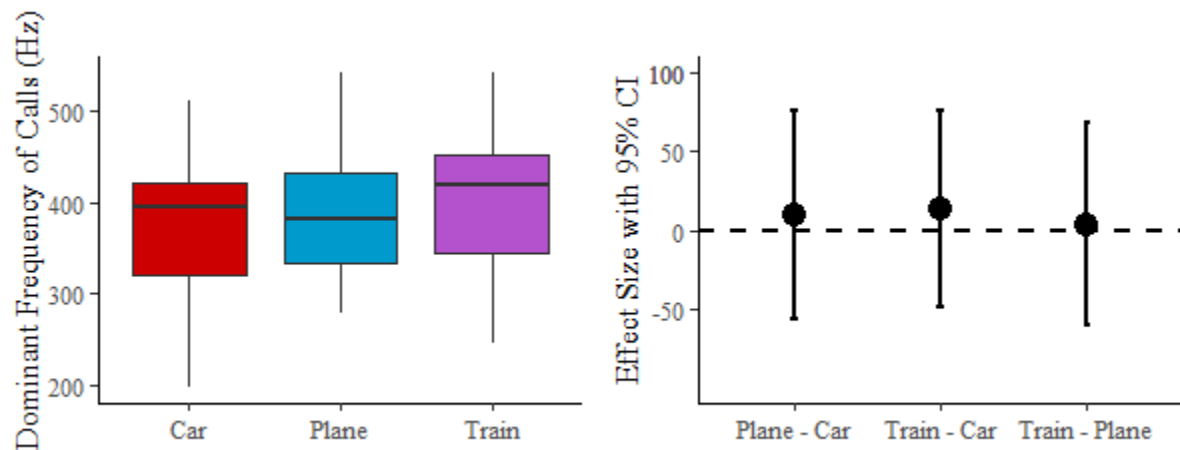
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix VV. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the green frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.

Appendix WW. Effect of intensity of car noise on dominant frequency for the green frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
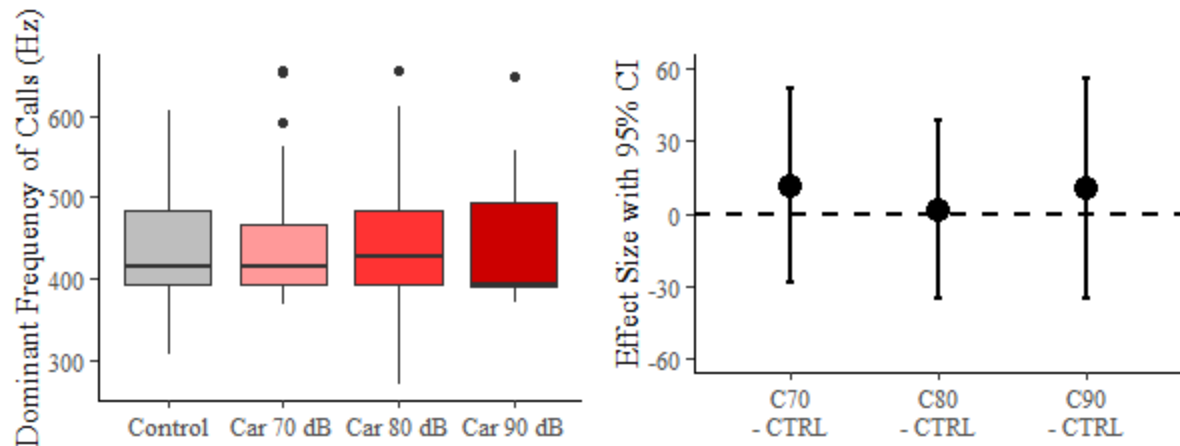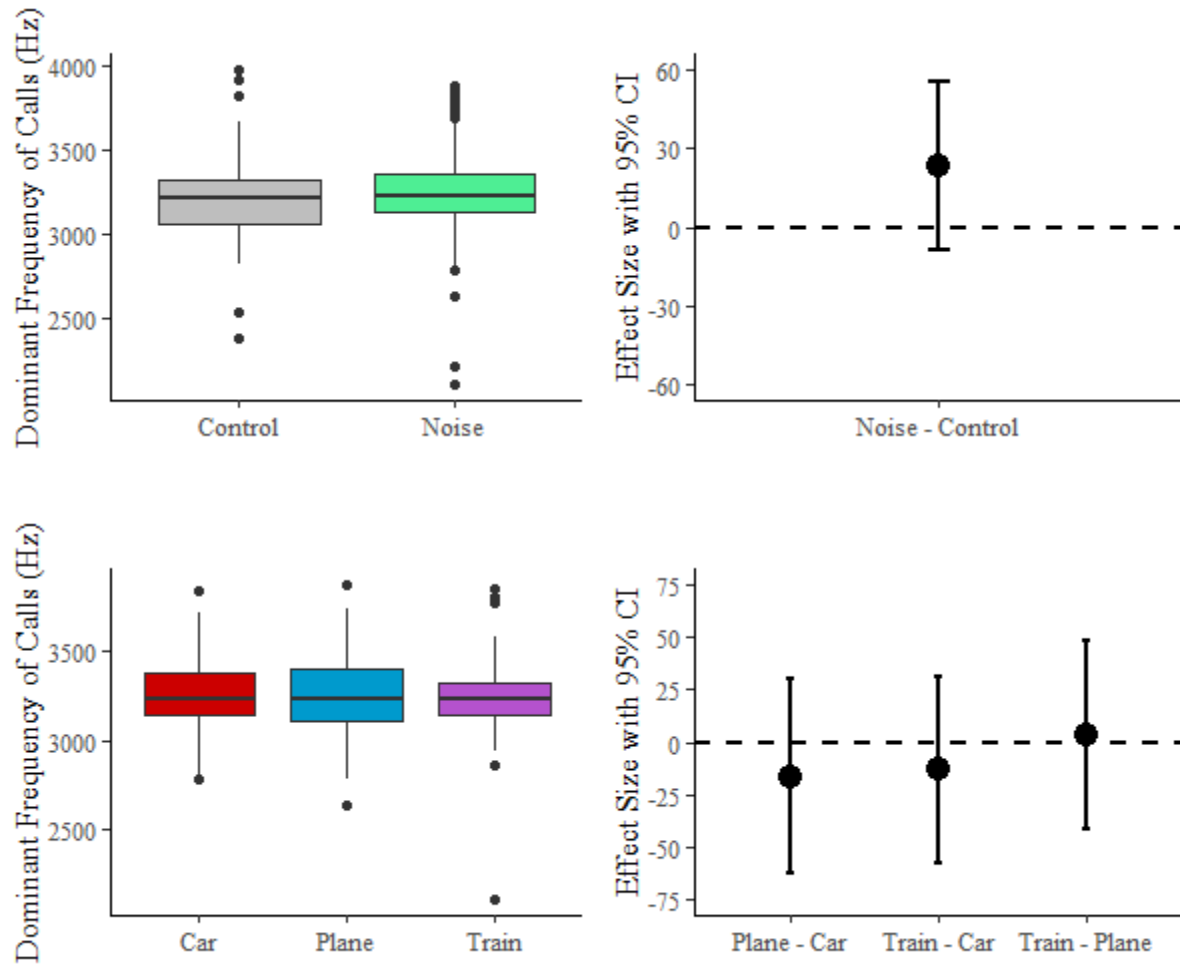
A) Control *vs.* Noise

Appendix XX. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the western chorus frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment + (1|initial conditions)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment + (1|initial conditions)).For more detail of effect size estimates, refer to Appendix A.
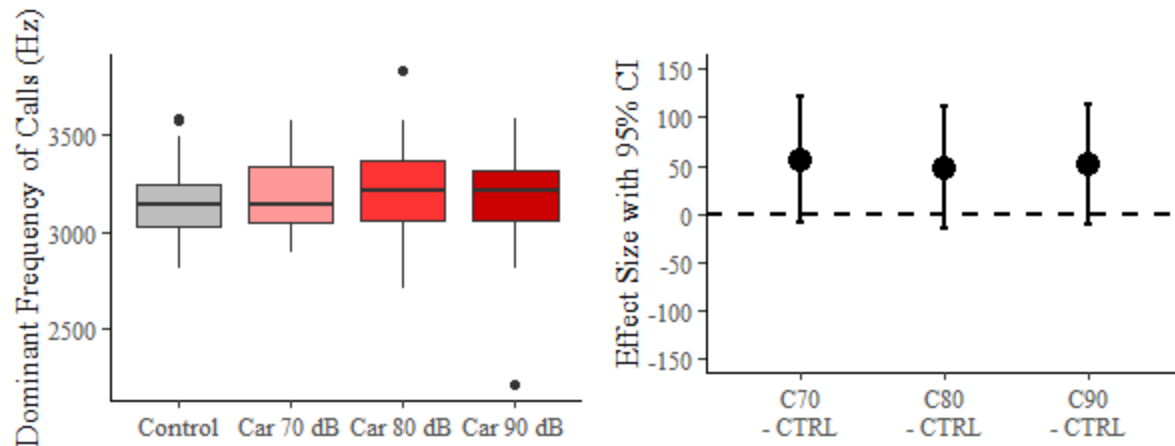
Appendix YY. Effect of intensity of car noise on dominant frequency for the western chorus frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lmer (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.

A) Control *vs.* Noise
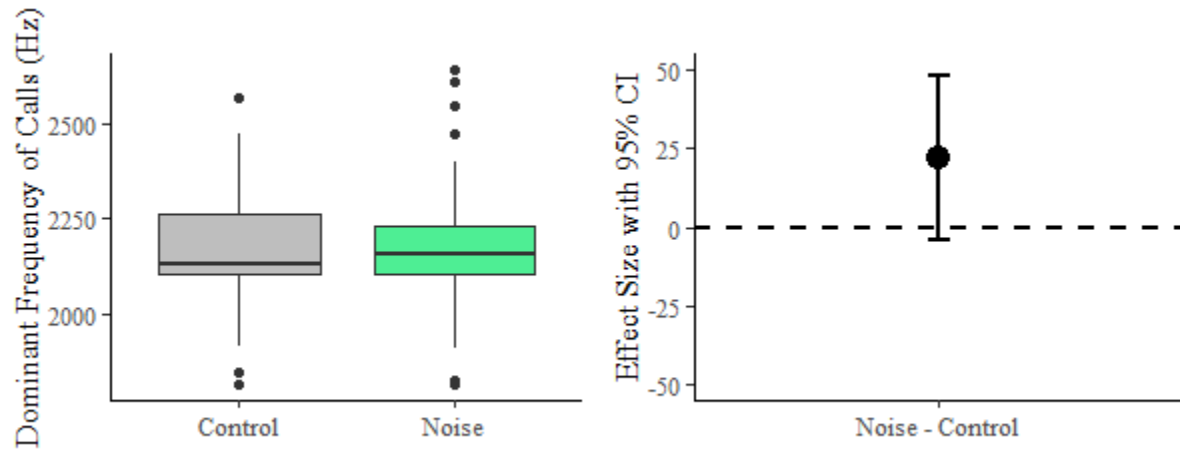

B) Car *vs.* Plane *vs.* Train

Appendix ZZ. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the eastern gray treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment + (1|initial conditions)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment).For more detail of effect size estimates, refer to Appendix A.
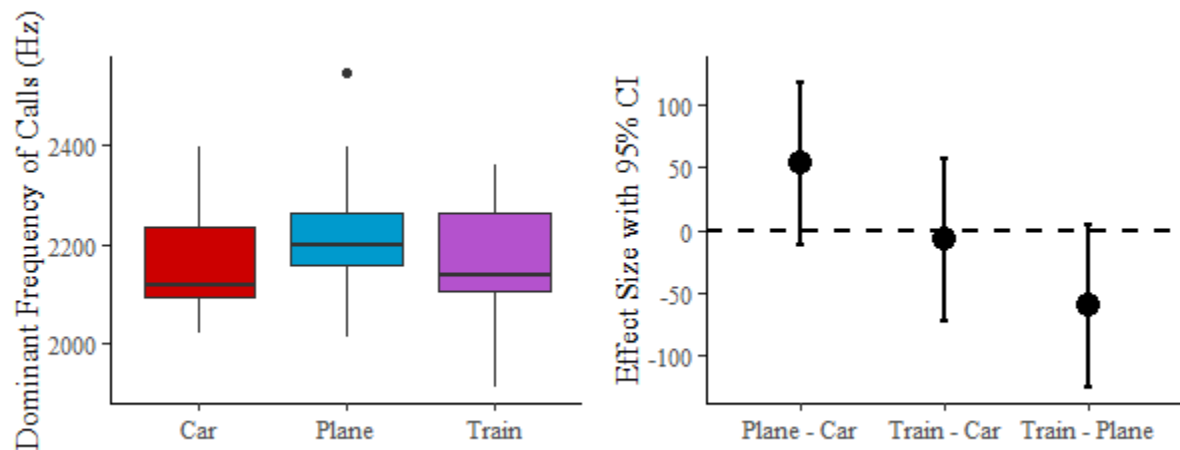
Appendix AAA. Effect of intensity of car noise on dominant frequency for the eastern gray treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lmer (dominant frequency ~ treatment + (1|block)). For more detail of effect size estimates, refer to Appendix A.

A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix BBB. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the wood frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment + offset(initial conditions)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment).For more detail of effect size estimates, refer to Appendix A.

Appendix CCC. Effect of intensity of car noise on dominant frequency for the wood frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
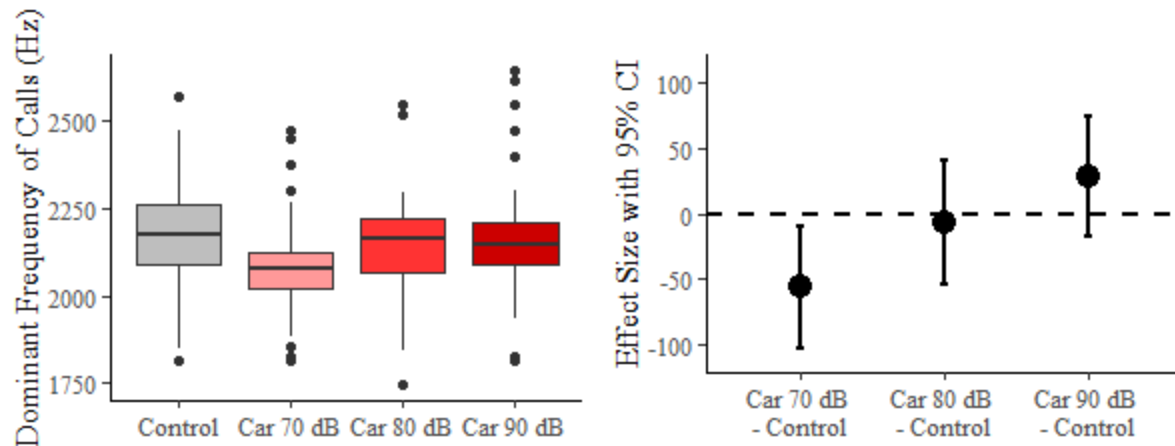
A) Control *vs.* Noise
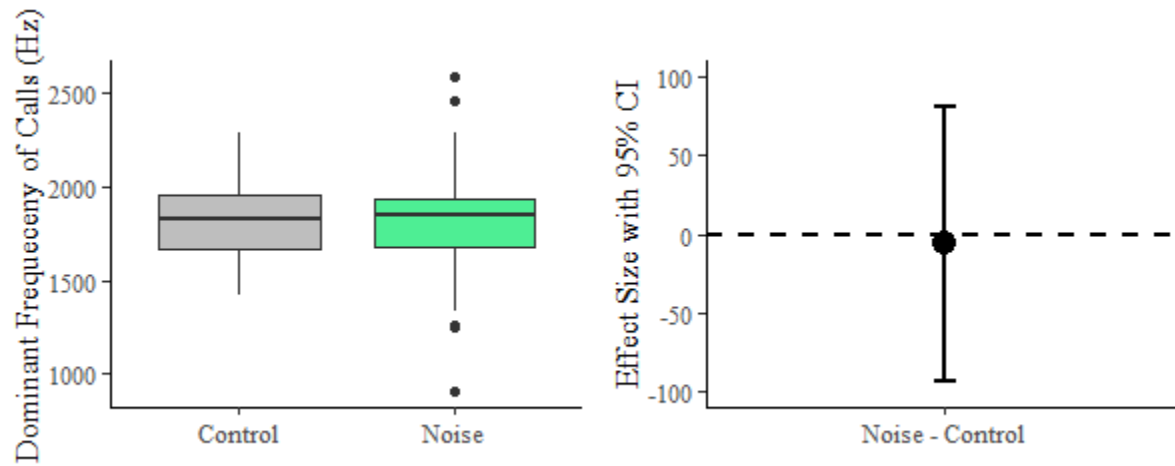

