

# Do birds vocalize at higher pitch in noise, or is it a matter of measurement?

Alejandro A. Ríos-Chelén<sup>1,2</sup> · Ambria N. McDonald<sup>1,3</sup> · Ayala Berger<sup>1</sup> · Anna C. Perry<sup>1</sup> · Alan H. Krakauer<sup>1</sup> · Gail L. Patricelli<sup>1</sup>

Received: 19 May 2016 / Revised: 1 October 2016 / Accepted: 3 October 2016 / Published online: 28 December 2016  
© Springer-Verlag Berlin Heidelberg 2016

## Abstract

Studies have found that some birds use vocalizations with higher minimum frequency in noisy areas. Minimum frequency is often measured by visual inspection of spectrograms (“by-eye practice” (BEP)), which is prone to bias, e.g., if low-frequency components are masked by noise. We tested for this bias by comparing measurements of minimum frequency obtained with the BEP for the same set of red-winged blackbird vocalizations (songs and two call types “checks” and “cheers”) played back under ambient, medium, and high noise conditions using a dual playback experiment where both vocalizations and noise were introduced. We compared BEP measurements to those obtained from power spectrum analyses using a pre-set amplitude threshold (“threshold method” (TM)). The BEP was biased when measuring the minimum frequencies of songs and checks, which are masked by noise, but not when measuring cheers, which are higher pitched and thus not masked. Measures using the TM were not affected by

noise, but this method may fail to identify the vocalizations’ lowest frequency if noise necessitates a low (i.e., conservative) threshold. Using the BEP, we also found a bias toward shorter-duration measurements for songs in increasing noise, and for checks, a bias toward increased measures of an energy distribution parameter (Freq5%), likely in correlation with increased measured minimum frequency. Measures taken from the unmasked cheers were similar regardless of the technique used. We discuss limitations of each approach and encourage the use of the TM, as studies using the BEP may lead to spurious results.

## Significance statement

Noise from human activities is ubiquitous. Researchers have found that some birds vocalize at higher frequency (pitch) in noise, hypothesizing that this may improve signal detection in low-frequency noise. Noise may also hinder detection of signal components by researchers using the most common measurement technique (the BEP), which may be mistaken for increasing frequency. To examine this bias, we conducted a dual playback experiment, in which we broadcast the same vocalizations at three background noise levels. We found that BEP measures of minimum frequency increased with increasing noise even though the vocalizations did not change. We recommend the TM which yielded similar measures across noise levels, although it excluded some lower-frequency elements included by the BEP. We encourage researchers to use the TM over the BEP and to validate their methods across noise levels of interest.

Communicated by H. Brumm

**Electronic supplementary material** The online version of this article (doi:10.1007/s00265-016-2243-7) contains supplementary material, which is available to authorized users.

✉ Alejandro A. Ríos-Chelén  
aarios@ecologia.unam.mx

<sup>1</sup> Department of Evolution and Ecology, University of California, Davis, CA, USA

<sup>2</sup> Centro Tlaxcala de Biología de la Conducta, Universidad Autónoma de Tlaxcala, Carretera Tlaxcala Km. 1.5 C. P. 90062 Puebla, Mexico

<sup>3</sup> Department of Biology, Howard University, Washington, D.C., USA

**Keywords** Noise · Bird song · Pitch · Minimum frequency · By-eye practice · Threshold method

## Introduction

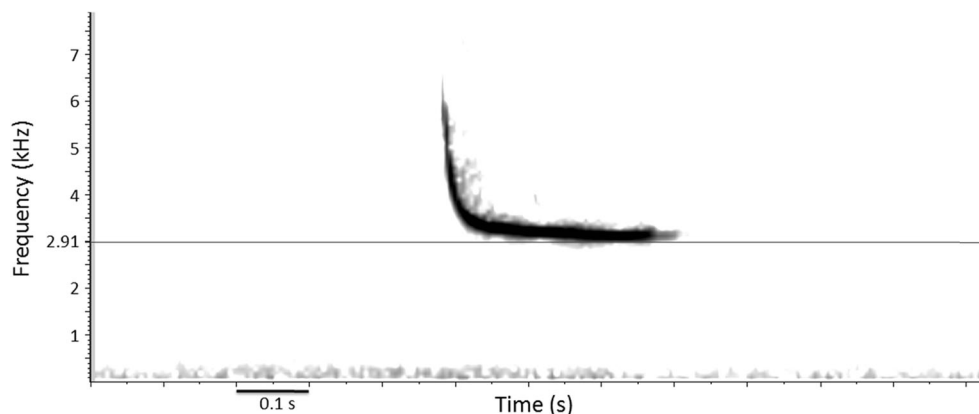
Since Slabbekoorn and Peet (2003) published their study on the positive association between urban noise intensity and the minimum (lowest) frequency of great tit (*Parus major*) songs, researchers have looked for similar relationships in several animal species (reviewed in Brumm and Slabbekoorn 2005; Patricelli and Blickley 2006; Ríos-Chelén 2009; Brumm and Zollinger 2013; Slabbekoorn 2013; Gil and Brumm 2014). A positive correlation between environmental noise levels and minimum song frequency has been reported for several bird species, though not for others (Brumm and Zollinger 2013; Ríos-Chelén et al. 2015). Several researchers have hypothesized that birds may tactically increase the minimum frequency of their songs to avoid masking by low-frequency anthropogenic noise (reviewed in Brumm and Slabbekoorn 2005; Patricelli and Blickley 2006; Ríos-Chelén 2009; Gil and Brumm 2014). Indeed, studies have found that higher-pitched songs are easier to detect and discriminate and evoke a stronger female response (i.e., they emerge more often from their nest boxes), under noisy conditions (Halfwerk et al. 2011; Pohl et al. 2012). Species that sing at a higher frequency in noise might have a fitness advantage over species that lack this ability. However, data relating reproductive success to the production of higher-pitched songs in noisy conditions is still lacking, and the extent to which an increase in pitch is a response to noise or a by-product of singing louder in noise has been the subject of debate (Nemeth and Brumm 2010; Cardoso and Atwell 2011a, 2012; Nemeth et al. 2012, 2013; Slabbekoorn et al. 2012; Zollinger et al. 2012; Potvin and Mulder 2013).

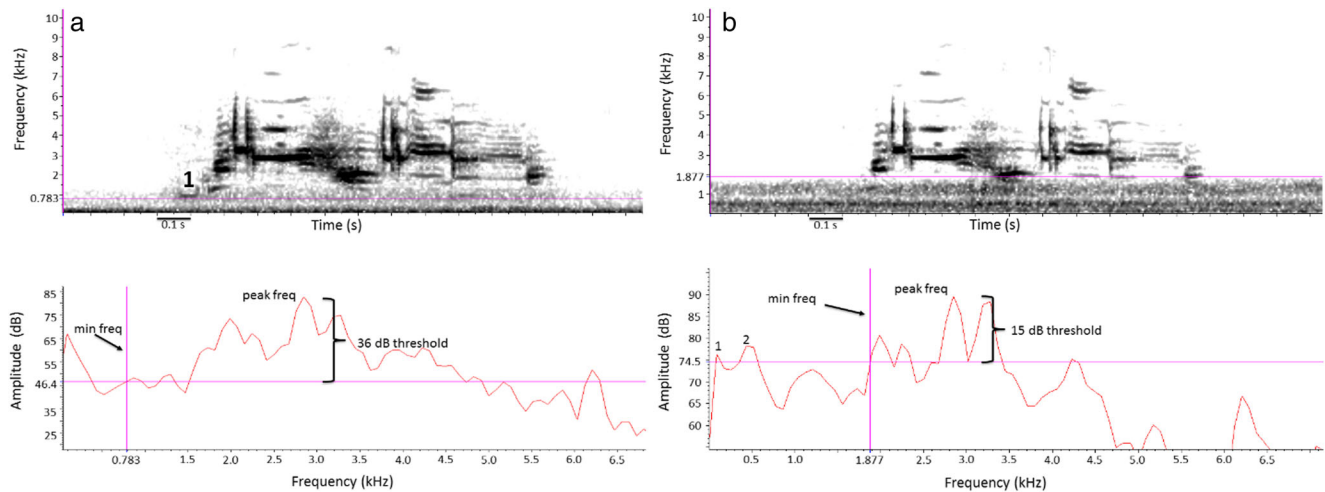
Measurements of minimum frequency, and other attributes of vocalizations, are more accurate when recordings have a high signal-to-noise ratio, so that the signal is easy to detect above the noise. However, to address whether birds change the minimum frequency of their vocalizations in noisy conditions, researchers must often make measurements on recordings with low signal-to-noise ratio. To do so, researchers have relied on two main approaches. The first approach involves