Appendix DDD. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the spring peeper. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment).For more detail of effect size estimates, refer to Appendix A.
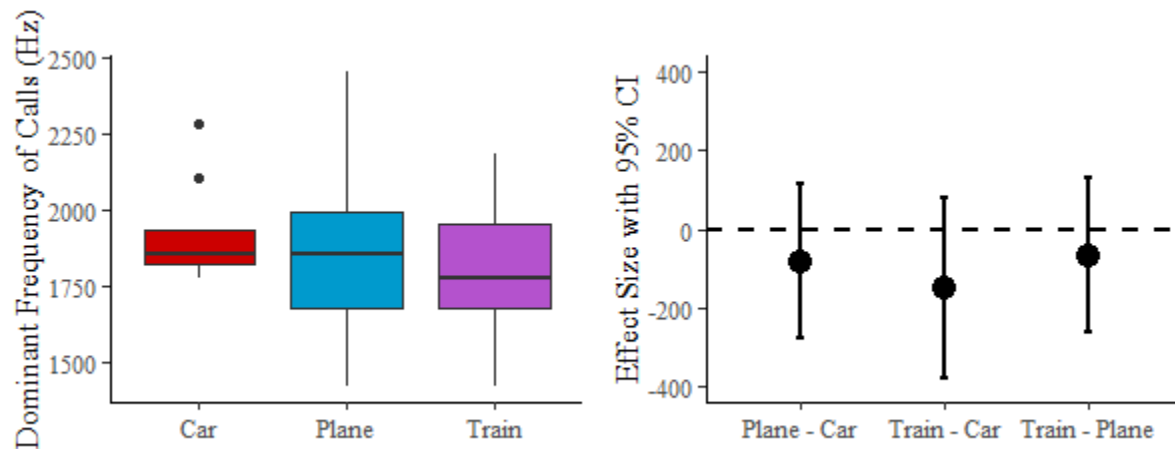
Appendix EEE. Effect of intensity of car noise on dominant frequency for the spring peeper. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
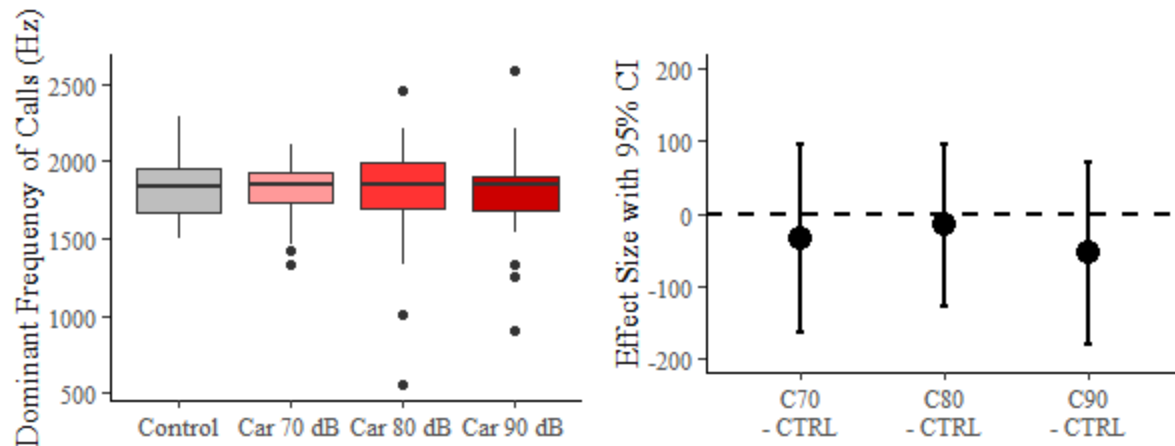
A) Control *vs.* Noise
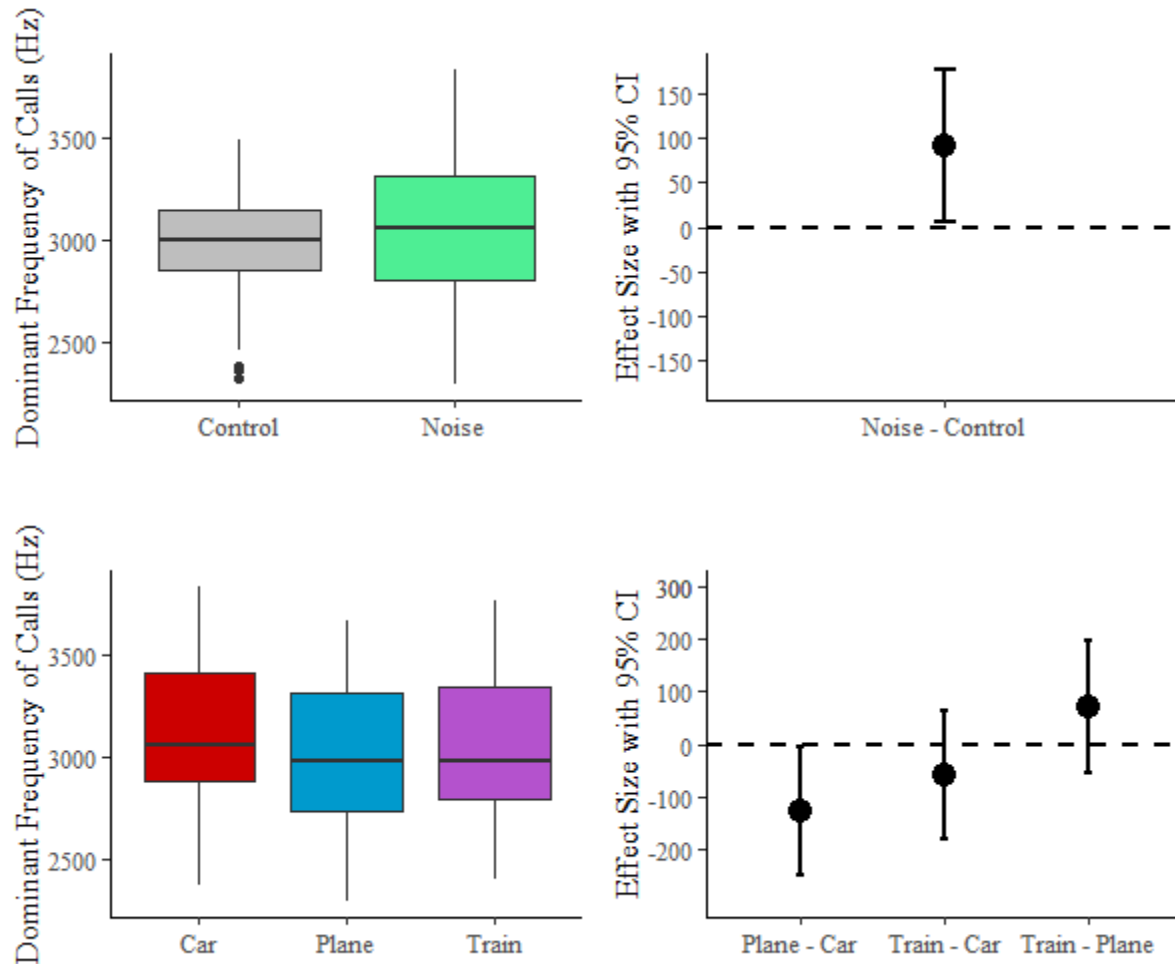

B) Car *vs.* Plane *vs.* Train
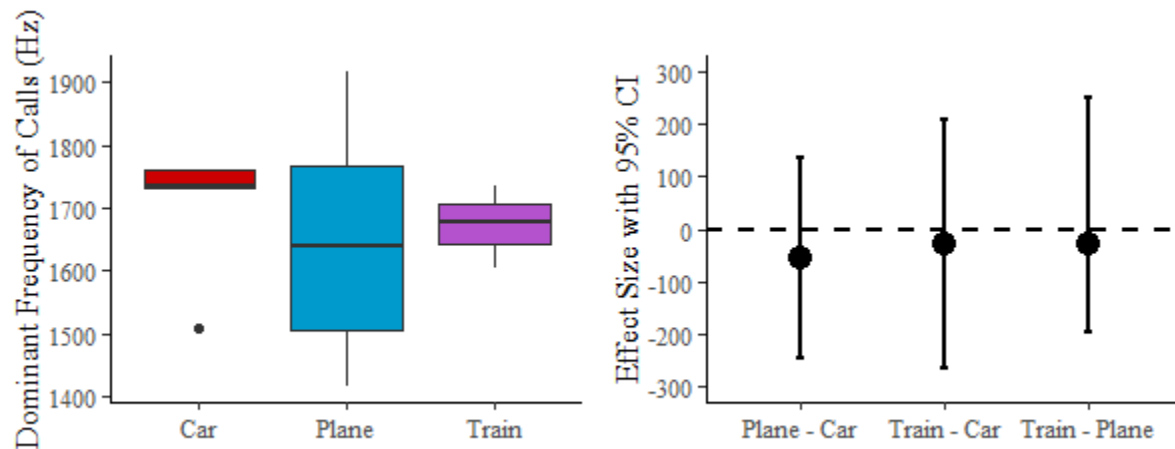
Appendix FFF. Impact of three types of noise (car, plane, and train) at 90 dB on dominant frequency for the American toad. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (dominant frequency~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (dominant frequency ~ treatment).For more detail of effect size estimates, refer to Appendix A.
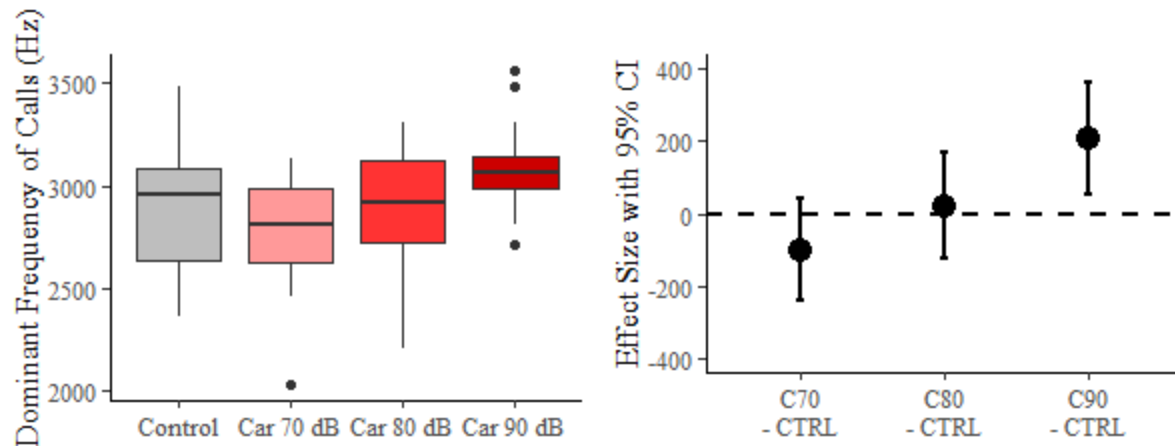
Appendix GGG. Effect of intensity of car noise on dominant frequency for American toad. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (dominant frequency ~ treatment). For more detail of effect size estimates, refer to Appendix A.
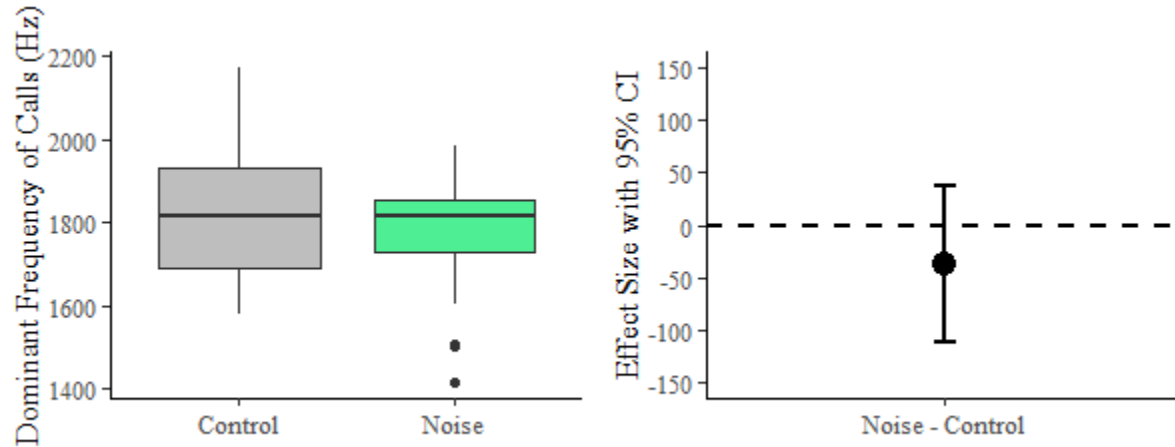
A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train
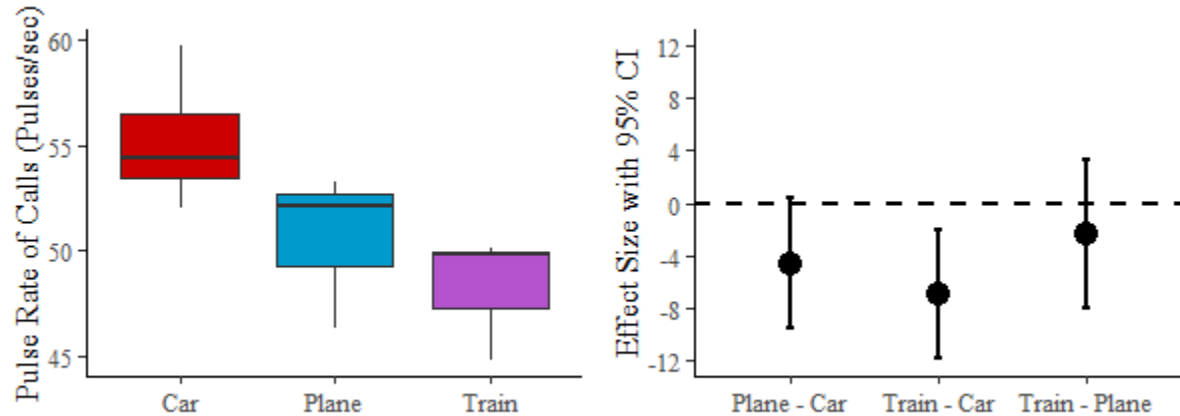
Appendix HHH. Impact of three types of noise (car, plane, and train) at 90 dB on pulse rate for the Cope's gray treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (pulse rate~ treatment + (1|block)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (pulse rate ~ treatment).For more detail of effect size estimates, refer to Appendix A.

Appendix III. Effect of intensity of car noise on pulse rate for Cope's gray treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (pulse rate ~ treatment). For more detail of effect size estimates, refer to Appendix A.

A) Control *vs.* Noise

B) Car *vs.* Plane *vs.* Train

Appendix JJJ. Impact of three types of noise (car, plane, and train) at 90 dB on pulse rate for the northern cricket frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (pulse rate~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lmer (pulse rate ~ treatment + (1|block)).For more detail of effect size estimates, refer to Appendix A.
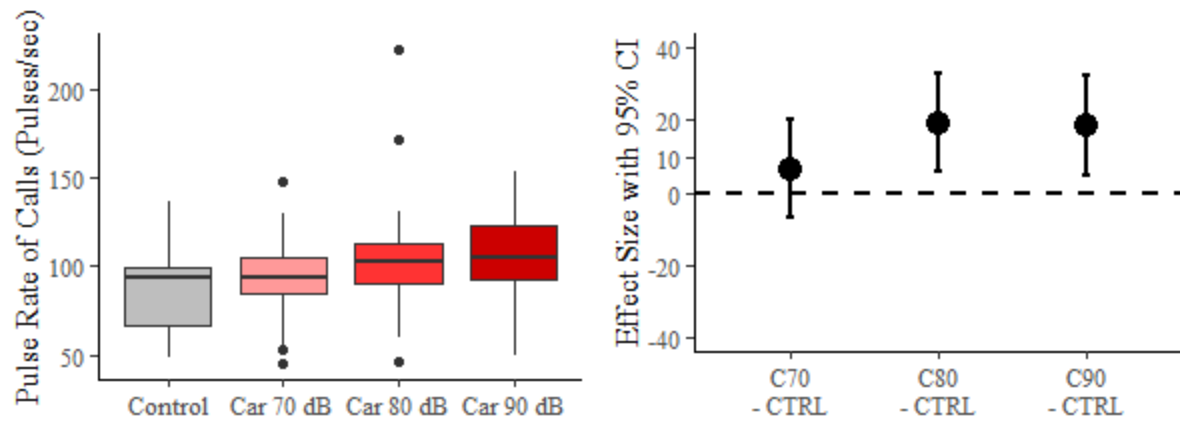
Appendix KKK. Effect of intensity of car noise on pulse rate for northern cricket frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (pulse rate ~ treatment). For more detail of effect size estimates, refer to Appendix A.

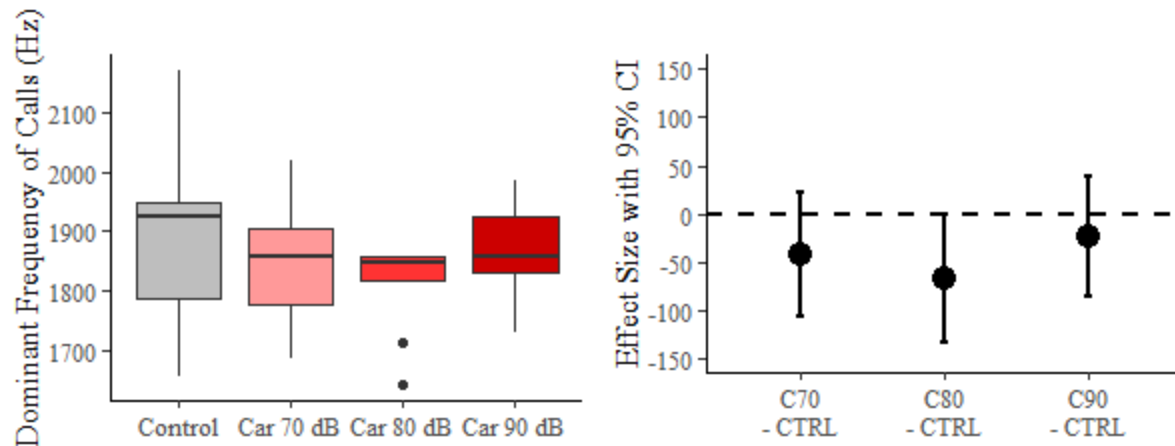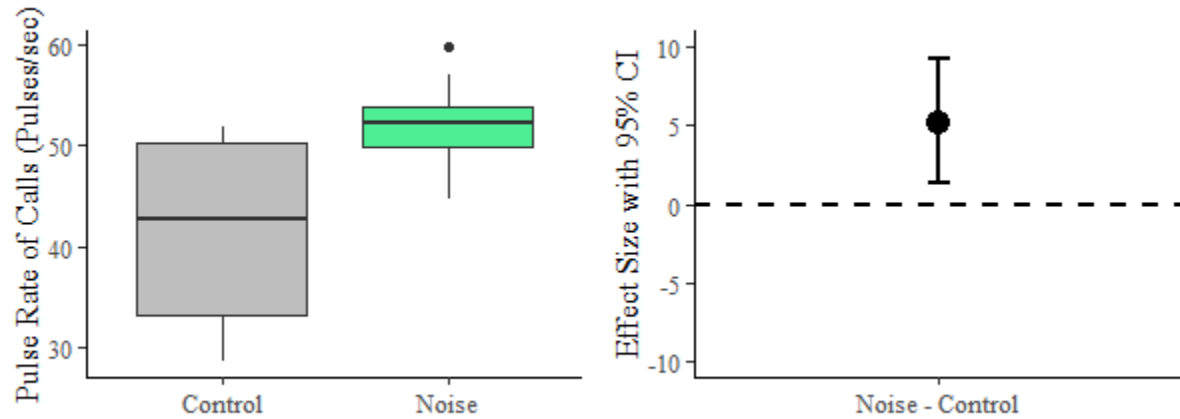A) Control *vs.* Noise
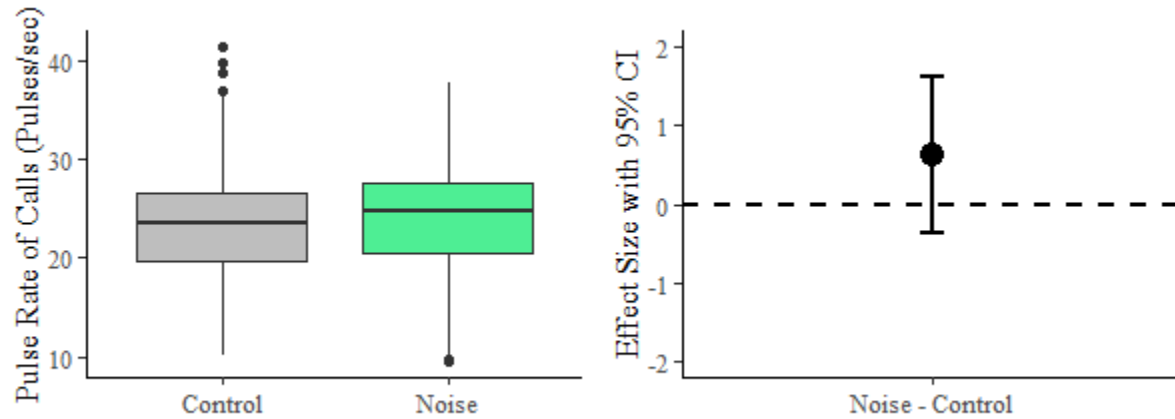
B) Car *vs.* Plane *vs.* Train
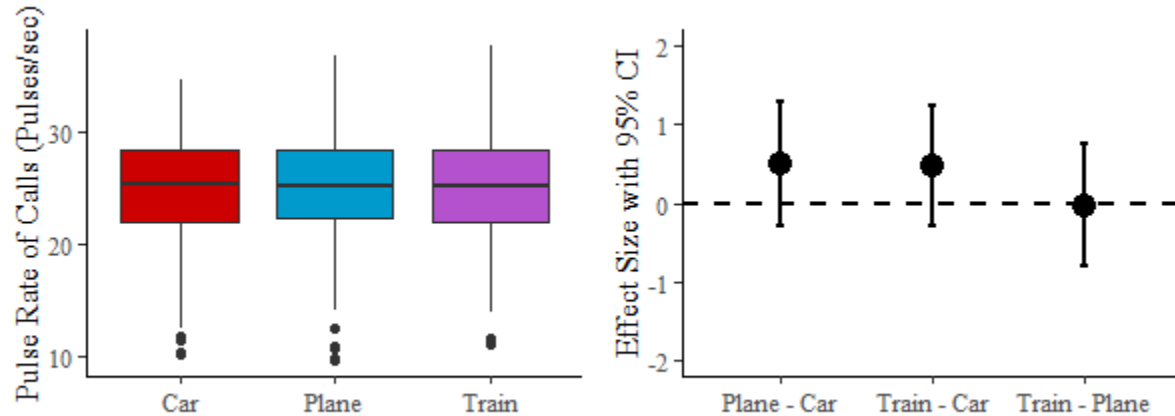
Appendix LLL. Impact of three types of noise (car, plane, and train) at 90 dB on pulse rate for the western chorus frog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lm (pulse rate~ treatment). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lmer (pulse rate ~ treatment + (1|block)).For more detail of effect size estimates, refer to Appendix A.
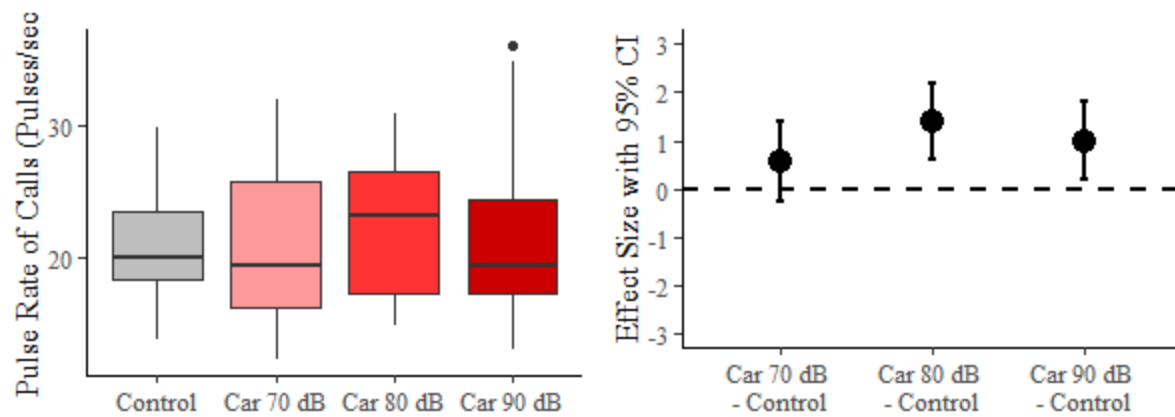
Appendix MMM. Effect of intensity of car noise on pulse rate for western chorus frog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (pulse rate ~ treatment + (1|block)). For more detail of effect size estimates, refer to Appendix A.
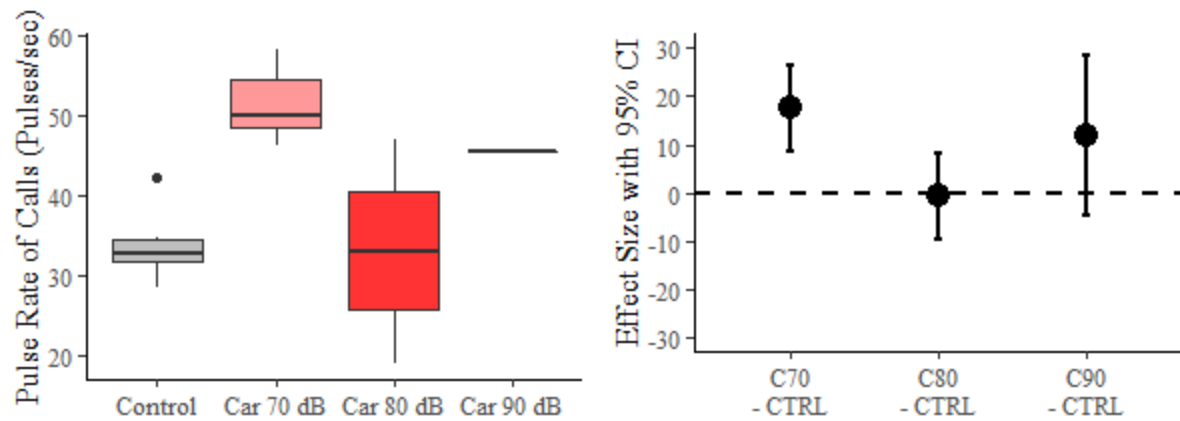
A) Control *vs.* Noise
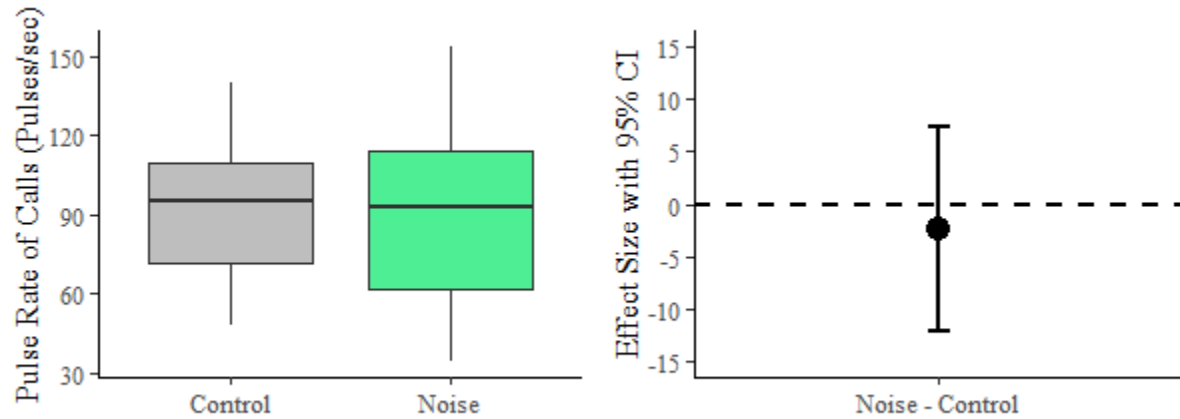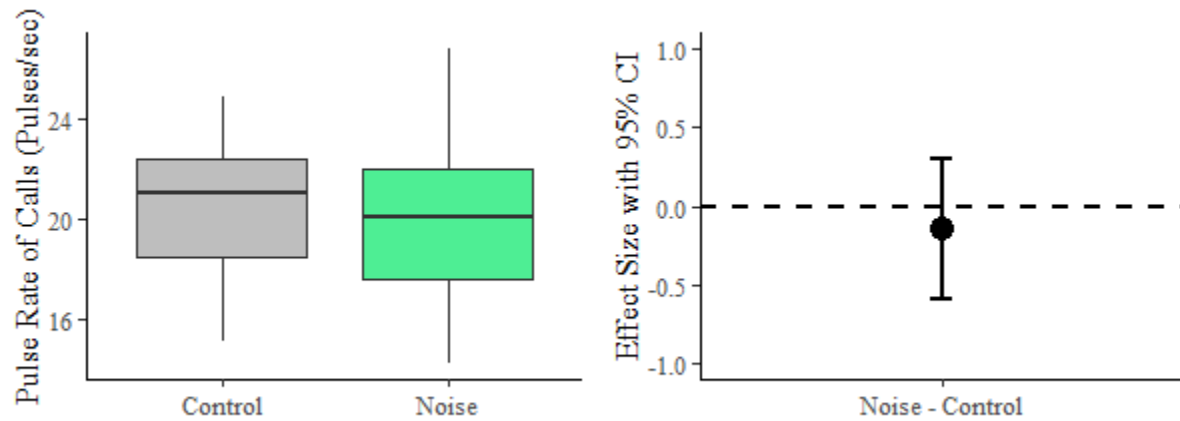

B) Car *vs.* Plane *vs.* Train
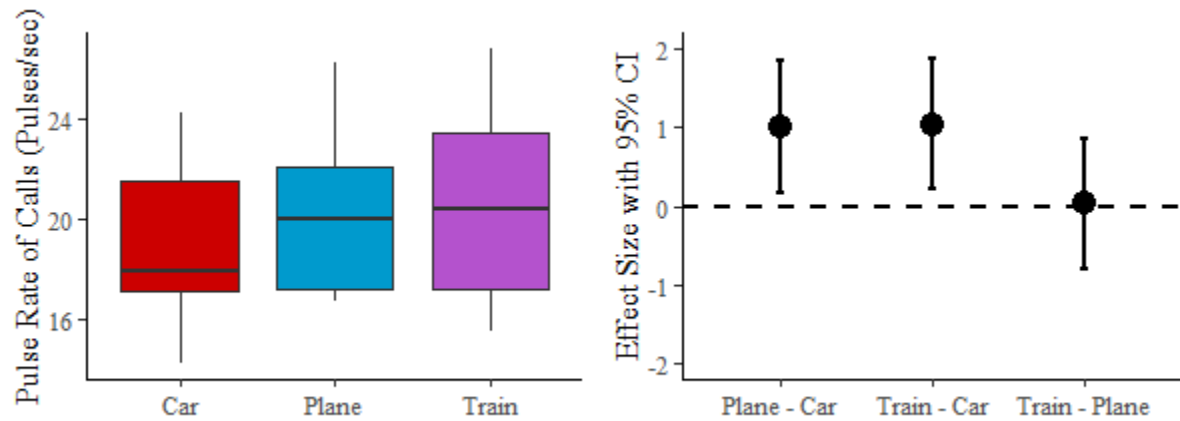
Appendix NNN. Impact of three types of noise (car, plane, and train) at 90 dB on pulse rate for the eastern gray treefrog. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lmer (pulse rate~ treatment + (1|block)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lmer (pulse rate ~ treatment + (1|block)).For more detail of effect size estimates, refer to Appendix A.
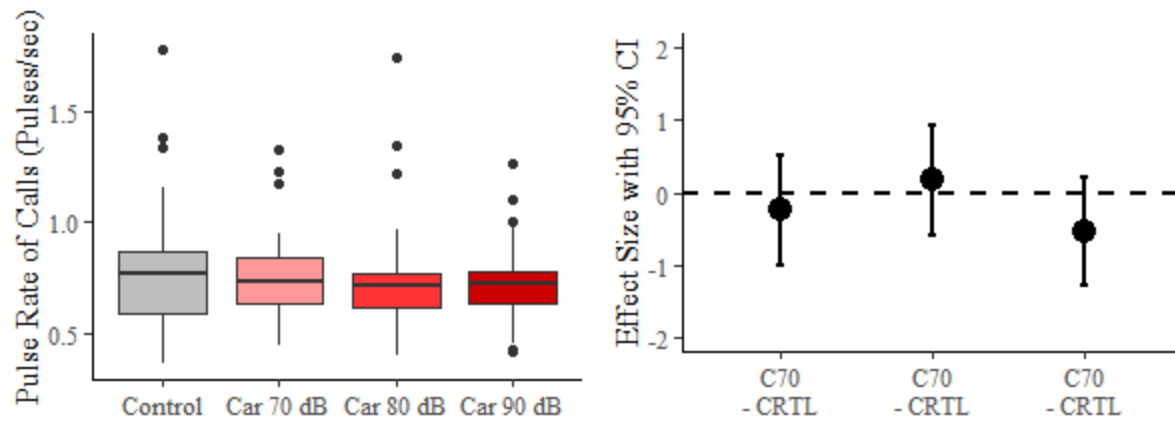
Appendix OOO. Effect of intensity of car noise on pulse rate for the eastern gray treefrog. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (pulse rate ~ treatment). For more detail of effect size estimates, refer to Appendix A.