visually inspecting digital spectrograms and measuring the minimum frequency using a cursor (Fig. 1), which we will call the “by-eye practice” (the BEP). An alternative approach, which we refer to as the “threshold method” (the TM), involves using a threshold value on a power spectrum to select the minimum frequency of the sound (Podos 1997). Researchers using the TM select a specific drop in the amplitude (the threshold value) in advance, as measured from the peak frequency, and the lowest-frequency value at which the energy of the vocalization exceeds this threshold is defined as the minimum frequency (Fig. 2a; see also Fig. 2 in Podos 1997). In studies where there is masking noise with energy in lower frequencies, a researcher would choose a threshold by inspecting the distribution of power spectra for each type of vocalization and then selecting a threshold that allows measurement of the lowest detectable vocal elements in the noisiest recordings. By using this threshold, the researcher has an objective means to analyze only the parts of a vocalization which are loud enough to be detected in recordings made at the highest noise level of interest. Therefore, if a low-frequency note would be undetectable under noisy conditions, it will not be measured under quiet conditions, even when it is detectable.

Previous research has shown that the BEP may result in biased measurements if the signal-to-noise ratio is too low (e.g., Beecher 1988; Zollinger et al. 2012). However, even with high signal-to-noise recordings, there are other ways in which frequency measures taken with the BEP can be misleading, including error or bias due to variation in spectrogram settings or recording levels (Zollinger et al. 2012). Additionally, as the relative intensity of low-pitched noise increases, the likelihood that vocalizations with low-frequency sounds will be masked by noise, and thus undetectable by eye on spectrograms, also increases. This is particularly problematic because we typically predict a positive correlation between noise levels and the minimum frequency of bird vocalizations, and the same relationship could arise as an artifact of the BEP. This raises the possibility that studies relying on the BEP could overestimate the effects of noise

**Fig. 1** Minimum frequency labeled on spectrogram with the BEP. The minimum frequency of this cheer, as measured with the BEP, is 2.91 kHz





**Fig. 2** Measuring the minimum frequency of a song using the TM, in a recording with low (i.e., ambient) noise (**a**) and in a recording where the same song is masked by noise playback (**b**). In **a**, the signal-to-noise ratio would allow us to use a threshold level of 36 dB (marked by a horizontal line in **a**) to obtain a measure of minimum frequency. However, in **b**, masking caused by background noise prevents the observer from obtaining an accurate measure of minimum frequency using a threshold of 36 dB. In order to retain this high noise recording in the dataset, the

observer must use a low threshold, in this case 15 dB (see “Methods” section for a discussion of choosing a threshold). This conservative threshold is then applied to all measures of this song type, including the recording in **a**, yielding consistent measures of minimum frequency across noise levels. Peaks of energy (marked as 1 and 2 in **b** power spectrum) result from background noise. Note that the lowest syllable (marked as “1” in **a** spectrogram) is masked by noise in **b** spectrogram

on the minimum frequency, and potentially other measures, when such an effect exists, or may lead to the appearance of a frequency shift in noise when no such shift occurs. While the TM is not expected to be influenced by background noise levels to the same extent as the BEP, measuring the minimum frequency of a vocalization accurately may simply not be possible if the signal-to-noise ratio is not high enough (see “Methods” section and Fig. 2b). These issues also can affect temporal measurements, as the beginning and end of the sounds may be difficult to discern in background noise (Zollinger et al. 2012). In order to understand the impact of anthropogenic and natural noise on animal communication, it is critical that we measure signal characteristics accurately. Therefore, it is important to understand and minimize potential biases introduced by common procedures of measuring acoustic signals.

To determine the proportion of studies using the BEP and TM, we surveyed published research exploring the impact of ambient noise on bird vocalizations. Additionally, to objectively quantify how the BEP and the TM performs with different types of vocalizations under noisy conditions, we carried out a dual playback experiment. We broadcast an identical set of songs and calls from male red-winged blackbirds (*Agelaius phoeniceus*) while simultaneously varying background noise by means of a recording of traffic noise played at different amplitudes from a separate speaker. We then used both BEP and TM to measure time and frequency attributes of the re-recorded vocalizations. We chose red-winged blackbird vocalizations due to a previous study (Ríos-Chelén et al. 2016), which found evidence suggesting that red-winged

blackbirds adjust the minimum frequency of their vocalizations in noise when these frequency measures are taken with the BEP (likely a spurious result), but not with the TM. This suggests that measures from vocalizations of the red-winged blackbird, some of which overlap significantly in frequency with anthropogenic noise, might be susceptible to bias when using the BEP. Zollinger et al. (2012) and Grace and Anderson (2015) demonstrated the same type of bias when using the BEP to measure synthetic signals.

## Methods

### Literature survey

We began by including studies cited in Brumm and Zollinger’s review (2013, combining Tables 7.1, 7.2, and 7.3); we also searched ISI Web of Science using “bird song” and “noise” as keywords for subject, back to 2003 (the year that Slabbekorn and Peet published their influential paper). We focused only on studies that addressed changes in vocal minimum frequency with noise or between habitats differing (or expected to differ) in noise.

### Playback field site

We conducted playback experiments at the Davis Wetlands (38°35' N, 121°39' W), a 400 ac habitat preservation and restoration area in California’s Central Valley, USA. Under the auspices of the City of Davis, the Wetlands site consists

of man-made permanent open water, seasonal wetland, native grassland, riparian woodland, and mudflats. The wetland is interspersed with dirt roads and parking areas. Our experiments were conducted on a dirt parking area next to a small hill leading to a permanent open water wetland.

### Preparation of noise and vocalization files

We recorded urban noise for playback using a shotgun microphone (Sennheiser K6/ME66) connected to a Marantz solid state recorder (PMD670; 16-bit, 44.1-kHz linear PCM). At a distance of ca. 20 m, we aimed the microphone at a busy highway near Davis, CA, USA and recorded ambient urban noise for about 20 min. Traffic levels were mostly continuous during the recording time. Our noise recordings mainly contained noise produced by motor vehicles. For our noise playbacks, we randomly selected 1 min from our noise recording and the file was normalized to the peak amplitude (where the gain is changed to bring the highest PCM sample to the loudest level allowed in the digital system) using Audacity v. 2.05. This 1 min file was played continuously using the “looping play” function in Raven Pro v. 1.4 (Cornell Laboratory of Ornithology 2011). Because we used a noise file recorded at a single location, we should be cautious about generalizing to other noise sources (though the power spectrum of this noise was typical of highway and urban noise and the mechanism for masking should be the same with other, similar noise sources). We used a Larson Davis System 824 sound level meter (SLM) (slow response, A-weighted decibels re 20  $\mu$ Pa; with wind screen) to adjust noise amplitude at 10 m from the speaker to three levels, an ambient noise treatment measured at approximately 45 dB(A) (this was ambient noise with the speaker on but the volume set at zero), a medium noise treatment with noise playback at approximately 55 dB(A), and a high noise treatment with noise playback at approximately 65 dB(A).

We recorded some of our red-winged blackbird vocalizations for playback near Davis, CA, USA, between May and June 2012, using the same equipment that we used to record traffic noise (see Ríos-Chelén et al. 2015 for detailed methods). Other vocalizations were recorded from a population near Ithaca, NY, USA, between April and June 2003, using an omnidirectional microphone (Sennheiser K6/ME62) placed 1 m from the bird and recorded with a MOTU 896 Audio Interface connected to a PowerMac G3 laptop (see Patricelli et al. 2007 for detailed methods). We chose ten different vocalizations of each type (Fig. 3), cheer calls (all from the California population), check calls (all from the California population), and songs (five from New York and five from California populations). Each vocalization, selected on the basis of its quality, was processed in Raven Pro to reduce background noise using a technique similar to a high-pass filter. To do this, we selected the lower part of the

spectrogram (including noise but no vocalization) using the cursor and removed this selection with the “filter out active selection” command. Cheers were filtered at 2000 Hz (any sound below this frequency was removed), while checks and songs were filtered at 200 Hz (using Raven Pro defaults Kaiser window, transition bandwidth of 0.02 times the Nyquist frequency, and stop band attenuation of 100 dB). Each vocalization was normalized to 0 dB using Audacity v.2.05. We then concatenated each vocalization into a single sound file separated by 1-s silent intervals. To remove sharp breaks in amplitude between the recordings and silent portions that can result in clipping, the first and last ~15 ms surrounding each vocalization was faded using the “fade in” and “fade out” commands in Audacity v. 2.05.

### Playback and audio recording

We broadcast pre-recorded highway noise with a rock-shaped speaker (300 W Outdoor Rock Speakers; TIC Corporation, City of Industry, CA). This traffic noise was predominantly low in frequency (<2 kHz). During each noise treatment (see above), we played red-winged blackbird vocalizations using another speaker (CRATE TX15 portable 15-W speaker). Checks, cheers, New York (NY) songs, and California (CA) songs were played at an unweighted sound pressure level of 80, 95, 85, and 85 dB, respectively, measured at 1 m from the speaker using the SLM. These amplitudes correspond to the species-typical amplitudes of each type of vocalization (Patricelli et al. 2007). These vocalizations were then re-recorded using the same equipment that was used to record noise (see above). The person re-recording blackbird vocalizations in traffic noise pointed the microphone toward the speaker playing bird sounds from 10 m away, while facing away from the speaker broadcasting highway noise, which was located 10 m directly behind them. This setup allowed us to collect data on an identical set of vocalizations under different, standardized noise levels.

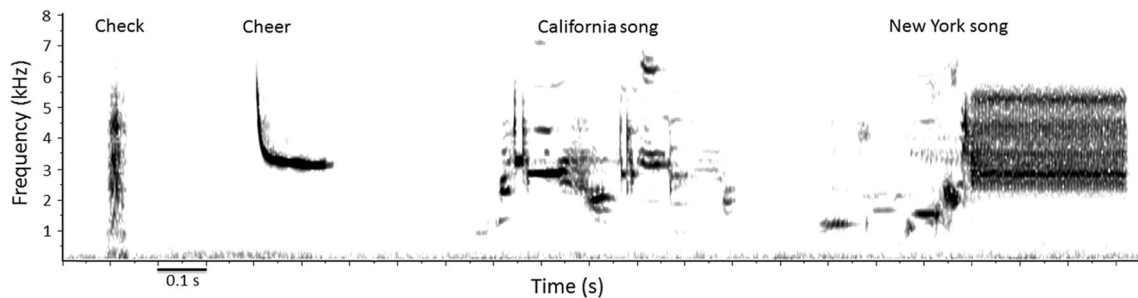
### Acoustic analyses

WAV files were visualized as spectrograms (Hann window, DFT size 512, overlap 50%, frequency resolution 86.1 Hz) in Raven Pro. To minimize observer bias, we used a single-blind study design, where all acoustic measures were collected by a single observer who was unaware of the study’s objectives.

### BEP

The BEP is a simple technique whereby the minimum (or maximum) frequency of a sound is estimated by visual inspection of the spectrogram representation of the signal. In Raven Pro, the boundaries of the vocalization were determined with a





**Fig. 3** Four different red-winged blackbird vocalizations used in this study, two call types (checks and cheers), California songs, and New York songs

“selection box” and the minimum frequency of the vocalization was defined as the bottom edge of this rectangle.

### TM

The TM involves a power spectrum to obtain the peak frequency (the frequency containing the most energy) of the entire sound of interest. For each type of vocalization (checks, cheers, songs), a threshold of sound energy (measured in decibels) was selected prior to measurement that left most of the signal within the selection, while excluding most of the noise (Fig. 2; see also Fig. 2b in Podos 2001). This threshold value was then subtracted from the maximum amplitude value found at the peak frequency for each vocalization. Therefore, the TM defines the minimum frequency as the frequency value at the intersection, to the left of the peak frequency, between the threshold line and the power spectrum line, and the maximum frequency as the value at the intersection to the right of the peak frequency (Fig. 2). The threshold for each type of vocalization is determined by the researcher, based on the difference between the amplitude of the noise and the amplitude of the signal in the noisiest recording included in the set for analysis (the closer the noise amplitude is to the amplitude of the signal of interest, the smaller the threshold must be). For cheer calls, which are high in amplitude and peak frequency, thus only slightly masked by experimental noise, we were able to use a relatively high threshold of 30 dB. For the remaining three types of vocalizations where the minimum frequency was low and masked by experimental noise (i.e., songs and checks), we used a lower threshold value of 15 dB.

### Other measures

In addition to the minimum frequency measurements, the difference between the time values at the beginning and end of the selection (see above, by-eye practice) were used to measure the duration of each vocalization. Because our goal was to investigate how low-frequency anthropogenic noise would affect the accuracy of our BEP temporal measurements, we did not filter noise from our spectrograms and collected all

duration measurements without the aid of a waveform (Zollinger et al. 2012).

We also wanted to evaluate how noise would affect other commonly used frequency measures. We obtained our additional frequency measures automatically with Raven Pro from the selections previously made using the BEP under different experimental noise levels. These additional measures were first quartile frequency ( $Q_1$ ) and third quartile frequency ( $Q_3$ ), the frequencies below which 25 and 75%, respectively, of the energy in the targeted sound are contained, and frequency 5% (Freq5%) and frequency 95% (Freq95%), the frequencies below which 5% and 95%, respectively, of the energy are contained. These additional frequency measures could be used to address whether the distribution of energy in animal vocalizations changes under different environmental conditions. Unlike the minimum frequency of vocalizations, we had no a priori hypotheses about the direction of bias for these frequency measures.

### Statistical analyses

We checked whether measures of CA and NY songs differed from each other. We found that the difference in measured song attributes between the ambient and medium noise levels and between the medium and high levels did not differ between CA and NY song attributes (sample size 5 songs from CA, 5 songs from NY; song duration ambient vs. medium noise levels,  $t$  test,  $t = 2.1$ ,  $df = 8$ ,  $P = 0.069$ ; medium vs. high noise levels,  $t$  test,  $t = 0.16$ ,  $df = 8$ ,  $P = 0.876$ ; song  $Q_1$  ambient vs. medium noise levels,  $t$  test,  $t = 1.1$ ,  $df = 8$ ,  $P = 0.302$ ; medium vs. high noise levels, Mann-Whitney rank sum test,  $U = 10$ ,  $T = 30$ ,  $P = 0.690$ ; song  $Q_3$  ambient vs. medium noise levels, Mann-Whitney rank sum test,  $U = 10$ ,  $T = 30$ ,  $P = 0.690$ ; medium vs. high noise levels, Mann-Whitney rank sum test,  $U = 7.5$ ,  $T = 22.5$ ,  $P = 0.310$ ; song Freq5% ambient vs. high noise levels,  $t$  test,  $t = -0.535$ ,  $df = 8$ ,  $P = 0.608$ ; medium vs. high noise levels, Mann-Whitney rank sum test,  $U = 12$ ,  $T = 28$ ,  $P = 1.000$ ; song Freq95% ambient vs. medium noise levels, Mann-Whitney rank sum test,  $U = 10$ ,  $T = 25$ ,  $P = 0.690$ ; medium vs. high noise levels, Mann-Whitney rank sum test,  $U = 8$ ,  $T = 32$ ,  $P = 0.421$ ). Thus, we pooled our data from CA and NY songs in further analyses.