A) Control *vs.* Noise
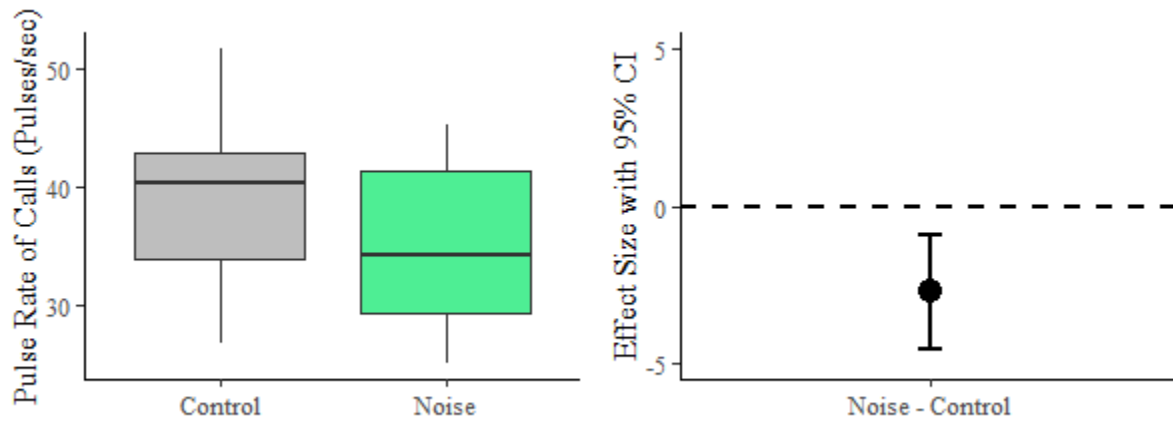
B) Car *vs.* Plane *vs.* Train
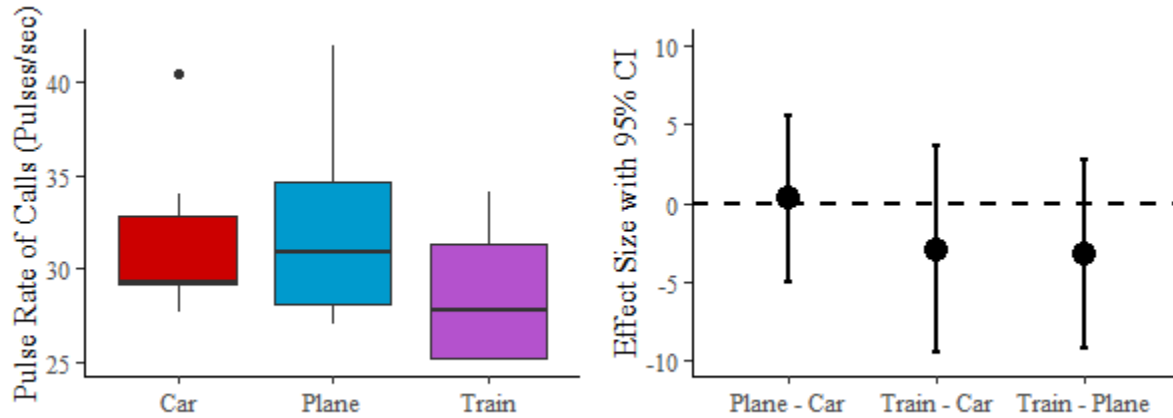
Appendix PPP. Impact of three types of noise (car, plane, and train) at 90 dB on pulse rate for the American toad. All box plots represent call counts from a single sample, separated by treatment. (A) Impact all three types of anthropogenic noise pooled. Effect size estimates are from the model lmer (pulse rate~ treatment + (1|block)). (B) Comparison of the three types of anthropogenic noise. Effect size estimates are from the lm (pulse rate ~ treatment).For more detail of effect size estimates, refer to Appendix A.
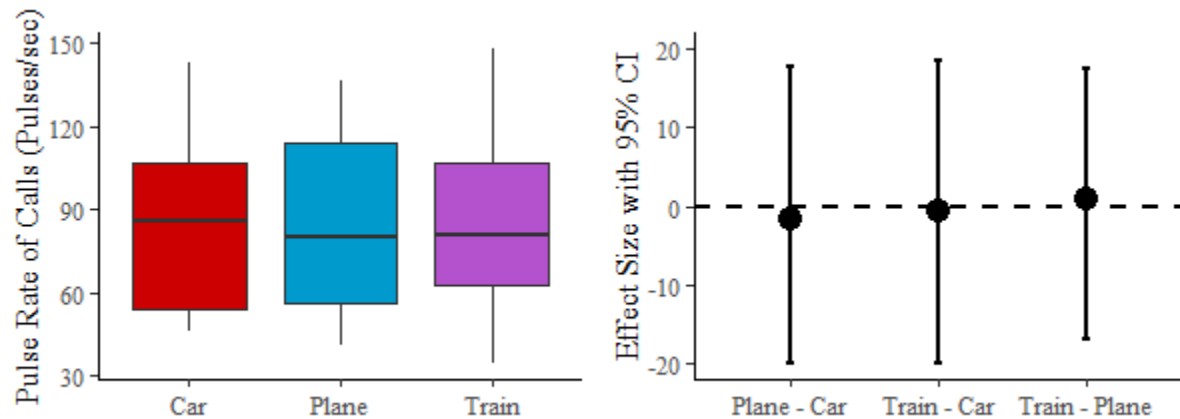
Appendix QQQ. Effect of intensity of car noise on pulse rate for the American toad. Noise power levels were at 70, 80, and 90 dB. All box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (pulse rate ~ treatment + (1|block). For more detail of effect size estimates, refer to Appendix A.

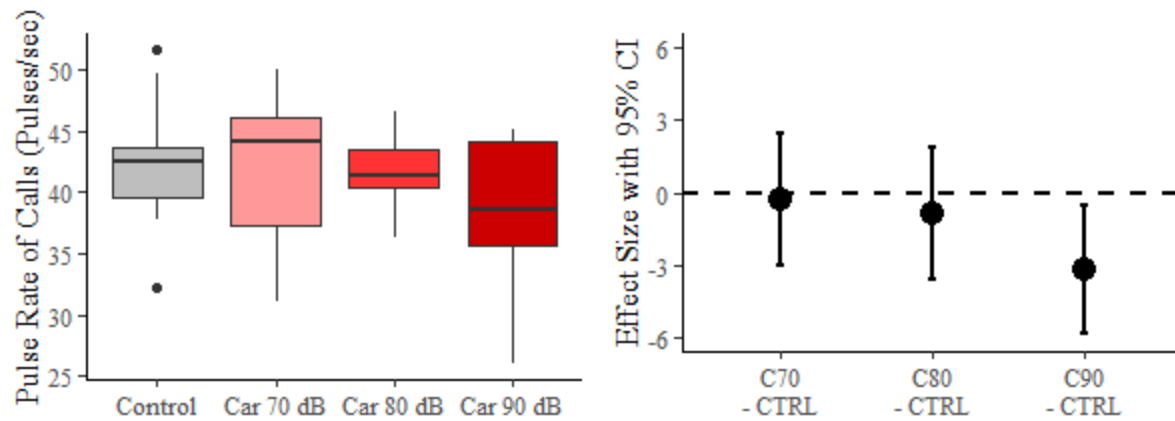| exp.id	initials	date	prese	place	n.coord	w.coord	exp.t	rec#	sp	grp1.genus	grp2.seas	grp3.habitat	grp4.brler	heter	aci	ndsi	period	trt	sample	trt_sar	initial	calls	call in	noise
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	1	i	1	i_1	75	37	2	na
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	1	i	2	i_2	75	38	2	na
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	2	t	1	t_1	75	58	2	yes
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	2	t	2	t_2	75	2	1	yes
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	3	e	1	e_1	75	0	0	no
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	3	e	2	e_2	75	47	0	no
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	4	p	1	p_1	75	0	0	yes
1	nb	04.07.2017	org	tpm	41.58163	87.8619	cpt	1	wcf	pseudacris	early	temporary	long	1	425.447	-0.603	4	p	2	p_2	75	0	0	yes

Appendix RRR. Data frame for call rate. A partial view of the raw data collected from the playback experiments. The full dataset is available online through the Center for Open Science data repository at https://osf.io/xwzjk/ under the file name "DF1pop.csv".

exp.id	exp.type	sp	grp1.genu	grp2.seas	grp3.habi	grp4.brler	hetero	aci	ndsi	period	trt	sample	sample_tr	initial.call	length	p.num	domfreq	p.rate	noise
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	1	i	2	i_2	75	0.567	13	3229.93	22.92769	no
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	1	i	2	i_2	75	0.686	12	3229.93	17.49271	no
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	1	i	1	i_1	75	0.762	17	3238.75	22.30971	no
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	1	i	1	i_1	75	0.782	14	3321.8	17.90281	no
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	1	i	1	i_1	75	0.814	14	3321.8	17.19902	no
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	1	i	2	i_2	75	0.65	12	3331.66	18.46154	no
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	2	t	2	t_2	75	0.176	5	2860.58	28.40909	yes
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	2	t	2	t_2	75	0.388	7	2952.43	18.04124	yes
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	2	t	1	t_1	75	0.8	15	3229.93	18.75	yes
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	2	t	1	t_1	75	0.593	10	3230.95	16.86341	yes
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	2	t	1	t_1	75	0.779	15	3306.23	19.25546	yes
1	cpt	wcf	pseudacri	early	temporan	long	1	425.447	-0.60325	3	e	2	e_2	75	0.653	16	3050.11	24.5023	no

Appendix SSS. Data frame for call length, dominant frequency, and pulse rate. A partial view of the raw data collected from the playback experiments. The full dataset is available online through the Center for Open Science data repository at https://osf.io/xwzjk/ under the file name "DF2pop.csv".

exp.id	date	n.coord	w.coord	exp.ty	rec#	sp	grp1.genu	grp2.seas	grp3.habit	grp4.brler	heter	aci	ndsi	period	trt	noise	sample	trt_sam	initial.	calls	length	domfreq	p.rate
1	04.07.2017	41.58163	87.8619	cpt	1	wcf	pseudacri	early	temporary	long	1	425.447	-0.60325	1	i	i	1	i_1	75	75	0.710167	3278.978	19.38225
1	04.07.2017	41.58163	87.8619	cpt	1	wcf	pseudacri	early	temporary	long	1	425.447	-0.60325	2	t	yes	1	t_1	75	60	0.5472	3116.024	20.26384
1	04.07.2017	41.58163	87.8619	cpt	1	wcf	pseudacri	early	temporary	long	1	425.447	-0.60325	3	e	e	1	e_1	75	47	0.663667	3271.15	22.14875
1	04.07.2017	41.58163	87.8619	cpt	1	wcf	pseudacri	early	temporary	long	1	425.447	-0.60325	4	p	yes	1	p_1	75	0	na	na	na
1	04.07.2017	41.58163	87.8619	cpt	1	wcf	pseudacri	early	temporary	long	1	425.447	-0.60325	5	c	yes	1	c_1	75	1	0.1	3062.17	30
1	04.07.2017	41.58163	87.8619	cpt	1	wcf	pseudacri	early	temporary	long	1	425.447	-0.60325	6	f	f	1	f_1	75	101	0.4775	3247.022	26.49728
2	04.07.2017	41.58163	87.8619	cpt	1	nlf	lithobates	middle	both	medium	1	425.447	-0.60325	1	i	i	1	i_1	2	2	2.7375	na	16.24796
2	04.07.2017	41.58163	87.8619	cpt	1	nlf	lithobates	middle	both	medium	1	425.447	-0.60325	2	t	yes	1	t_1	2	0	na	na	na
2	04.07.2017	41.58163	87.8619	cpt	1	nlf	lithobates	middle	both	medium	1	425.447	-0.60325	3	e	e	1	e_1	2	0	na	na	na
2	04.07.2017	41.58163	87.8619	cpt	1	nlf	lithobates	middle	both	medium	1	425.447	-0.60325	4	p	yes	1	p_1	2	0	na	na	na
2	04.07.2017	41.58163	87.8619	cpt	1	nlf	lithobates	middle	both	medium	1	425.447	-0.60325	5	c	yes	1	c_1	2	0	na	na	na
2	04.07.2017	41.58163	87.8619	cpt	1	nlf	lithobates	middle	both	medium	1	425.447	-0.60325	6	f	f	1	f_1	2	12	1.721333	855.96	19.50946

Appendix TTT. Data frame for the multivariate data. A partial view of the raw data collected from the playback experiments. The full dataset is available online through the Center for Open Science data repository at https://osf.io/xwzjk/ under the file name "MVpop.csv".

```
####BF, noise vs no noise models####
mydata<-filter(noisedf, noisedf$sp=="bf")
par(mfrow=c(1,1))
hist(mydata$call.effort,breaks = 10)

#hurdle
modelh1<-hurdle(mydata$call.effort~mydata$noise, dist = "negbin")
summary(modelh1)
rootogram(modelh1, max = 50)
qqrplot(modelh1)
plot(modelh1$fitted.values,modelh1$residuals)
#good
```

Appendix UUU. Data frame for the multivariate data. An example of the R code used for data analysis. The full R code is available online through the Center for Open Science data repository at https://osf.io/xwzjk/ under the file name "mvAnalysis" for the multivariate analysis, "modelNumCalls" for the univariate analysis on the number of calls counted, "modelDomFreq" for the dominant frequency univariate analysis, "modelLength" for the call length univariate analysis, "modelPulseRate" for the pulse rate univariate analysis.

CHAPTER 3: ANTHROPOGENIC NOISE PLAYBACK EXPERIMENTS ON URBAN BULLFROG AND

GREEN FROG INDIVIDUALS

3.1 ABSTRACT

Noise playback experiments are often used to investigate changes in frog calling behavior

in response to anthropogenic noise. Studies can either record groups or individuals. Given that

the results from previous research are inconsistent, comparisons between different

methodologies may be helpful in interpreting patterns across studies. In this study, we conducted

noise playback experiments on American bullfrog (*Lithobates catesbeianus*) and green frog

(*Lithobates clamitans*) individuals at urban ponds in order to compare changes in calling

behavior to our previous study conducted on the same species at the population-level. Noise

treatments included car traffic at three magnitudes (peaks of 70, 80 and 90 dB). Our major

finding was a clear difference in changes to call length between the two studies. We found that

individuals from both species shortened their calls during the 80 and 90 dB treatments. This was

not seen for either species in the population-level study. We did not find major differences

between studies for changes in call rate or call length. We found evidence for the Lombard

Effect (louder calling) in the bullfrogs in response to louder car noise. For the green frog, there

was a slight positive trend between calling power and treatment power, but differences were well

within 95% confidence intervals. We did not measure call power in our previous study. This

comparison between methodologies revealed that changes in some call attributes seen in

individuals may not be translatable to a chorusing group, as aggregate patterns seen in

population-level studies may not match the varied responses from individuals, especially those

that reduce their call rate. Ideally, researchers should utilize both chorus-level and individual-

level methods. Population-level studies better capture the chaos that females will encounter when

making mate choices, whereas individual-level studies can identify the changes made by males in an attempt stand out amidst the chaos.

Key Words: acoustic species; anthropogenic noise; Anura; bioacoustics; call attributes; field playback experiments; individual; soundscape

3.2 INTRODUCTION

3.2.1 Approaches to anuran soundscape research

Research exploring the effect of soundscape on frog behavior has used various analytical approaches and environmental settings. Some have been laboratory studies, usually focusing on individuals. These have been conducted to answer specific questions about mechanisms that influence changes in male calls (Love & Bee 2010; Halfwerk *et al.* 2016), or mechanisms that dictate female preferences of male calls (Luddecke 2002; Reichert *et al.* 2016). Other projects have been conducted in the field, and data for these studies can either be collected on an individual basis (Sun & Narins 2005; Kaiser & Hammers 2009; Bleach *et al.* 2015; Zhu *et al.* 2017), or at the population level (Parris et al 2009; Vargas-Salinas & Amezquita 2013; Vargas-Salinas *et al.* 2014).

There are potential complications when interpreting results from this mixture of individual and community-level playback experiments. For instance, for most of the breeding season, females are forced to make choices from multiple nearby individuals in the community, while there are still many other males in the background producing calls (Ryan 1980; Gerhardt 1994; Phillips *et al.* 1999). This process is different than evaluating a call from a single male in an isolated arena with a controlled acoustic and spatial environment. In large choruses, female choice must be made in a chaotic environment which may reduce or change optimal behavior

strategies (Stratman & Hobel 2019). Thus, results from population-level studies are not necessarily interchangeable with individual-level studies. There are advantages and disadvantages to both approaches. Individual-level studies can investigate individual variation within males, whereas data from population-level studies can capture the chaotic setting in which female choice and mating occur amidst a chorus.

3.2.2 Our study

In order to determine if behavioral responses to noise differ when monitoring single individuals compared to full choruses, comparative studies should be made with the same species, locations and methods. However, playback studies have chosen one level to test at (population or individual) to make their conclusions. This study investigates individual responses of male bullfrogs (*Lithobates catesbeianus*) green frogs (*Lithobates clamitans*) to noise playback and is a follow-up to a previous playback study that measured population-wide responses from ponds (Bielinski Dissertation Chapter 2). The purpose of this study is to compare results between these two approaches to explore whether responses to noise change at the individual-level versus the population-level.

3.2.3 Predictions

Comparisons of individuals versus populations

Call attribute data from population-level studies should include proportionally more data from individuals with higher call rates. Thus, aggregate responses from population-level studies may not fully capture individual variation. We predict that patterns of response variables will show more effects and higher magnitudes of those effects compared to the previous study.