We evaluated whether noise treatment (ambient, medium, or high noise) and the measurement procedure (BEP or TM) had an effect on the minimum frequency measures for each type of vocalization. We analyzed the minimum frequency for each vocalization type (i.e., songs, checks, and cheers) with a two-way repeated measure ANOVA, with noise treatment and measurement procedure included as factors. We assessed whether the residuals for each model were normally distributed as tested with a Shapiro-Wilk test ( $P > 0.05$ ) and whether the assumption of homoscedasticity was achieved. The residuals for our two-way repeated measure ANOVA were not normally distributed, so we analyzed our data with one-way repeated measure ANOVAs, with noise level and measurement method as factors in separate analyses. If the residuals from these models failed our test for normality, then we analyzed our data with non-parametric Friedman tests for related samples, with noise level or method as independent factors in separate tests. When comparing between two groups (e.g., measures obtained with the TM vs. measures obtained with the BEP), we used paired  $t$  tests when residuals were normally distributed or Wilcoxon signed test for related samples when they were not. For post hoc comparisons, we used the Holm-Sidak test.

We also investigated if our additional sound measures (vocalization duration,  $Q_1$ ,  $Q_3$ , Freq5%, and Freq95%) were influenced by noise level. For this set of analyses, we did not include the measurement procedure as a factor, as all measures were done on selections of vocalizations made using the BEP. We used a one-way repeated measure ANOVA with noise treatment as a factor and our vocalization measures as dependent variables when the residuals reached normality and the assumption of homoscedasticity was met. In cases where normality was not achieved, we used the Friedman test for related samples. Thus, for song duration, check  $Q_1$ , check Freq5%, cheer  $Q_3$ , and cheer Freq95%, we used one-way repeated measure ANOVA. For song  $Q_1$ , song  $Q_3$ , song Freq5% and song Freq95%, check duration, check  $Q_3$ , check Freq95%, cheer duration, cheer  $Q_1$ , and cheer Freq5%, we used Friedman test for related samples. Results are reported as means  $\pm$  SE.

## Results

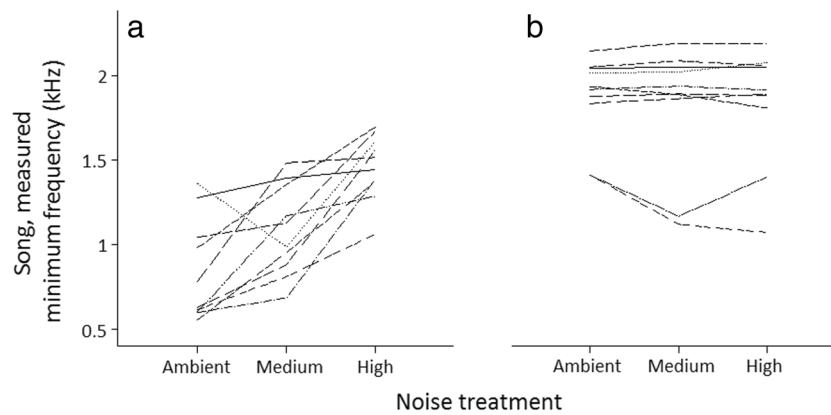
Our review of existing literature identified 49 studies that were focused on detecting changes in minimum frequency across variation in natural or anthropogenic noise (see [Supplementary material](#)). Some of these studies (39 studies) either directly reported their measurement procedure or provided enough information to infer the procedure used (e.g., authors stated that spectrograms were produced and measured likely used the BEP, while authors reporting the use of a chart of frequency and spectral intensities likely used the TM, even when a specific threshold was not reported). Of the 49 studies

that tested for changes in minimum frequency, we found that 20 (40.8%) used or strongly suggested using the BEP and only 13 (26.5%) used or strongly suggested using the TM. Of the remaining 16 studies, 9 (18.3%) did not report a procedure, 1 (2%) reported a procedure but we could not determine which technique was used based on the description, and 6 (12.2%) used neither the BEP nor TM. If we consider that (1) it is likely that some, if not all, of the studies that did not report a procedure probably used the BEP and (2) studies that used a combination of measuring minimum frequencies by eye on spectrograms with the aid of power spectra are equivalent to the BEP in practice because they did not use a pre-set threshold value, then our determination that 40.8% of the literature surveyed used the BEP is likely an underestimate. Our survey shows that the BEP was utilized more frequently than any other approach (including in some of our own studies).

## Songs

We found a significant effect of noise treatment on minimum frequency when measured with the BEP (one-way repeated measure ANOVA  $F_{9,2} = 26.4$ ,  $P < 0.001$ ; Fig. 4a). The Holm-Sidak post hoc test showed that measures of song minimum frequency were higher for the medium noise treatment than for the ambient noise treatment ( $P = 0.012$ ) and highest during the high noise treatment when compared to the ambient and medium noise treatments ( $P < 0.001$ ). The increase in song minimum frequency, as measured with the BEP, was  $239 \pm 94$  Hz between the medium and ambient noise levels,  $375 \pm 81$  between the high and medium noise levels, and  $615 \pm 78$  between the high and low noise levels, respectively. In contrast, we found no effect of noise on measured minimum frequency for data collected with the TM (one-way repeated measure ANOVA  $F_{9,2} = 0.806$ ,  $P = 0.462$ ; Fig. 4b). Results from multiple paired  $t$  tests show that the TM yielded significantly higher minimum frequencies than the BEP; this was true for the ambient noise, medium noise, and high noise treatments (ambient noise BEP =  $845 \pm 95$  Hz, TM =  $1867 \pm 80$  Hz,  $t = -11.95$ ,  $df = 9$ ,  $P < 0.001$ ; medium noise BEP =  $1085 \pm 84$  Hz, TM =  $1823 \pm 117$  Hz,  $t = -7.71$ ,  $df = 9$ ,  $P < 0.001$ ; high noise BEP =  $1461 \pm 61$  Hz, TM =  $1836 \pm 109$  Hz,  $t = -4.67$ ,  $df = 9$ ,  $P = 0.001$ ).

There was also a significant effect of noise treatment on the measures of song duration (one-way repeated measure ANOVA  $F_{9,2} = 5.79$ ,  $P = 0.011$ ; Fig. 5). The Holm-Sidak post hoc test revealed that the measure of song duration did not differ between the ambient noise and the medium noise levels ( $P = 0.996$ ), but it was significantly shorter for the high noise level treatment when compared to both the medium and ambient noise treatments ( $P = 0.009$  in both; Fig. 5). The mean ( $\pm$ SE) difference in measured song duration was  $0.0001 \pm 0.01$  s between the medium and ambient noise treatment,  $0.06 \pm 0.02$  s between the high and medium noise



**Fig. 4** Minimum frequency of songs as measured with the BEP (a) and TM (b) in different noise conditions. The minimum frequency values measured for blackbird songs increased with noise when evaluated with the BEP but not when measured with the TM. Data points linked with a

line represent the values for the same vocalization measured under different noise conditions. *Ambient* = ambient noise level, *medium* = medium noise level, *high* = high noise level

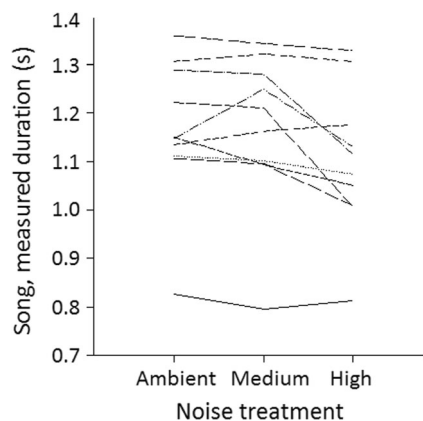
treatment, and  $0.06 \pm 0.02$  s between the high and ambient noise treatment. Noise had no effect on measures of song  $Q_1$ , song  $Q_3$ , song Freq5%, and song Freq95% (Friedman test for related samples song  $Q_1$ ,  $\chi^2 = 1.2$ ,  $df = 2$ ,  $P = 0.549$ ; song  $Q_3$ ,  $\chi^2 = 2.867$ ,  $df = 2$ ,  $P = 0.239$ ; song Freq5%,  $\chi^2 = 1.27$ ,  $df = 2$ ,  $P = 0.529$ ; song Freq95%,  $\chi^2 = 2.92$ ,  $df = 2$ ,  $P = 0.232$ ).

## Calls

### Checks

Noise had an effect on measurements of check minimum frequency when the BEP was used (one-way repeated measure ANOVA  $F_{9,2} = 135.97$ ,  $P < 0.001$ ). The Holm-Sidak test showed that measured minimum frequencies were higher during the high noise treatment than during either the ambient noise treatment ( $P < 0.001$ ) or the medium noise treatment

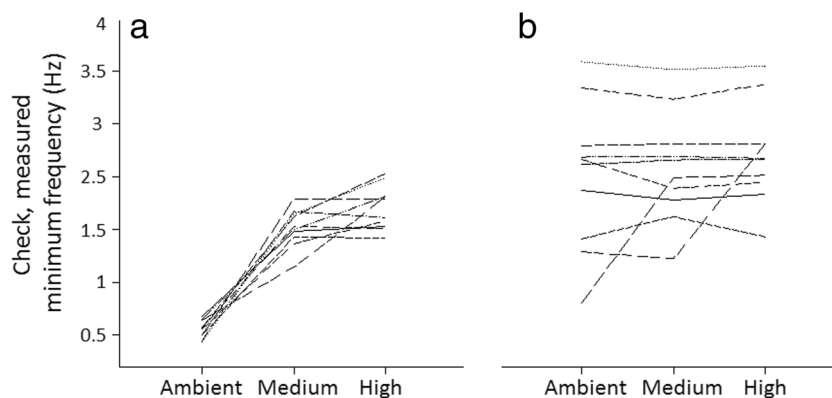
( $P = 0.021$ ) and also higher during the medium noise treatment when compared to the ambient noise treatment ( $P < 0.001$ ; Fig. 6a). Given that the same set of vocalizations was measured for each noise treatment, these results only emerge as an artifact of measurement bias. The increase in the average minimum frequency of checks, as measured with the BEP, was  $963 \pm 76$  Hz between the medium and ambient noise treatments,  $189 \pm 75$  Hz between the high and medium noise levels, and  $1152 \pm 72$  Hz between the high and low noise levels, respectively. Measurements of the minimum frequency for checks did not change with noise levels when measured with the TM (one-way repeated measure ANOVA  $F_{9,2} = 1.19$ ,  $P = 0.326$ ; Fig. 6b). Results from paired  $t$  tests show that measures of check minimum frequency were significantly higher when measured with the TM than with the BEP in the ambient, medium, and high noise treatments (ambient noise BEP =  $556 \pm 26$  Hz, TM =  $2007 \pm 219$  Hz,  $t = -6.4$ ,  $df = 9$ ,  $P < 0.001$ ; medium noise BEP =  $1519 \pm 57$  Hz, TM =  $2092 \pm 164$  Hz,  $t = -4.2$ ,  $df = 9$ ,  $P = 0.002$ ; high noise BEP =  $1709 \pm 65$  Hz, TM =  $2211 \pm 150$  Hz,  $t = -3.8$ ,  $df = 9$ ,  $P = 0.004$ ).



**Fig. 5** Measured song duration using spectrograms in different noise conditions. The measure of song duration decreased as noise increased. Three data points linked with a line represent the values collected for the same vocalization. *Ambient* = ambient noise level, *medium* = medium noise level, *high* = high noise level

Measures of check duration, check  $Q_1$ , check  $Q_3$ , and check Freq95% were not significantly different among noise treatments (check duration, Friedman test for related samples  $\chi^2 = 4.2$ ,  $df = 2$ ,  $P = 0.122$ ; check  $Q_1$ , one-way repeated measure ANOVA  $F_{9,2} = 0.438$ ,  $P = 0.652$ ; check  $Q_3$ , Friedman test for related samples,  $\chi^2 = 5.03$ ,  $df = 2$ ,  $P = 0.081$ ; check 95%, Friedman test for related samples,  $\chi^2 = 0.5$ ,  $df = 2$ ,  $P = 0.779$ ). The mean  $\pm$  SE difference in measured check duration was  $0.02 \pm 0.01$  between the medium and ambient noise treatments,  $0.003 \pm 0.01$  between the high and medium noise levels, and  $0.01 \pm 0.004$  between the high and ambient noise treatment. However, the measure for check Freq5% differed between noise treatments (one-way repeated measure ANOVA  $F_{9,2} = 8.66$ ,  $P = 0.002$ ). The Holm-Sidak test showed that the measure of Freq5% was

**Fig. 6** Minimum frequency of checks as measured with the BEP (a) and TM (b) in different noise conditions. Check minimum frequencies increased when measured with the BEP in noise but not when measured with the TM. Three data points linked with a line represent the values collected for the same vocalization. *Ambient* = ambient noise level, *medium* = medium noise level, *high* = high noise level



significantly higher during the high noise treatment than both the ambient noise treatment ( $P < 0.001$ ) and the medium noise treatment ( $P = 0.016$ ); the measure did not change between the medium and ambient noise levels ( $P = 0.168$ ).

### Cheers

When we measured cheers (which are relatively higher in pitch compared to songs and checks), we found that their minimum frequency did not significantly change with noise levels when it was measured using the BEP (one-way RM ANOVA,  $F_{9,2} = 2.90$ ,  $P = 0.080$ ; Fig. 7a) or the TM (one-way repeated measure ANOVA,  $F_{9,2} = 2.26$ ,  $P = 0.133$ ; Fig. 7b). Additionally, measurements of minimum frequency did not differ between the BEP and TM during the ambient noise treatment (BEP  $3784 \pm 219$  Hz, TM  $3759 \pm 211$  Hz, paired  $t$  test,  $t = 1.75$ ,  $df = 9$ ,  $P = 0.113$ ), medium noise level treatment (BEP  $3775 \pm 213$  Hz, TM  $3772 \pm 215$  Hz, paired  $t$  test,  $t = 0.44$ ,  $df = 9$ ,  $P = 0.67$ ), or high noise treatment (BEP  $3753 \pm 223$  Hz, TM  $3764 \pm 216$  Hz, paired  $t$  test,  $t = -1.07$ ,  $df = 9$ ,  $P = 0.310$ ).

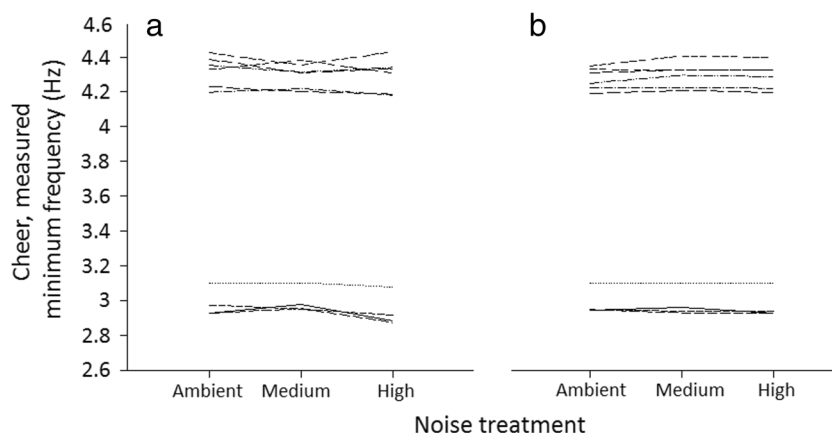
For cheer calls, we did not find any significant differences in the other acoustic measurements we made relative to noise levels (cheer duration, Friedman test for related samples  $\chi^2 = 5.4$ ,  $df = 2$ ,  $P = 0.067$ ; cheer  $Q_1$ , Friedman test for related

samples  $\chi^2 = 3$ ,  $df = 2$ ,  $P = 0.223$ ; cheer  $Q_3$ , one-way repeated measure ANOVA  $F_{9,2} = 1.34$ ,  $P = 0.286$ ; cheer Freq5%, Friedman test for related samples  $\chi^2 = 4$ ,  $df = 2$ ,  $P = 0.135$ ; cheer Freq95%, one-way repeated measure ANOVA  $F_{9,2} = 1.94$ ,  $P = 0.172$ ). Mean  $\pm$  SE difference in measured cheer duration was  $0.01 \pm 0.01$  between medium and ambient noise treatments,  $0.003 \pm 0.01$  between high and medium noise levels, and  $0.01 \pm 0.004$  between high and ambient noise levels.