Call rate

Call rate is a dynamic property that experiences high selection pressure from females (Gerhardt 1991). From the female prospective, it is one of the more distinguishable properties of a call (Stratman & Hobel 2019). In accordance with the predictions and results of the population-level study, individual males should still reduce their call rate, in order to avoid wasting energy when calling in unfavorable conditions for females to assess male suitors. Due to the intermittent nature of anthropogenic noise, frogs should be willing to temporarily forgo calling in the present, for a future time with more favorable soundscape conditions. Therefore, our prediction is that call rates will be reduced.

Dominant frequency

Another method for males to cope with noise is to avoid spectral interference by shifting their call frequency away from the noise frequency range. Documented frequency shifts in response to noise have been in both directions (up: Kruger and Du Preez 2016; Shen & Xu 2016; or down: Hanna *et al.* 2014; Caorsi *et al.* 2017). If only less-active callers alter their dominant frequency, we should be able to detect a response in this study. However, frequency shifts in general were not noticeable for the bullfrog and the green frog at the population-level. This may be because they lack the plasticity to completely shift their call out of the range of anthropogenic noise. We expect to see the same non-response for call frequency in this study.

Call length

Our predicted changes in call length mirrors that of call rate, as they are both dynamic properties that are directionally selected for by females but are limited by energy. Therefore, we expect call length to be reduced, especially if this is a strategy implemented by less-active callers.

Threshold effect

For individuals that do react to the anthropogenic noise treatments, the changes in attributes should be more extreme in louder noise treatments. Traffic noise at 90 dB should elicit greater call changes than 80 dB or 70 dB. It will be more difficult for females to accurately interpret mating calls during the noisier treatments, so it would benefit the males to stand out more by compensating through whatever mechanism they implement, whether it be through call rate, frequency, or length. If there are any changes, the magnitude of the change will be greatest during the loudest treatment.

Call power

Boosting call power is very energetically taxing, so males do not call at their maximum amplitude even though females prefer louder calls (Halfwerk *et al.* 2016). One common prediction is that males should respond to noise by increasing their call power so their calls can be heard by females over the noise, known as the Lombard Effect. Previous studies have tested for the Lombard effect, with differing responses (for: Halfwerk *et al.* 2016, Yi & Sheridan 2019; against: Love & Bee 2010). Regardless of how males interpret the projected anthropogenic noise (as either biologically irrelevant noise or as an upswing in competitive calling), when males *do* decide to vocalize, calls should be louder during the noisiest treatments. It would be a waste of energy to call too quietly and have the signal be washed out by noise. Therefore, calls should be the loudest for the 90 dB treatment, and the amplitude should diminish for the 80 dB, 70 dB, and finally the control treatment. Call power will be interpreted separately because we could not analyze call power in the previous study since the distance from each frog to the recorder was not controlled.

3.3 MATERIALS AND METHODS

184

3.3.1 Study sites.

We conducted this study at 5 frog breeding sites around the south side of the Chicago. Four sites were in forest preserves and one site was a Chicago park. These sites were originally surveyed for a large-scale population-level playback study. There were two species involved in this study, the American bullfrog (n=6) and the green frog (n=5), both of which were included in the first study. These two species were chosen for this type of study because they are territorial. This meant that individuals were not too close to each other to interfere with the test subject, and that the test subject remained stationary for the duration of the experiment. We attempted to include other species in this study, but they were either too mobile (*Hyla chrysoscelis*) or called too closely to competitors (*Acris crepitans*) to responsibly collect individual-level data.

3.3.2 Playback experiments

We conducted the playback experiments from 24 July 2018 to 12 August 2018. Sites were visited during peak calling hours as indicated by previous research (Bielinski Dissertation Chapter 1 and 2) and natural history information from field guides (Phillips *et al.* 1999). Sites were not visited if rain was forecasted, in order to maintain a consistent amount of background noise from geophonies. A recorder (H1 Handy Recorder, Zoom North America, Hauppauge, New York, USA) was placed one meter over the head of the focal male using a tripod. We placed a Bluetooth speaker (Turtle Shell 3.0, Outdoor Tech, Laguna Hills, California, USA) along the pond's edge 10 m away from the recorder.

There was a considerable amount of disturbance when setting up the equipment. However, it was observed that all subjects maintained their location and started calling again, which is consistent with past experiments and observations (Ryan 1980; Gerhardt 1991; Sun & Narins 2005, Kaiser & Hammers 2009). The playback experiment had the same exact structure

185

as the "Noise Power Experiment" from the previous study, which included three treatments of car noise at 70, 80 and 90 dB (Bielinski Dissertation Chapter 2).

3.3.3 Acoustic analysis

We used Song Scope bioacoustics software (version 4.1.5; Wildlife Acoustics, Inc., Concord Massachusetts, USA) to listen to and view each treatment in spectrograms and to count the number of calls produced by each species. Recognizers were used at times to find calls within recordings, but all calls were manually checked (Bielinski Dissertation Chapter 1). Incomplete calls at the beginning or end of the treatment were still counted. After counting the calls, we selected a maximum of 6 calls at random from each sample to calculate the following call attributes: length, dominant frequency, and power. Calls can have multiple levels of spectral shaping, called formants, which result from acoustic resonance of the vocal sac. For the dominant frequency, we gathered data solely on the loudest of the formants.

All call attributes were collected in the previous population-level study (Bielinski Dissertation Chapter 2) besides power, which could only be collected in this study because we controlled for the distance between the individual and the recorder. Changes in all common response variables were compared to the previous study. Patterns in call power are included in this study as another objective to look for the Lombard Effect. Call attributes were measured by using the spectrogram, log-transformed waveform, and slice plot views in the Song Scope program.

3.3.4 Statistical analysis

The Car treatments (Car at 90 dB, 80 dB, and 70 dB) were the same as the previous study (Bielinski Dissertation Chapter 2). All treatments were compared against the control treatment in order to determine if a threshold effect exists in either species. In other words, we are trying

to determine at what amplitude can car noise invoke changes in calling behavior. These patterns were then compared to the American bullfrog and green frog data from the previous study to gain insight on potential differences in results between individualized playback experiments and population-level experiments.

Multivariate analyses included a Principle Coordinate Analysis (PCoA) and an UPGMA cluster analysis of call attributes grouped by species and treatment. For each univariate analysis of each species × response variable combination, we calculated the effect size and 95% CIs between the noise treatment versus the control treatment. Statistical models included hurdle, linear, and linear mixed-effects models. Details for the multivariate and univariate analyses methods are available in chapter 2. We used a summary table to compare each effect size and its respective significance between this individual-level playback experiment to the previous population-level playback experiment.

We conducted all statistical analyses on R version 3.5.0. (R Core Team 2019).

3.4 RESULTS

3.4.1 Multivariate Analysis

For the bullfrog we ran 6 playback experiments on unique individuals and analyzed the attributes of 216 calls. For the green frog we ran 5 playback experiments on unique individuals and analyzed 180 calls. The PCoA of our full data set shows an obvious separation by species (Figure 1). For the bullfrog, there is some separation along axis 1 between the calls from control treatments and the calls from car treatments, specifically showing the 90 dB treatment having the greatest separation from the control. This pattern was not seen in the Green frog.

Clustering the multivariate data using the UPGMA method yielded a dendrogram with similar grouping by species, but no grouping by treatment (Figure 2). Like in our population-level study, the multivariate approach indicated a need for a univariate analysis.

3.4.2 Univariate Comparison Between Individual and Population-Level Experiments

Bullfrog

A summary figure shows the differences in results between the individual and population level-studies (Figure 3) and model outputs for all response variables are found in Table 1. For call rate, although we counted fewer calls in the Car 90 dB treatment compared to the control, we found no effect of treatment (Figure 4 A & C). Estimates from the model do show an increase in proportion of zeros for all car treatments. The data for the number of calls from the population-level study shows similar patterns to the individuals, except that the drop in calls was indicated by the count model (Figure 4 B & D). This difference is not a surprise, as it is more difficult to register zeros if there are more males in the recording (Bielinski Dissertation Chapter 2), and easier to register zeros when focusing on just one calling male (this study).

Individual bullfrogs shortened their calls during the treatment 80 and 90 dB treatments (Figure 5). The population level study did not yield similar results as no effect was registered. Individual bullfrogs did not alter their dominant frequency during any of the car treatments (Figure 6). The population-level study yielded a negative relationship between treatment power and frequency, but evidence of this relationship was weak considering the overlap of the 95% CIs.

Green frog

A summary figure shows the differences in results between the individual and population level-studies (Figure 3) and model outputs for all response variables are found in Table 1.

188

Individuals called less during the 90 dB treatment, as evident in the effect sizes from our statistical model (Figure 7). In comparison, the population-level study showed a reduction in calling in response to the car treatments, but with considerable overlap from the 95% confidence intervals.

Individual green frogs shortened their calls during the 80 and 90 dB treatments (Figure 8). However, the population-level study did not yield similar results as no effect was registered. The individual-level study showed a weak reduction in dominant frequency during the 90 dB treatment only, whereas the population-level study showed no effect of any treatment (Figure 9).

3.4.3 Call power

For the bullfrog, we found an effect during the Car 70 dB treatment (+2.5 dB) and car 90 dB treatment (+7.6 dB; Figure 10). For the green frog, we found no effect of the noise treatments (Figure 11). However, there is a noticeable positive trajectory across the three treatments.

3.5 DISCUSSION

3.5.1 Call length

The major finding in our comparative procedure was the clear difference in call length changes between the two studies. In this individual-level study, we found a threshold effect for the reduction in call length starting at 80 dB for both species. This was not seen for either species in the population-level study. In response to anthropogenic noise, it may be that some individuals choose to shorten their calls, but in aggregate the main strategy may be to maintain the usual call length. Similarly, Stratman & Hobel (2019) found that call length was the trait in individual males for which females had the highest preference, but in a chorus setting was less detectable to females compared to call rate.

189

3.5.2 Call rate

The differences in call rate patterns between the individual and population-level studies were not major. The individual call data showed a large reduction in calling for the green frog and no change for the bullfrog, whereas the population-level experiments showed the opposite. Because only one datum was collected in an entire treatment (as opposed to the other attributes where data were collected from several calls *within* a treatment), the dataset was much smaller for this attribute. Considering that both species reduced their call counts in response to *pooled* noise in the population study, we anticipate that both species would show a reduction in call count in the individual-level experiment if we had a larger sample size. However, it is unknown if this would be the case.

An alternative explanation of these discrepancies could be variation in calling of individual frogs. Individual frogs can raise their level of attractiveness in competitive scenarios (Bee *et al.* 2000; Martinez-Rivera & Gerhardt 2008, Zhu *et al.* 2017). The calling behavior of individual frogs varies enough so that males and females can distinguish between individuals (Bee & Gerhardt 2001; Bee *et al.* 2012). Therefore, chorus-level measurements may show an aggregate reduction in call rate if this was the most common response, but some individuals may take advantage of lower amounts of conspecific calls by maintaining their normal call rate.

3.5.3 Dominant frequency

The small reduction in dominant frequency detected in the green frogs during the 90 dB treatment could indicate misinterpretation of the playback noise. Green frogs can reduce their frequency in a competitive environment (Bee *et al.* 2000). Lower frequency calls are more attractive to females as it is a dependable signal for larger size (Tarano & Herrera 2003). If a frog perceived the low frequency noise as an upswing in chorusing at the pond, then lowering his

own call frequency would help him stay competitive in a larger field of suitors. However, considering the median dominant frequency during our 90 dB treatment (~ 450 Hz) was *higher* than the 90 dB treatment for the population-level study (~ 400 Hz) the more convincing explanation simply is that this effect was due to natural variation across individuals.

The complete absence of increased frequencies for both species was a meaningful result of this study. A possible explanation lies in the fact that low-frequency callers lack the plasticity to raise their frequencies higher than the range of anthropogenic noise, meaning it is a wasteful strategy. This result is accordant with the population-level study. Therefore, it is a strong indication that even though many previous studies have noted changes in frequencies in response to noise (up: Kruger and Du Preez 2016; Shen & Xu 2016; or down: Hanna *et al.* 2014; Caorsi 2017), some species may not have the spectral plasticity to adjust to a range that would stand out amongst the noise.

3.5.4 Call power

Call power can only be measured in either lab or individual-level field studies, as it requires a recorder to continually be the same distance away from a single frog for the entirety of the playback experiment. Thus, this individual-level study added one more response variable to our bullfrog and green frog datasets that we would have otherwise been overlooked (Bielinski Dissertation Chapter 2 Table 1). We found evidence for the Lombard Effect in the American bullfrogs. Individuals made louder calls (+ 7.6 dB) during the Car 90 dB treatment only. The natural power for bullfrogs is about 80 dB (Megela-Simmons 1984; Boatright-Horowitz 2000). Thus, it seems that bullfrogs raised their SPL once the peak treatment amplitude surpassed their natural power level. We did not see an increase in power in the green frogs. There was a slight positive trend between calling SPL and treatment SPL, but differences were well within 95%

CIs. Cunnington & Fahrig (2010) found that green frogs actually reduced their call power in response to noise. Modifying the amplitude of calls seems to be a strategy that is different across species, as past research provides mixed results (for: Halfwerk *et al.* 2016, Yi & Sheridan 2019; against: Love & Bee 2010). Potentially, these strategies could be different from across populations of the same species, depending on their natural level of exposure to noise (Cunnington & Fahrig 2010).

3.5.5 Conclusion

Overall, the comparison between our two studies revealed several key findings. First, patterns found in some attributes of individual frog behavior may not be translatable to a chorusing group. In our case, this was seen most clearly in call length. Aggregate patterns seen in population-level studies may not indicate the response across all individuals, as those with lower call rates tend to be overlooked. Individual-based experiments should be able to capture this variation. We should also be hesitant to extrapolate patterns or mechanisms of individual studies to a loud chorus environment. Stratman & Hobel (2019) mirrors this sentiment when they warn about arbitrariness of decision-making in choruses found in the wild.

Ideally, researchers should utilize both chorus-level and individual-level methods on the same species. Chorus-level studies should better capture the chaos and arbitrariness that females will encounter when making mate choices, whereas individual-level studies should be better at identifying changes made by individual males in an attempt stand out amidst the chaos.

3.6 LITERATURE CITED

Bee, M.A., and H.C. Gerhardt. 2001. Neighbour-stranger discrimination by territorial male bullfrogs (rana catesbeiana): I. Acoustic basis. Animal Behaviour 62:1129-1140.

Bee, M.A., S.A. Perrill, and P.C. Owen. 2000. Male green frogs lower the pitch of acoustic signals in defense of territories: A possible dishonest signal of size? Behavioral Ecology 11:169-177.

Bee, M.A., A. Velez, and J.D. Forester. 2012. Sound level discrimination by gray treefrogs in the presence and absence of chorus-shaped noise. Journal of the Acoustical Society of America 131:4188-4195.

Bleach, I.T., C. Beckmann, C. Both, G.P. Brown, and R. Shine. 2015. Noisy neighbours at the frog pond: Effects of invasive cane toads on the calling behaviour of native australian frogs. Behavioral Ecology and Sociobiology 69:675-683.

Boatright-Horowitz, S.L., S.S. Horowitz, and A.M. Simmons. 2000. Patterns of vocal interactions in a bullfrog (rana catesbeiana) chorus: Preferential responding to far neighbors. Ethology 106:701-712.

Caorsi, V.Z., C. Both, S. Cechin, R. Antunes, and M. Borges-Martins. 2017. Effects of traffic noise on the calling behavior of two neotropical hylid frogs. Plos One 12.

Gerhardt, H.C. 1991. Female mate choice in treefrogs - static and dynamic acoustic criteria. Animal Behaviour 42:615-635.

Gerhardt, H.C. 1994. The evolution of vocalization in frogs and toads. Annual Review of Ecology and Systematics 25:293-324.

Halfwerk, W., A.M. Lea, M.A. Guerra, R.A. Page, and M.J. Ryan. 2016. Vocal responses to noise reveal the presence of the lombard effect in a frog. Behavioral Ecology 27:669-676.

Hanna, D.E.L., D.R. Wilson, G. Blouin-Demers, and D.J. Mennill. 2014. Spring peepers pseudacris crucifer modify their call structure in response to noise. Current Zoology 60:438-448.

Kaiser, K., and J.L. Hammers. 2009. The effect of anthropogenic noise on male advertisement call rate in the neotropical treefrog, dendropsophus triangulum. Behaviour 146:1053-1069.

Kruger, D.J.D., and L.H. Du Preez. 2016. The effect of airplane noise on frogs: A case study on the critically endangered pickersgill's reed frog (hyperolius pickersgilli). Ecological Research 31:393-405.

Love, E.K., and M.A. Bee. 2010. An experimental test of noise-dependent voice amplitude regulation in cope's grey treefrog, hyla chrysoscelis. Animal Behaviour 80:509-515.

Luddecke, H. 2002. Male and female responses to call playback in the andean frog colostethus subpunctatus. Amphibia-Reptilia 23:141-150.

Martinez-Rivera, C.C., and H.C. Gerhardt. 2008. Advertisement-call modification, male competition, and female preference in the bird-voiced treefrog hyla avivoca. Behavioral Ecology and Sociobiology 63:195-208.

Megela-Simmons, A. 1984. Behavioral vocal response thresholds to mating calls in the bullfrog, r anacatesbeiana. The Journal of the Acoustical Society of America 76:676-681.

Parris, K.M., M. Velik-Lord, and J.M.A. North. 2009. Frogs call at a higher pitch in traffic noise. Ecology and Society 14.

Phillips, C.A., R.A. Brandon, and E.O. Moll. 1999. Field guide to amphibians and reptiles of illinois. Illinois Natural History Survey.

Reichert, M.S., L.B. Symes, and G. Hobel. 2016. Lighting up sound preferences: Cross-modal influences on the precedence effect in treefrogs. Animal Behaviour 119:151-159.

Ryan, M.J. 1980. The reproductive behavior of the bullfrog (rana catesbeiana). Copeia:108-114.

Shen, J.X., and Z.M. Xu. 2016. The lombard effect in male ultrasonic frogs: Regulating antiphonal signal frequency and amplitude in noise. Scientific Reports 6.

Stratman, K.D., and G. Höbel. 2019. Opportunity for female choice in the wild is frequently curtailed by low male trait variation in hyla versicolor. Behavioral Ecology and Sociobiology 73:59.

Sun, J.W.C., and P.A. Narins. 2005. Anthropogenic sounds differentially affect amphibian call rate. Biological Conservation 121:419-427.

Tarano, Z., and E.A. Herrera. 2003. Female preferences for call traits and male mating success in the neotropical frog physalaemus enesefae. Ethology 109:121-134.

Vargas-Salinas, F., and A. Amezquita. 2013. Traffic noise correlates with calling time but not spatial distribution in the threatened poison frog andinobates bombetes. Behaviour 150:569-584.

Vargas-Salinas, F., G.M. Cunnington, A. Amézquita, and L. Fahrig. 2014. Does traffic noise alter calling time in frogs and toads? A case study of anurans in eastern ontario, canada. Urban Ecosystems 17:945-953.

Yi, Y.Z., and J.A. Sheridan. 2019. Effects of traffic noise on vocalisations of the rhacophorid tree frog kurixalus chaseni (anura: Rhacophoridae) in borneo. Raffles Bulletin of Zoology 67:77-82.

Zhu, B.C., J.C. Wang, Z.X. Sun, Y. Yang, T.L. Wang, S.E. Brauth, Y.Z. Tang, and J.G. Cui. 2017. Competitive pressures affect sexual signal complexity in kurixalus odontotarsus: Insights into the evolution of compound calls. Biology Open 6:1913-1918.

3.7 TABLES

Table 1. Effect sizes and model statistics of call attributes for the car noise treatments *vs* control treatments. Noise treatments include car noise at 70 dB (C70), 80 dB (C80), and 90 dB (C90). The split components of the hurdle models include the zero model (HZ) and the count model (HC). Other models include linear models (LM) and linear mixed-effect models (LMEM). HZ effect sizes are the percent change between a 1:1 ratio of zeros and the control : treatment ratio of zeros. LM call rate and call length effect sizes are the percent change from the control calls. Dominant frequency effect sizes are listed as the change in hertz from the control calls. Call power effect sizes are the change in SPL (dB) from the control calls.

| | | | **Car at 70, 80 and 90 dB *vs.* Control** | | | | | | |
Response Variable	Species	Model	C70 *vs.* Control Effect Size, P-Value		C80 *vs.* Control Effect Size, P-Value		C90 *vs.* Control Effect Size, P-Value		F-Statistic or Hurdle Model df
Call Rate	Bullfrog	HZ	-12%	0.514	-12%	0.514	-44%	0.130	9
	Bullfrog	HC	+6%	0.850	+9%	0.764	+4%	0.923	9
	Green frog	LM	+13%	0.685	+4%	0.907	-59%	0.086	$F_{3,16} = 2.1$
Dominant Frequency	Bullfrog	LM	-19 Hz	0.379	-1 Hz	0.974	-34 Hz	0.214	$F_{3,79} = 0.8$
	Green frog	LMEM	-3 Hz	0.800	-19 Hz	0.209	-29 Hz	0.098	$F_{3,87} = 1.3$
Call Length	Bullfrog	LM	-1%	0.857	-14%	< 0.001	-19%	< 0.001	$F_{3,86} = 7.6$
	Green frog	LM	-2%	0.697	-15%	0.033	-29%	< 0.001	$F_{3,100} = 6.7$
Call Power	Bullfrog	LM	+2.5 dB	0.010	+2 dB	0.155	+7.6 dB	< 0.001	$F_{3,81} = 5.4$
	Green frog	LMEM	-1 dB	0.215	-0.2 dB	0.788	+0.4 dB	0.621	$F_{3,96} = 1.0$
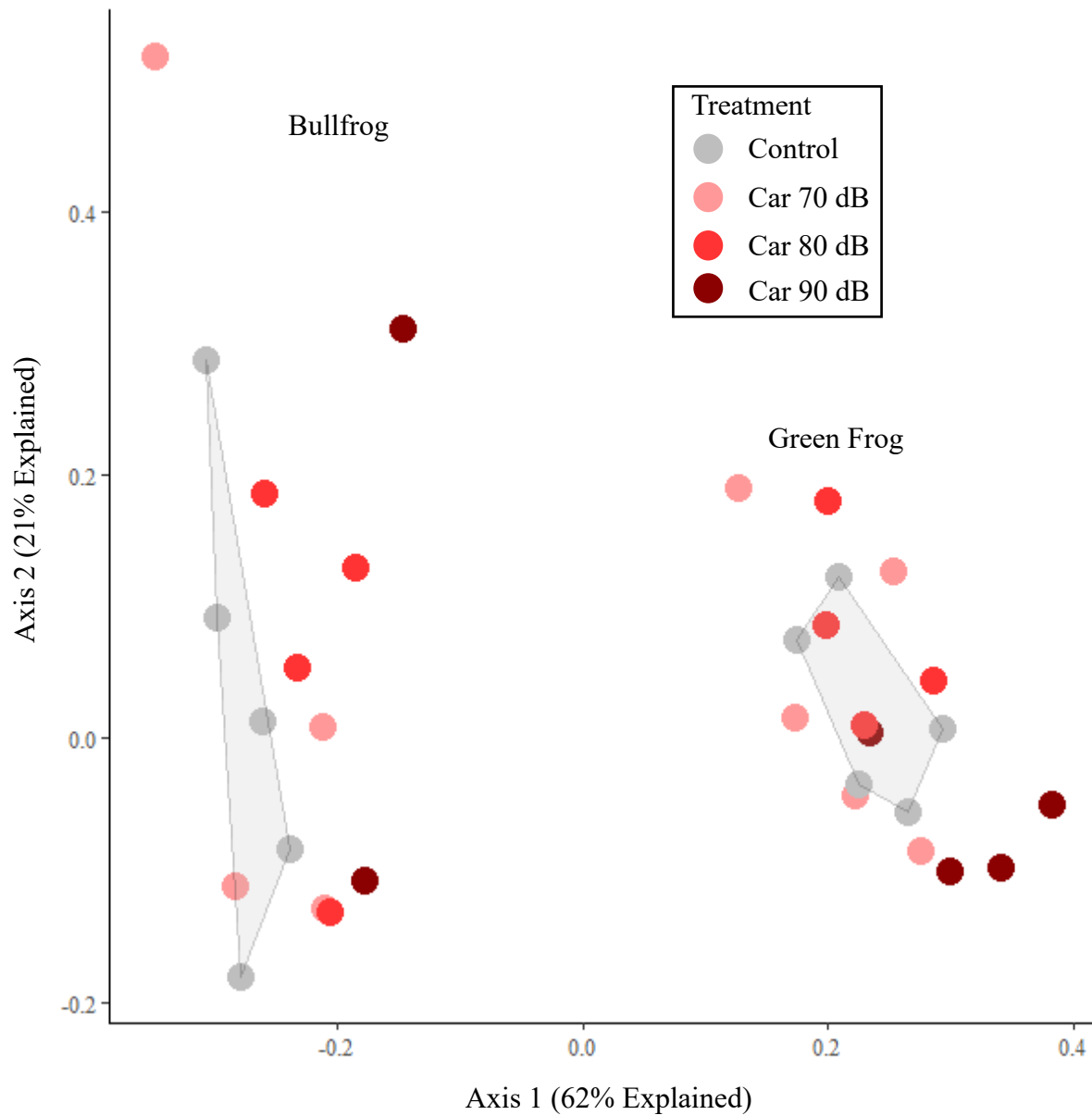
Figure 1. Principal Coordinates Analysis of call attributes for both target species. The call attributes in multivariate space clearly separate between species, as bullfrogs and green frogs have different dominant frequencies and call lengths. Polygons are drawn around control treatments. There is clear grouping for the bullfrog, but no separation for the green frog.
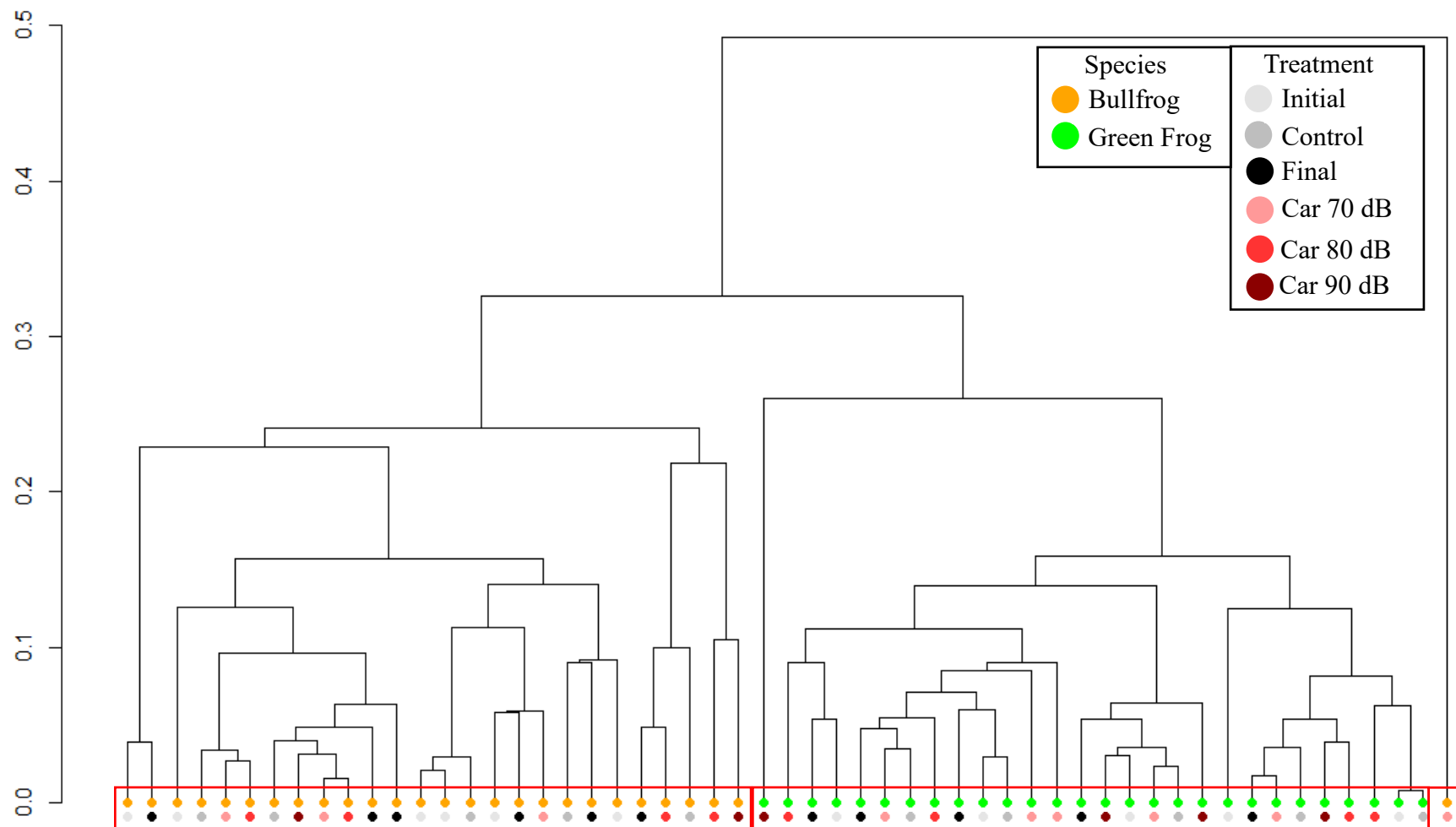
Figure 2. Cluster Analysis. The UPGMA clustering method was used to create a dendrogram of the multivariate data of both species combined. Red boxes represent suggested number of groups, as indicated by the silhouette widths that measure the intensity of the link between objects and their groups. There is clear grouping across species. There is no grouping across treatments.

Study Level	Common Name	# Calls			Pitch			Length		
		70	80	90	70	80	90	70	80	90
Population	Bullfrog			⬇ (>-50%)						
Individual	Bullfrog								⬇ **	⬇ **
Population	Green Frog					⬇				
Individual	Green Frog			⬇ *		⬇			⬇ *	⬇ ***

Effect Sizes
- ☐ No Effect
- ⬇ - 0-10%
- ⬇ - 11-50%
- ⬇ > - 50%

Significance
- (blank) p < .1
- * p < .05
- ** p < .01
- *** p < .00

Figure 3. Comparison of effect sizes between individual playback experiments and population-level playback experiments. Population level experiments were conducted previously. For the Bullfrog, individuals reduced the length of their calls by 14% and 18% for the 80 dB and 90 dB car treatment, respectively. For the Green Frog, individuals reduced the number of calls by 59% at the 90 dB car treatment. Green frogs reduced their pitch by 6% at the 90 dB car treatment. Green frogs reduced their call length by14% and 29% for the 80 dB and 90 dB car treatment, respectively. Most notably, the population level studies did not show a reduction in call length, whereas there were clear signs of shorter calls in the individual level studies.

Figure 4. Bullfrog call rate. A comparison of the effect of car noise intensity on call number between individual-level and population-level experiments. Noise power levels were at 70, 80, and 90 dB. Effect size estimates are from the model hurdle (number of calls ~ treatment, distribution = negative binomial). Top graphs (A & B) show the proportion of zeros. Bottom graphs (C & D) show the change in non-zeros. Left graphs (A & C) are from the current individual-level study. Left graphs (B & D) are from the previous population-level study. For more detail of effect size estimates, refer to Chapter 2 Appendix A.
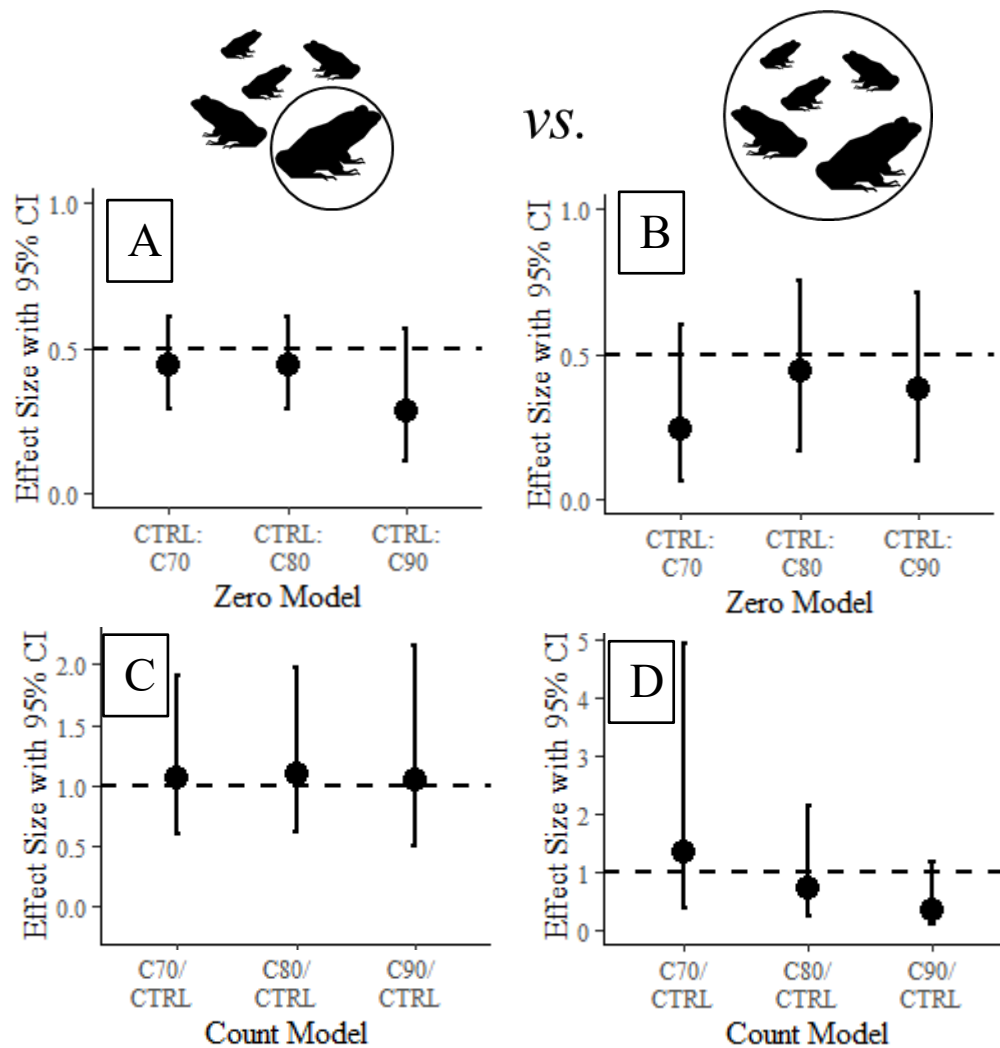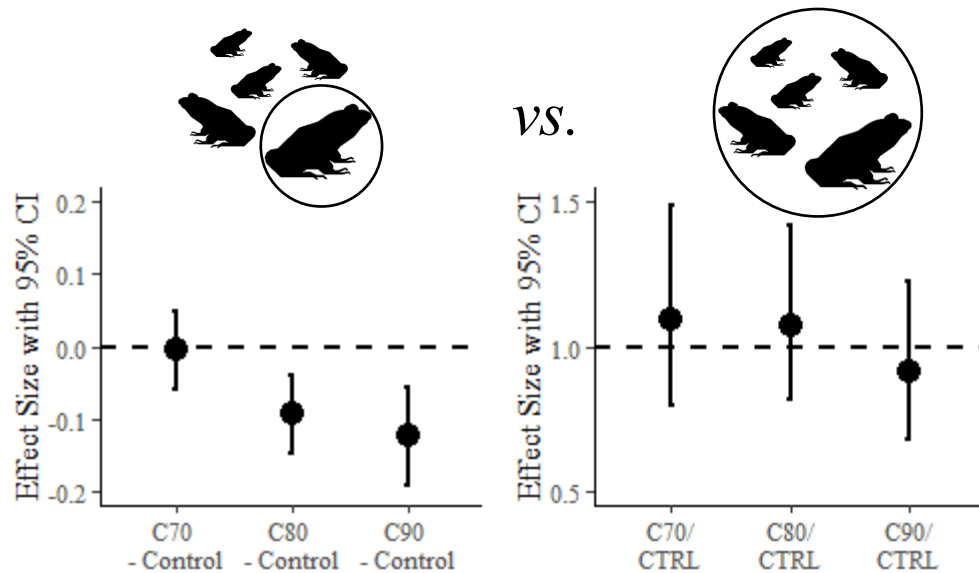
Figure 5. Bullfrog call length. A comparison of the effect of car noise intensity on call length between individual-level and population-level experiments. Noise power levels were at 70, 80, and 90 dB. Effect size estimates are from the model lm (call length ~ treatment). The left graph is from the current individual-level study. The right graph is from the previous population-level study. For more detail of effect size estimates, refer to Chapter 2 Appendix A.