### Discussion

This study compares the performance of two techniques used to measure minimum frequency under varying levels of anthropogenic noise—measuring by eye from spectrograms (the BEP) and using preset thresholds on power spectra (the TM). To do so, we played back the same set of vocalizations while simultaneously playing back recordings of traffic noise at different levels, then measured the re-recorded vocalizations using both procedures. We found that, for most vocalization types, the BEP yielded increasing measures of minimum frequency when noise increased despite there being no actual increases in the minimum frequency of these vocalizations. In other words, we detected a positive relationship between

**Fig. 7** Minimum frequency of cheers as measured with the BEP (a) and TM (b) in different noise conditions. The measure of cheer minimum frequency did not change with noise regardless of the approach used. *Ambient* = ambient noise level, *medium* = medium noise level, *high* = high noise level





minimum frequency and noise, similar to the relationship found in many recent studies, as an artifact of the measurement technique. We found that the potential for noise-related bias increased when we used the BEP to measure vocalizations that overlapped in frequency with our highway noise playback, which is loudest in lower frequencies (see Fig. 2). Thus, we detected strong bias in measures of songs and checks, which have low minimum frequencies (mean in ambient noise with BEP, songs 845 Hz, checks 556 Hz). However, we found no evidence of bias in measures of cheers, which have a higher minimum frequency (mean in ambient noise with the BEP, 3784 Hz) and other acoustic features, discussed below, which make them less prone to masking and therefore easier to measure by eye. Grace and Anderson (2015) conducted a similar study using synthetic sounds resembling Carolina chickadee (*Poecile carolinensis*) calls. Although these sounds were higher pitched (with minimum frequencies of >1300 Hz) than our checks and songs, Grace and Anderson also found a noise-related bias when they measured minimum frequency by eye. They reported a 1750-Hz increase in measures in noise (Fig. 3 in Grace and Anderson 2015) compared to our 615-Hz mean maximum increase in measures of songs and 1152-Hz mean maximum increase in measures of checks in noise. Taken together, these studies suggest that while the BEP is intuitive and quick, the obfuscation of low frequencies by masking noise may make it difficult to accurately detect the minimum frequency of the vocalization and may therefore yield false positives in studies looking for a relationship between the minimum frequency of vocalizations and the amplitude of background noise.

In addition to bias in measures of minimum frequency, we also found that noise had a small, but significant, effect on the measures of song duration using the BEP, where duration measures decreased as noise levels increased (mean  $\pm$  SE maximum difference in song duration between treatments was  $0.06 \pm 0.02$  s). This supports the contention that background noise may influence temporal measures from spectrograms (Zollinger et al. 2012). This bias appeared to be largest when the first and last syllables of the song were low in frequency and/or had relatively narrow bandwidths at the onset or offset (see Fig. 3). Low frequencies and narrow bandwidths may make it more difficult for the observer to detect the beginning and end of the vocalization in a noisy recording. On the other hand, we found no effect of noise on the measured duration of checks; this call type has a comparatively wider bandwidth and thus can be detected above the noise (Fig. 3). We also found a non-significant tendency ( $P = 0.067$ ) for measures of cheers to be shorter in duration as noise increases (mean  $\pm$  SE maximum difference in cheer duration between treatments was  $0.01 \pm 0.01$  s). Cheers have relatively narrow bandwidths and are above most of the noise; however, the

gradual offset of amplitude at the tail may have made the duration more difficult to detect. To reduce this type of bias in temporal measures, recordings can be cleaned by filtering and waveforms can be used to better visualize the onset and offset of vocalizations (Zollinger et al. 2012).

We did not find an effect of noise on the other frequency measures that we examined using the BEP, with the exception of check Freq5%; this measure increased as noise increased. If this result was due to the direct effect of noise on the recording, then we would expect to find Freq5% measures decreasing with increasing noise levels since the additional noise would increase sound energy more at lower frequencies. Instead, we found the opposite pattern. Therefore, we believe that our results are better explained by noise-related bias introduced by the BEP. The measured minimum frequency of check calls increased with noise, which could explain the correlated increase in check Freq5%. Thus, studies addressing the relationship between low-frequency background noise and the distribution of energy in the lower part of a vocalization may find misleading results when using the BEP if there is an artificial positive association between noise level and measures of the minimum frequency of the vocalization.

In contrast with the BEP, we found that the TM yielded more consistent estimates of minimum frequency that were not biased by background noise, which is consistent with expectations for the TM (Zollinger et al. 2012). Using the TM, one will not find spurious correlations between minimum frequency and noise because only the components of the vocalization that would be detectable in the noisiest conditions of the study are measured. However, this method may miss the lowest components of some vocalizations. Indeed, we found that the TM yielded significantly higher minimum frequency measures than the BEP in songs and checks. However, these low-frequency components falling below the threshold likely contain very little of the total signal energy and it is unclear whether they would be meaningful in animal communication (Gil and Brumm 2014). When measuring cheers, however, the BEP and TM performed equally well (i.e., the measures of minimum frequency did not change with noise and did not differ between measurement procedures). Cheers are very loud (95 dB; Patricelli et al. 2007), very tonal, and higher in frequency than most of the energy of our noise playbacks, making the minimum frequency easy to detect visually using the BEP. In addition, these acoustic features allowed us to use a higher, less-conservative threshold to measure the minimum frequency of cheers (30 dB), whereas the poorer signal-to-noise ratio of songs and checks necessitated a relatively low, more-conservative threshold (15 dB). These results suggest that studies in which the threshold value must be conservative may exclude the lowest notes of relatively low-frequency vocalizations. Because many studies of

vocal adjustment in noise have found that birds increase their minimum frequency, failure to measure the lowest frequency notes may lead researchers to conclude that no vocal adjustment occurs, when in fact such adjustments are present. To reduce the likelihood of such a false negative, the recordings with the lowest signal-to-noise ratio could be dropped from the study, allowing one to use a higher threshold. This, however, may reduce overall sample sizes and exclude from the analysis important and interesting cases where background noise is most intense. The decision about whether or not to exclude noisier recordings to increase the threshold will depend on the goals of the study. Our goal here was to compare the BEP and TM across a controlled range of noise conditions, so we did not exclude any recordings from our analyses. To inform such decisions, playback studies such as ours can be used to test for the risk of detection errors from the TM (as well as false positives from the BEP) for vocalizations in a species of interest (Grace and Anderson 2015). Note that methods which allow adjustment of the threshold based on the noise levels in different recordings risk the same bias we detected with the BEP, as researchers (or algorithms) may be unable to detect vocalizations masked by noise on power spectra.