Figure 6. Bullfrog dominant frequency. A comparison of the effect of car noise intensity on dominant frequency between individual-level and population-level experiments. Noise power levels were at 70, 80, and 90 dB. Effect size estimates for individuals are from the model lm (call length ~ treatment). Effect size estimates for populations are from the model lmer (dominant frequency ~ treatment + (1|initial conditions)). The left graph is from the current individual-level study. The right graph is from the previous population-level study. For more detail of effect size estimates, refer to Chapter 2 Appendix A.
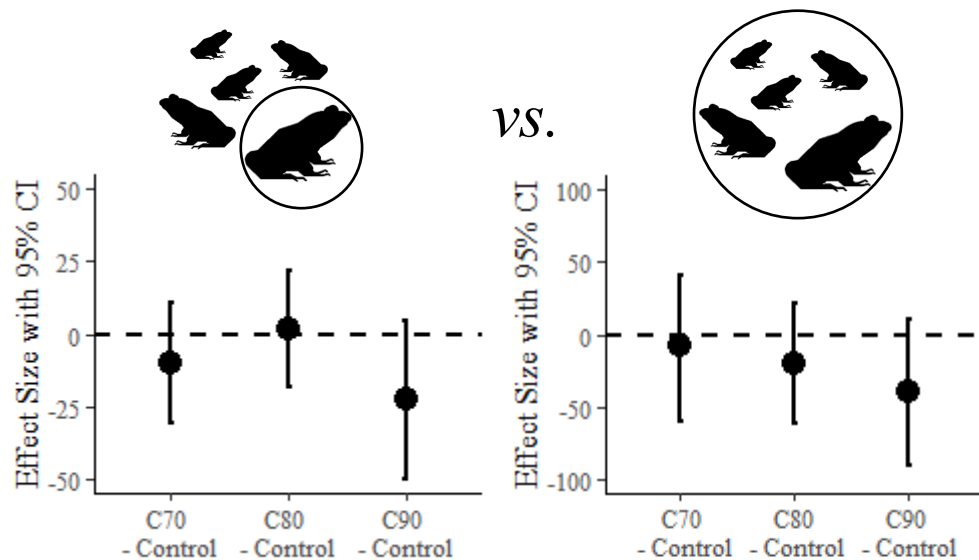
Figure 7. Green frog call rate. A comparison of the effect of car noise intensity on call count between individual-level and population-level experiments. Noise power levels were at 70, 80, and 90 dB. Individual-level effect size estimates are from the model lmer (number of calls ~ treatment + (1|individual). Population-level effect size estimates are from the model hurdle (number of calls ~ treatment, distribution = negative binomial). (B) shows the proportion of zeros. The bottom graph (C) show the change in non-zeros. The Left graph (A) is from the current individual-level study. Right graphs (B & C) are from the previous population-level study. For more detail of effect size estimates, refer to Chapter 2 Appendix A.

Figure 8. Green frog call length. A comparison of the effect of car noise intensity on call length between individual-level and population-level experiments. Noise power levels were at 70, 80, and 90 dB. Effect size estimates are from the model lm (call length ~ treatment). The left graph is from the current individual-level study. The right graph is from the previous population-level study. For more detail of effect size estimates, refer to Chapter 2 Appendix A.
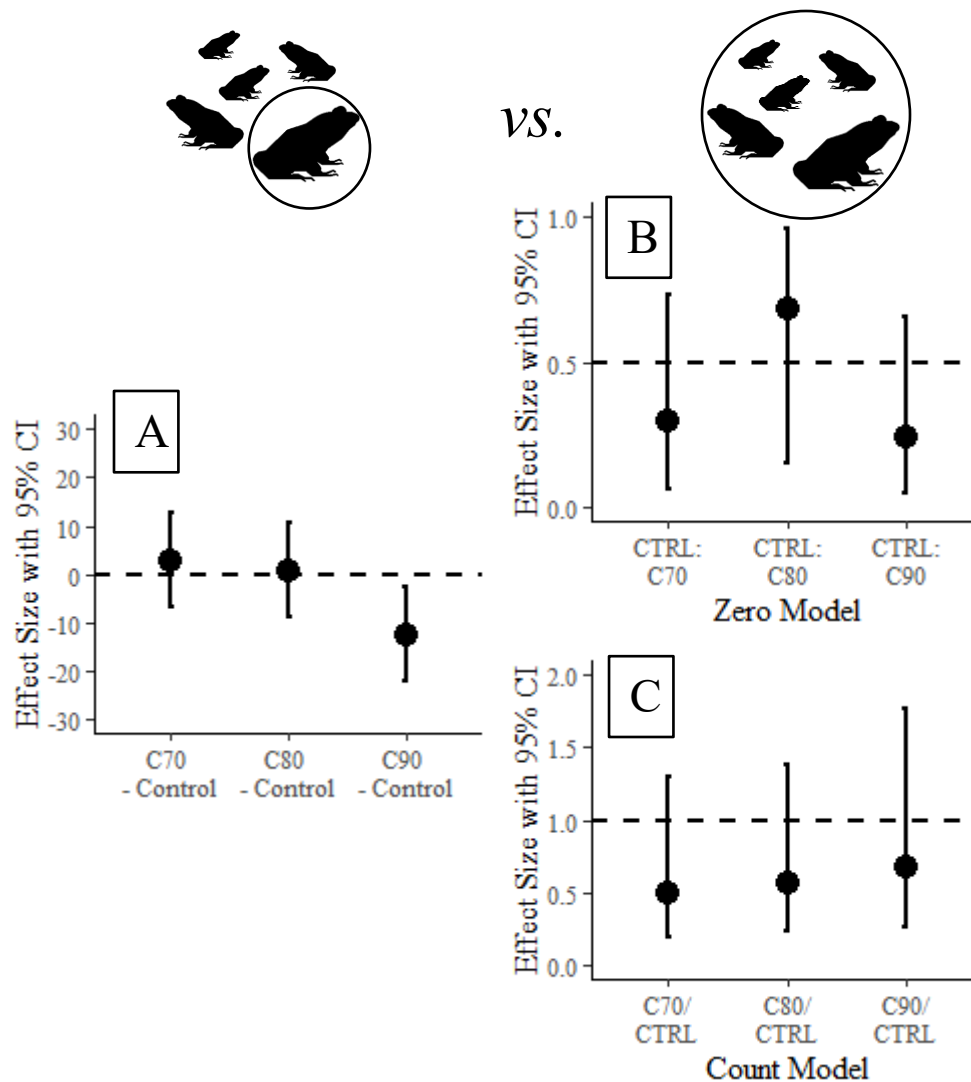

Figure 9. Green frog dominant frequency. A comparison of the effect of car noise intensity on dominant frequency between individual-level and population-level experiments. Noise power levels were at 70, 80, and 90 dB. Effect size estimates are from the model lm (call length ~ treatment). The left graph is from the current individual-level study. The right graph is from the previous population-level study. For more detail of effect size estimates, refer to Chapter 2 Appendix A.
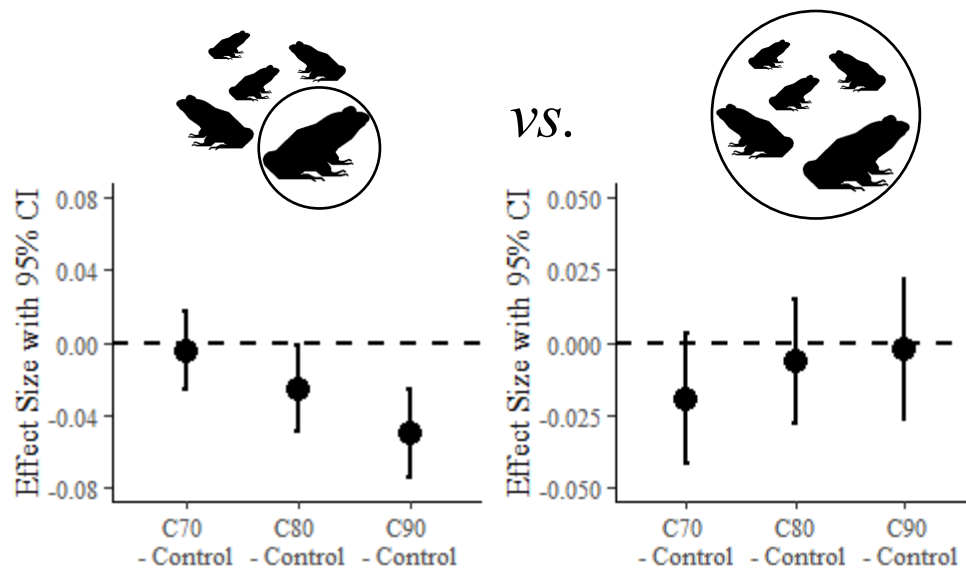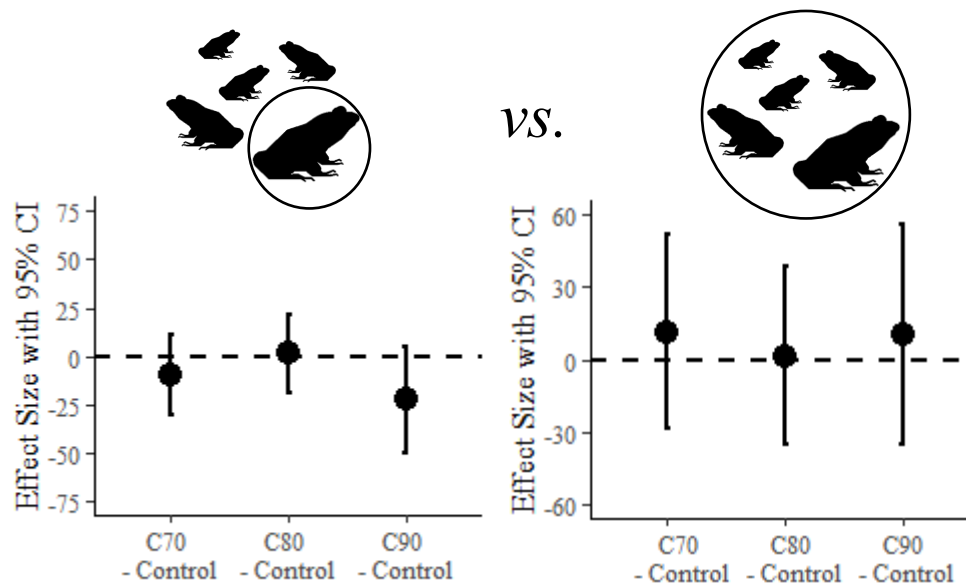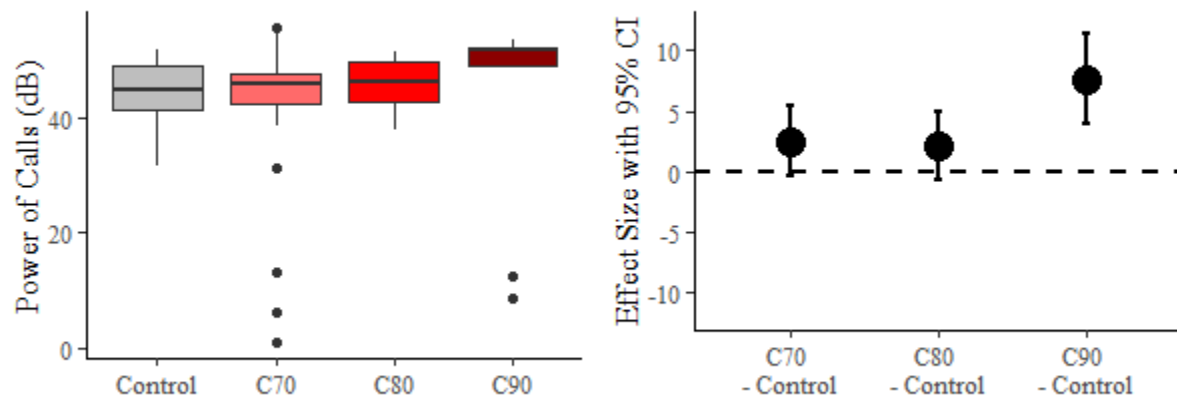
Figure 10. Effect of intensity of car noise on call power for American bullfrog individuals. Noise power levels were at 70, 80, and 90 dB. Box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lm (call power ~ treatment). This data provides evidence for the Lombard Effect. For more detail of effect size estimates, refer to Chapter 2 Appendix A.
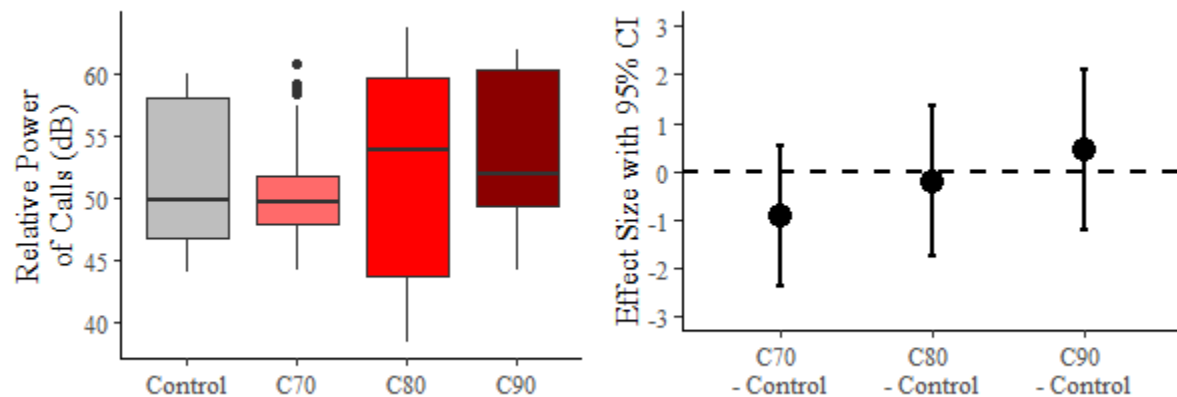
Figure 11. Effect of intensity of car noise on call power for green frog individuals. Noise power levels were at 70, 80, and 90 dB. Box plots represent call counts from a single sample, separated by treatment. Effect size estimates are from the model lmer (call power ~ treatment + (1|block)). This data does not provide evidence for the Lombard Effect. For more detail of effect size estimates, refer to Chapter 2 Appendix A.
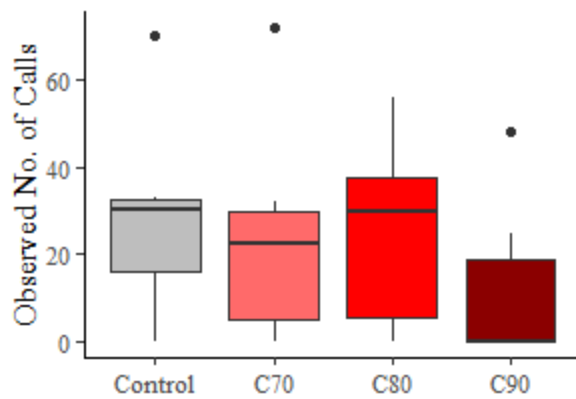
Appendix A. Call count data across treatments for the American bullfrog. Car noise power levels were at 70, 80, and 90 dB (C70, C80, and C90 respectively). Box plots represent call counts from a single sample, separated by treatment.
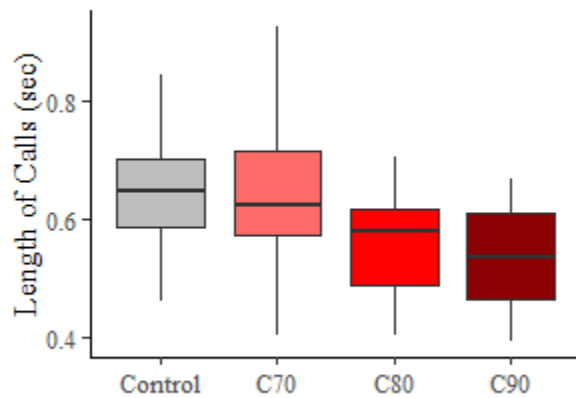
Appendix B. Call length data across treatments for the green frog. Car noise power levels were at 70, 80, and 90 dB (C70, C80, and C90 respectively). Box plots represent measured length of calls from a single call, separated by treatment.
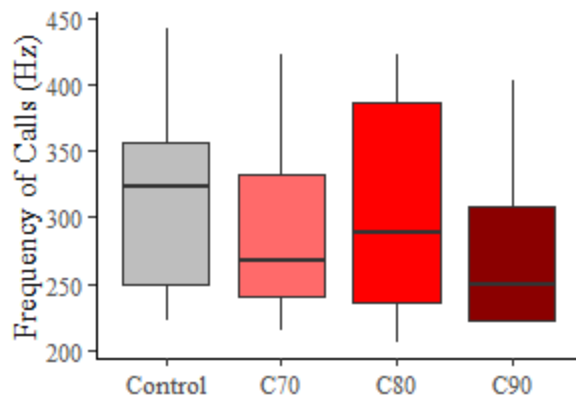
Appendix C. Dominant frequency data across treatments for the American bullfrog. Car noise power levels were at 70, 80, and 90 dB (C70, C80, and C90 respectively). Box plots represent measured dominant frequency from a single call, separated by treatment.

Appendix D. Call count data across treatments for the green frog. Car noise power levels were at 70, 80, and 90 dB (C70, C80, and C90 respectively). Box plots represent call counts from a single sample, separated by treatment.

Appendix E . Call length data across treatments for the Green Frog. Noise power levels were at 70, 80, and 90 dB. Box plots represent measured call length from a single call, separated by treatment.
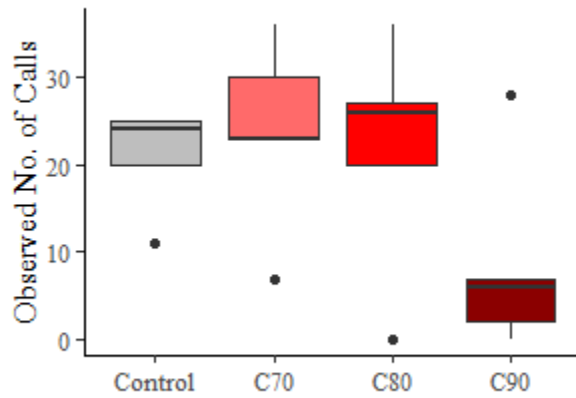

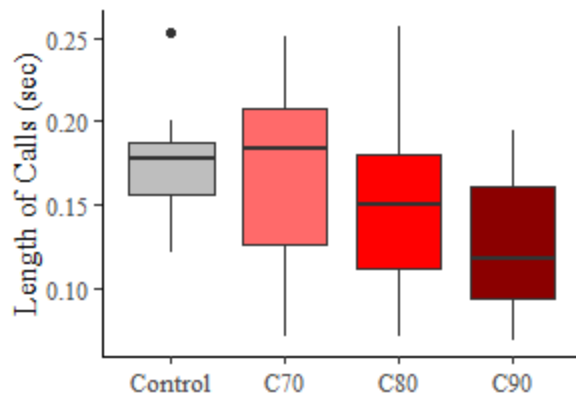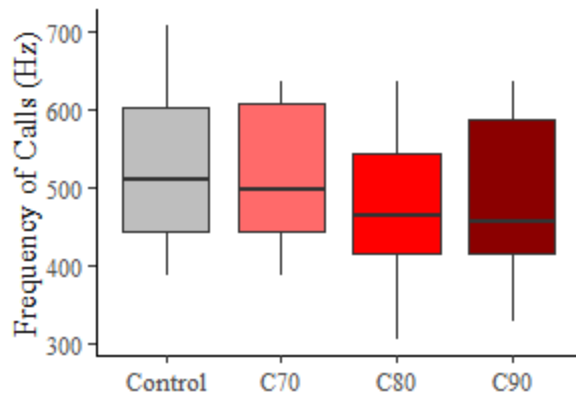Appendix F. Dominant frequency data across treatments for the green frog. Car noise power levels were at 70, 80, and 90 dB (C70, C80, and C90 respectively). Box plots represent measured dominant frequency from a single call, separated by treatment.

exp.id	sp	hetero	aci	ndsi	period	trt	initial.call	calls
i.1	bf	0	324.8986	0.664858	1	i	3	3
i.1	bf	0	324.8986	0.664858	2	c90	3	0
i.1	bf	0	324.8986	0.664858	3	c70	3	0
i.1	bf	0	324.8986	0.664858	4	c80	3	0
i.1	bf	0	324.8986	0.664858	5	e	3	0
i.1	bf	0	324.8986	0.664858	6	f	3	4
i.2	bf	0	326.8492	-0.4773	1	i	14	14
i.2	bf	0	326.8492	-0.4773	2	c70	14	24
i.2	bf	0	326.8492	-0.4773	3	c80	14	23
i.2	bf	0	326.8492	-0.4773	4	e	14	29
i.2	bf	0	326.8492	-0.4773	5	c90	14	25

Appendix G. Data frame for call counts. A partial view of the raw data collected from the individual playback experiments. The full dataset is available online through the Center for Open Science data repository at https://osf.io/xwzjk/ under the file name "DF1ind.csv".

exp.id	sp	hetero	index1	index2	period	trt	dif.if	initial.call	length	power	p.num	domfreq
i.1	bf	0	324.8986	0.664858	1	i	-1	3	0.505	42.24	1	166.09
i.1	bf	0	324.8986	0.664858	1	i	-1	3	0.414	42.49	1	55.36
i.1	bf	0	324.8986	0.664858	1	i	-1	3	0.434	41.48	1	110.73
i.1	bf	0	324.8986	0.664858	1	i	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	1	i	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	1	i	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	2	c90	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	2	c90	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	2	c90	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	2	c90	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	2	c90	-1	3	na	na	na	na
i.1	bf	0	324.8986	0.664858	2	c90	-1	3	na	na	na	na

Appendix H. Data frame for call length, dominant frequency, and power. A partial view of the raw data collected from the individual playback experiments. The full dataset is available online through the Center for Open Science data repository at https://osf.io/xwzjk/ under the file name "DF2ind.csv".

exp.id	sp	hetero	aci	ndsi	period	trt	initial.ce	calls	length	power	domfreq
i.1	bf	0	324.8986	0.664858	1	i	3	3	0.451	28.93	110.7267
i.1	bf	0	324.8986	0.664858	2	c90	3	0	na	na	na
i.1	bf	0	324.8986	0.664858	3	c70	3	0	na	na	na
i.1	bf	0	324.8986	0.664858	4	c80	3	0	na	na	na
i.1	bf	0	324.8986	0.664858	5	e	3	0	na	na	na
i.1	bf	0	324.8986	0.664858	6	f	3	4	0.445	33.1775	83.0425
i.2	bf	0	326.8492	-0.4773	1	i	14	14	0.617833	12.195	249.63
i.2	bf	0	326.8492	-0.4773	2	c70	14	24	0.644	19.63167	237.7433
i.2	bf	0	326.8492	-0.4773	3	c80	14	23	0.628333	20.44	230.6767

Appendix I. Data frame for the multivariate data. A partial view of the raw data collected from the individual playback experiments. The full dataset is available online through the Center for Open Science data repository at https://osf.io/xwzjk/ under the file name "MVind.csv".

VITA

Nolan Kenneth Bielinski
2254 W 113th Street
Chicago, Illinois 60643
Mobile Phone: (815) 274-3247
Email: nolanbielinski@gmail.com

EDUCATION

University of Illinois at Chicago, Chicago IL, Fall 2014 to Present
       PhD in Ecology and Evolutionary Biology, anticipated Spring 2020
       Certificate in the Foundations of College Instruction
       Current GPA: 3.90
Rhodes College, Memphis TN, Fall 2009 to Spring 2013
       Bachelor of Science in Biology, History Minor
       GPA: 3.69/4.0, *cum laude*
Jagiellonian University, Krakow, Poland, Fall 2012 (Study Abroad)
Teton Science School, Grand Teton National Park and Yellowstone National Park WY,
       Summer 2012 (Study Abroad)

TEACHING

Adjunct Professor, Indiana University Northwest, Fall 2018 to present
- BIO-L T570: Evolution (2 semesters). Created from scratch the Indiana University System's first graduate-level online course in evolution specifically for Indiana high school educators.
- BIO-L 100 Laboratory: Humans and the Biological World (1 semester).
- BIO-L 100: Humans and the Biological World (1 semester).

Teaching Assistant, UIC, Fall 2014 to Fall 2019.
- BIO 310: Genetics Laboratory (1 semester). Instructor in Charge.
- BIO 100: Biology of Cells and Organisms (1 semester)
- NATS 106: Chemical and Biological Systems (1 semester)
- BIO 272: Comparative Vertebrate Anatomy and Physiology (1 semester)
- BIO 331: Ecology Lab (2 semesters)
- BIO 101: Populations and Communities (5 semesters)

Private High School Math and ACT Tutor, Chicago IL, June 2017-December 2017.

PROFESSIONAL EXPERIENCE

Natural Resources, Forest Preserve District of Dupage County IL, May 2014-July 2014
- Helped manage forest preserves, including planting native species, seed collection, and clearing invasive species

Deer Management, Forest Preserve District of Dupage County IL, October 2013-March 2014
- Dressed deer and collected samples for Chronic Wasting Disease testing

Dermatopathology Lab Technician, Skin M.D. LLC, Orland Park IL, June-August 2012, June-October 2013
- Assisted in Mohs Surgery cancer analysis by cutting frozen specimens and staining slides

Data Entry, Skin M.D. LLC, Orland Park IL, Summer 2010

RESEARCH

2017-2018, performed 166 noise playback experiments in the field across 9 frog species in Chicago suburbs and northern Indiana.
- 53,410 frog calls counted
- 4,042 frog calls analyzed

2016, recorded frog calls at breeding ponds in Chicago suburbs and northern Indiana. Used soundscape software to develop recognizers that automatically detect frog calls in a recording.
- Worked with 10 undergraduate volunteers (3 current)
- Collected over 500 hours of soundscape recordings at ponds.

2015-2016, ran an experiment on eastern cottontail rabbits' giving up densities at different spatial scales at the UIC greenhouse

MANUSCRIPTS

- In prep: Bielinski, N., & D, Wise. Ribbit and racket: anthropogenic noise playback experiments on urban populations of nine frog species.
- Accepted: Bielinski, N., J. Pajda-Delao, A. Gorniak, & D. Wise. Improving automated detection of frog calls in noisy urban habitats using narrow-banded recognizers.
- Published: Bielinski, K., & N. Bielinski. 2014. UV Radiation Transmittance: Regular Clothing Versus Sun-Protective Clothing. *Am J Orthop*, 94(3), 135-138.

PRESENTATIONS

- Nolan Bielinski. 2019. Ribbit and racket: how does human-made noise affect frog calling behavior? Guest scientist for Science Saturday. Indiana Dunes National Park.
- Nolan Bielinski. 2019. Ecology: The good, the bad, and the muddy. Guest Scientist for Interpreters and Scientists Working on our Parks (iSWOOP) event: People of Color Leaders at the Dunes. Indiana Dunes National Park.
- Nolan Bielinski. 2019. Investigating trends in frog biodiversity. Guest lecturer for BIO 101. University of Illinois at Chicago.
- Nolan Bielinski. 2017. Breaking the Silence: How to Get Your Students Engaged in Discussions and Activities. University of Illinois at Chicago Teaching Assistant Orientation.
- Nolan Bielinski. 2017. Leading Group Discussions in the College Classroom. College Teaching Symposium. University of Illinois at Chicago
- Nolan Bielinski. 2012. Invasion Patterns of Chinese Privet (*Ligustrum sinense*) at Shelby Farms Park. Rhodes Undergraduate Research and Creative Activity Symposium. Memphis, TN.
- Nolan Bielinski. 2011. Spatial Use of the Gray Wolf (*Canis lupus*) Exhibit at Memphis Zoo. Memphis, TN.

GRANT SUPPORT

- Liberal Arts & Sciences Undergraduate Research Initiative, UIC. 2019-2020.Collaborated with and mentored two undergraduates to develop research skills and methods of inquiry in soundscape research. Total award for undergraduate researchers: $5,000.00. Total award for our lab: $1,000.00.
- Blueprints for Success Teaching Grant, Indiana University Northwest, Center for Innovation and Scholarship in Teaching and Learning. 2018. Built a graduate-level online evolution class specifically designed for high school science teachers. Total award: $3000.00

- Provost's Award for Graduate Research, UIC. 2017. "Ribbit and racket: examining frog call changes in response to anthropogenic noises." Total award: $1000.00.
- Society for the Study of Amphibians and Reptiles Grant-in-Herpetology. 2017. "Ribbit and racket: examining frog call changes in response to anthropogenic noises." Total award: $500.00.
- Elmer Hadley Graduate Research Grant, UIC. 2017. "Ribbit and racket: Examining frog call changes in response to anthropogenic noises." Total award: $2,000.00.
- Chicago Herpetological Society Grant. 2016. Graduate Student Research in Herpetology. "Recording the soundscape of frog habitats in Chicago's fragmented landscape." Total award: $500.00.
- Elmer Hadley Graduate Research Grant, UIC. 2016. "The effect of noise pollution on anuran calling behavior and species distributions." Total award: $2,000.00.

RESEARCH AND TECHNCAL SKILLS

Frog Call Identification
R programming for Statistical Computing
Song Scope Soundscape Ecology Software
Canvas Learning Management Platform
Google Earth GIS Software
Tableau Data Visualization Software
Audacity Audio Software
PRIMER-E, including PERMANOVA+ Statistical Computing
Microsoft Office
Gel Electrophoresis
PCR
Kaltura Video Platform
Giving-Up Densities for Foraging Ecology

GRADUATE HONORS

UIC BIOS Research Photo Competition Finalist 2016
Award for Excellence in Teaching 2015

UNDERGRADUATE HONORS

Beta Beta Beta, National Biological Honor Society
Gamma Sigma Epsilon, National Chemistry Honor Society
Phi Alpha Theta, National History Honor Society
Dean's List Fall 2010, Spring 2011, Spring 2013
Southern Collegiate Athletic Conference Academic Honor Roll Spring 2010, 2011 and 2012

SERVICE AND ACTIVITIES

One Million Degrees Volunteer Coach, Fall 2019-Spring 2020. Provided academic, professional and personal support to a low-income community college student. Helped the student set goals, build a resume, and prepare for interviews
What's that Noise? Interpreting the Wetland Soundscape, Participant, 2019. An iSWOOP-funded event looking to build connections between national park rangers and scientists

Investigational Research Design and Innovation Course Mentor, 2017-2018. Helped a high school student with independent research on the effect of crowd noise on captive royal starlings (*Lamprotornis regius*)
UIC Teaching Assistant Orientation Peer Facilitator
Chicago Calling Frog Survey Participant, 2015. Conducted through the Peggy Notebaert Nature Museum
Chicago Herpetological Society Member
Society for the Study of Amphibians and Reptiles Member
The Herpetologists' League Member
Society of Wetland Scientists Member
Soulsville Charter School Volunteer, 2013. Saturday morning middle school math program,
Slachetna Paczka (Christmas Package Project) Volunteer, Krakow, Poland Winter 2012
Rhodes International Peer Advisor Spring 2013