Our results indicate that caution is required when interpreting results that have used the BEP to measure low-frequency vocalizations and have found a positive association between noise and minimum frequency. Our survey (see [Supplementary material](#)) found that most studies addressing the influence of anthropogenic noise on avian vocalizations have used the BEP to measure minimum frequency. Our findings raise the possibility that some of these studies may have overestimated the effects of environmental noise on vocalization pitch. While this bias may lead researchers to incorrectly identify an adaptive response in bird vocalization in response to noise exposure, several studies that have found a positive association between the minimum frequency of vocalizations and noise have used the TM (e.g., Hu and Cardoso 2009; Mockford and Marshall 2009; Luther and Derryberry 2012; Potvin and Mulder 2013; Luther et al. 2015), while others also using the TM have not found this association (Hanna et al. 2011; Ríos-Chelén et al. 2013; Cartwright et al. 2014; Gough et al. 2014; Ríos-Chelén et al. 2015). Therefore, we are not arguing that the observed relationships between noise levels and the minimum frequencies of bird vocalizations are always an artifact. Rather, we argue that caution and validation are required due to the potential for artifacts in common techniques for measurement.

In conclusion, we showed that the BEP is prone to noise-related bias, especially when measuring vocalizations with low minimum frequencies in noisy conditions.

The TM can provide more consistent minimum frequency measurements but may omit some lower-frequency components of vocalizations when the signal-to-noise ratio is poor. Thus, the TM may fail to detect shifts in minimum frequency even when such shifts occur. However, for high-pitched and tonal vocalizations that are less likely to be masked by noise, such as red-winged blackbird cheers, both methods yield comparable results. However, we still encourage using the TM instead of the BEP in these cases, because the BEP is an intrinsically subjective procedure, prone to other sources of error (e.g., variation in spectrogram settings or recording levels), which could compromise our ability to make comparisons among recordings and studies (Zollinger et al. 2012). These results also emphasize the importance of high signal-to-noise recordings, though collecting such recordings may not always be possible when recording animals in the field—especially when the goal of the study is to address how animals respond to high noise conditions. In some species, the structure of vocalizations may be sufficiently consistent that minimum frequency can be accurately inferred by measuring less masked components of the vocalization. Regardless of the metric or procedure used, we strongly encourage the use of playbacks of the vocalizations of interest under different noise conditions to assess the possibility of bias and detection errors (e.g., Grace and Anderson 2015). In the absence of careful validation, researchers run the risk of detecting both false positives and false negatives in studies of noise impacts on animal signals.

**Acknowledgments** We thank the Wetlands reserve for allowing us to work in the Wetlands. We thank members of the Patricelli Lab, especially Allison Injaian, for the feedback and help in the field. We thank Carole Hom, Rick Grosberg, and the UC Davis Evolution and Ecology Graduate Admissions Pathways program for bringing ANM to UC Davis for a summer and providing excellent feedback and training. We thank Laura Corkovic for the help with translation during the literature survey. We thank the reviewers and editor who improved this paper with their valuable comments.

#### Compliance with ethical standards

**Funding** AAR-CH was supported by a UC MEXUS-CONACYT fellowship, and ANM was supported by the UC Davis Evolution and Ecology Graduate Admissions Pathways program. Support for the data collection was provided by UC MEXUS-CONACYT to AAR-CH, by UC Davis and National Science Foundation grant (IOS-1258217) to GLP and AHK, and an NSF Graduate Research Fellowship to ACP (DGE-1148897).

**Conflict of interest** The authors declare that they have no conflict of interest.

**Informed consent** Not applicable.

**Ethical approval** This article does not contain any studies with human participants or animals performed by any of the authors.

## References

- Azar JF, Bell BD, Borowiec M (2014) Temporal change of the song of a local population of the Grey warbler (*Gerygone igata*): has its song changed over time? *Emu* 114:80–85
- Beecher MD (1988) Spectrographic analysis of animal vocalizations: implications of the ‘uncertainty principle’. *Bioacoustics* 1:187–208
- Bermúdez-Cuamatzin E, Ríos-Chelén AA, Gil D, Macías García C (2009) Strategies of song adaptation to urban noise in the house finch: syllable pitch plasticity or differential syllable use? *Behaviour* 146:1269–1286
- Bermúdez-Cuamatzin E, Ríos-Chelén AA, Gil D, Macías García C (2011) Experimental evidence for real-time song frequency shift in response to urban noise in a passerine bird. *Biol Lett* 7:36–38
- Brumm H, Slabbekoom H (2005) Acoustic communication in noise. *Adv Stud Behav* 35:151–209
- Brumm H, Slater PJB (2006) Ambient noise, motor fatigue, and serial redundancy in chaffinch song. *Behav Ecol Sociobiol* 60:475–481
- Brumm H, Zollinger SA (2013) Avian vocal production in noise. In: Brumm H (ed) *Animal communication and noise. Animal signals and communication, vol 2*. Springer-Verlag, Berlin, pp. 187–227
- Cardoso GC, Atwell JW (2011a) On the relation between loudness and the increased song frequency of urban birds. *Anim Behav* 82:831–836
- Cardoso GC, Atwell JW (2011b) Directional cultural change by modification and replacement of memes. *Evolution* 65:295–300
- Cardoso GC, Atwell JW (2012) On amplitude and frequency in birdsong: a reply to Zollinger et al. *Anim Behav* 84:e10–e15
- Cartwright LA, Taylor DR, Wilson DR, Chow-Fraser P (2014) Urban noise affects song structure and daily patterns of song production in red-winged blackbirds (*Agelaius phoeniceus*). *Urban Ecosyst* 17: 561–572
- Derryberry EP, Danner RM, Danner JE, Derryberry GE, Phillips JN, Lipschutz SE, Gentry K, Luther DA (2016) Patterns of song across natural and anthropogenic soundscapes suggest that white-crowned sparrows minimize acoustic masking and maximize signal content. *PLoS One* 11:e0154456
- Dowling JL, Luther DA, Marra PP (2012) Comparative effects of urban development and anthropogenic noise on bird songs. *Behav Ecol* 23:201–209
- Fernández-Juricic E, Poston R, De Collibus K, Morgan T, Bastain B, Martin C, Jones K, Treminio T (2005) Microhabitat selection and singing behavior patterns of male house finches (*Carpodacus mexicanus*) in urban parks in a heavily urbanized landscape in the western U.S. *Urban Habitats* 3:49–69
- Francis CD, Ortega CP, Cruz A (2011a) Vocal frequency change reflects different responses to anthropogenic noise in two subspecies tyrant flycatchers. *Proc R Soc Lond B* 278:2025–2031
- Francis CD, Ortega CP, Cruz A (2011b) Different behavioural responses to anthropogenic noise by two closely related passerine birds. *Biol Lett* 7:850–852
- Gil D, Brumm H (2014) Acoustic communication in the urban environment: patterns, mechanisms, and potential consequences of avian song adjustments. In: Gil D, Brumm H (eds) *Avian urban ecology*. Oxford University Press, Oxford, pp. 69–83
- Giraudeau M, Nolan PM, Black CE, Earl SR, Hasegawa M, McGraw KJ (2014) Song characteristics track bill morphology along a gradient of urbanization in house finches (*Haemorhous mexicanus*). *Front Zool* 11:83
- Gough DC, Mennill DJ, Nol E (2014) Singing seaside: Pacific wrens (*Troglodytes pacificus*) change their songs in the presence of natural and anthropogenic noise. *Wilson J Ornithol* 126:269–278
- Grace MK, Anderson RC (2015) No frequency shift in the “D” notes of Carolina chickadee calls in response to traffic noise. *Behav Ecol Sociobiol* 69:253–263
- Gross K, Pasinelli G, Kunc HP (2010) Behavioral plasticity allows short-term adjustment to a novel environment. *Am Nat* 176:456–464
- Halfwerk W, Bot S, Buikx J, van der Velde M, Komdeur J, ten Cate C, Slabbekoom H (2011) Low-frequency songs lose their potency in noisy urban conditions. *P Natl Acad Sci USA* 108:14549–14554
- Hamao S, Watanabe M, Mori Y (2011) Urban noise and male density affect songs in the great tit *Parus major*. *Ethol Ecol Evol* 23:111–119
- Hanna D, Blouin-Demers G, Wilson DR, Mennill DJ (2011) Anthropogenic noise affects song structure in red-winged blackbirds (*Agelaius phoeniceus*). *J Exp Biol* 214:3549–3556
- Hu Y, Cardoso GC (2009) Are bird species that vocalize at higher frequencies preadapted to inhabit noisy urban areas? *Behav Ecol* 20: 1268–1273
- Hu Y, Cardoso GC (2010) Which birds adjust the frequency of vocalizations in urban noise? *Anim Behav* 79:863–867
- Ivanitskii VV, Antipov VA, Marova IM (2015) The thrush nightingale (*Luscinia luscinia*) in Moscow and Moscow suburbs: city noise influences the frequency parameters of its song. *Biol Bull* 42:724–727
- Lenz AK, La VT (2014) White-throated sparrows alter songs differentially in response to chorusing anurans and other background noise. *Behav Process* 105:28–35
- León E, Beltzer A, Quiroga M (2014) The saffron finch (*Sicalis flaveola*) modifies its vocalizations to adapt to urban habitats. *Rev Mex Biodivers* 85:546–552
- Leonard ML, Horn AG (2008) Does ambient noise affect growth and begging call structure in nestling birds? *Behav Ecol* 19:502–507
- Leonard ML, Horn AG, Oswald KN, McIntyre E (2015) Effect of ambient noise on parent–offspring interactions in tree swallows. *Anim Behav* 109:1–7
- Luther D, Baptista L (2010) Urban noise and the cultural evolution of bird songs. *Proc R Soc Lond B* 277:469–473
- Luther DA, Derryberry EP (2012) Birdsongs keep pace with city life: changes in song over time in an urban songbird affects communication. *Anim Behav* 83:1059–1066
- Luther DA, Phillips J, Derryberry EP (2015) Not so sexy in the city: urban birds adjust songs to noise but compromise vocal performance. *Behav Ecol* 27:332–340
- McLaughlin KE, Kunc HP (2013) Experimentally increased noise levels change spatial and singing behaviour. *Biol Lett* 9:20120771
- Mendes S, Colino-Rabanal VJ, Peris SJ (2011) Bird song variations along an urban gradient: the case of the European blackbird (*Turdus merula*). *Landscape Urban Plan* 99:51–57
- Mockford EJ, Marshall RC (2009) Effects of urban noise on song and response behavior in great tits. *Proc R Soc Lond B* 276:2979–2985
- Montague MJ, Danek-Gontard M, Kunc JP (2012) Phenotypic plasticity affects the response of a sexually selected trait to anthropogenic noise. *Behav Ecol* 24:343–348
- Narango DL, Rodewald AD (2016) Urban-associated drivers of song variation along a rural–urban gradient. *Behav Ecol* 27:608–616
- Nemeth E, Brumm H (2010) Birds and anthropogenic noise: are urban songs adaptive? *Am Nat* 176:465–475
- Nemeth E, Zollinger SA, Brumm H (2012) Effect sizes and the integrative understanding of urban bird song. *Am Nat* 180:146–152
- Nemeth E, Pieretti N, Zollinger SA, Geberzahn N, Partecke J, Miranda AC, Brumm H (2013) Bird song and anthropogenic noise: vocal constraints may explain why birds sing higher-frequency songs in cities. *Proc R Soc B* 280:20122798
- Oden AI, Brown MB, Burbach ME, Brandle JR, Quinn JE (2015) Variation in avian vocalizations during the non-breeding season in response to traffic noise. *Ethology* 121:472–479

- Patricelli GL, Blickley JL (2006) Avian communication in urban noise: causes and consequences of vocal adjustment. *Auk* 123:639–649
- Patricelli GL, Dantzker MS, Bradbury JW (2007) Differences in acoustic directionality among vocalizations of the male red-winged blackbird (*Agelaius phoeniceus*) are related to function in communication. *Behav Ecol Sociobiol* 61:1099–1110
- Podos J (1997) A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution* 51:537–551
- Podos J (2001) Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature* 409:185–188
- Pohl NU, Leadbeater E, Slabbekoorn H, Klump GM, Langemann U (2012) Great tits in urban noise benefit from high frequencies in song detection and discrimination. *Anim Behav* 83:711–721
- Potvin DA, MacDougal-Shackleton SA (2015) Traffic noise affects embryo mortality and nestling growth rates in captive zebra finches. *J Exp Zool* 323:722–730
- Potvin DA, Mulder RA (2013) Immediate, independent adjustment of call pitch and amplitude in response to varying background noise by silvereyes (*Zosterops lateralis*). *Behav Ecol* 24:1363–1368
- Potvin DA, Parris KM, Mulder RA (2011) Geographically pervasive effects of urban noise on frequency and syllable rate of songs and calls in silvereyes (*Zosterops lateralis*). *Proc R Soc Lond B* 278:2464–2469
- Potvin DA, Mulder RA, Parris KM (2014) Silvereyes decrease acoustic frequency but increase efficacy of alarm calls in urban noise. *Anim Behav* 98:27–33
- Redondo P, Barrantes G, Sandoval L (2013) Urban noise influences vocalization structure in the house wren *Troglodytes aedon*. *Ibis* 155:621–625
- Ríos-Chelén AA (2009) Birdsong: the interplay between urban noise and sexual selection. *Oecol Brasil* 13:153–164
- Ríos-Chelén AA, Salaberria C, Barbosa I, Macías Garcia C, Gil D (2012) The learning advantage: bird species that learn their song show a tighter adjustment of song to noisy environments than those that do not learn. *J Evol Biol* 25:2171–2180
- Ríos-Chelén AA, Quirós-Guerrero E, Gil D, Macías Garcia C (2013) Dealing with urban noise: vermilion flycatchers sing longer songs in noisier territories. *Behav Ecol Sociobiol* 67:145–152
- Ríos-Chelén AA, Lee GC, Patricelli GL (2015) Anthropogenic noise is associated with changes in acoustic but not visual signals in red-winged blackbirds. *Behav Ecol Sociobiol* 69:1139–1151
- Ríos-Chelén AA, Lee GC, Patricelli GL (2016) A comparison between two ways to measure minimum frequency and an experimental test of vocal plasticity in red-winged blackbirds in response to noise. *Behaviour* (published online, doi:10.1163/1568539X-00003390)
- Salaberria C, Gil D (2010) Increase in song frequency in response to urban noise in the great tit *Parus major* as shown by data from the Madrid (Spain) city noise map. *Ardeola* 57:3–11
- Seeger-Fullam KD, Rodewald AD, Soha JA (2011) Urban noise predicts song frequency in northern cardinals and American robins. *Bioacoustics* 20:267–276
- Slabbekoorn H (2013) Songs of the city: noise-dependent spectral plasticity in the acoustic phenotype of urban birds. *Anim Behav* 85:1089–1099
- Slabbekoorn H, den Boer-Visser A (2006) Cities change the songs of birds. *Curr Biol* 16:2326–2331
- Slabbekoorn H, Peet M (2003) Birds sing at higher pitch in urban noise. *Nature* 426:267–267
- Slabbekoorn H, Smith TB (2002) Habitat-dependent song divergence in the little greenbul: an analysis of environmental selection pressures on acoustic signals. *Evolution* 56:1849–1858
- Slabbekoorn H, Yeh P, Hunt K (2007) Sound transmission and song divergence: a comparison of urban and forest acoustics. *Condor* 109:67–78
- Slabbekoorn H, Yang X-J, Halfwerk W (2012) Birds and anthropogenic noise: singing higher may matter. *Am Nat* 180:142–145
- Verzijden MN, Ripmeester EAP, Ohms VR, Snelderwaard P, Slabbekoorn H (2010) Immediate spectral flexibility in singing chiffchaffs during experimental exposure to highway noise. *J Exp Biol* 213:2575–2581
- Wood WE, Yezerinac SM (2006) Song sparrow (*Melospiza melodia*) song varies with urban noise. *Auk* 123:650–659
- Zollinger SA, Podos J, Nemeth E, Goller F, Brumm H (2012) On the relationship between, and measurement of, amplitude and frequency in bird song. *Anim Behav* 84:e1–e9