

Behaviour 161 (2024) 167-194

Vocal performance of one affects that of another

Keith W. Sockman^{a,*}, Susan M. Lyons^b and Samuel P. Caro^c

 ^a Department of Biology, University of North Carolina at Chapel Hill, Coker Hall, 120 South Road, Chapel Hill, NC 27599, USA
^b Great Minds PBC, 55 M Street SE, Suite 340, Washington, DC 20003, USA
^c CEFE, University of Montpellier, CNRS, EPHE, IRD, 1919 Route de Mende, 34293 Montpellier cedex 5, France
*Corresponding author's e-mail address: kws@unc.edu ORCID iDs: Sockman: 0000-0002-3846-5386; Lyons: 0009-0002-2591-9868; Caro: 0000-0002-5405-7753

Received 23 June 2023; initial decision 30 August 2023; revised 20 September 2023; accepted 21 December 2023; published online 9 January 2024

Abstract

The trill elements of many bird species' songs have been hypothesized as honest signals of performance. However, the breadth of receiver responses to variation in the signaller's trill performance is unknown. We exposed wild male Lincoln's sparrows (*Melospiza lincolnii*) to two song treatments, one with low- and another with high-performance trills. We found no effect of treatment on measures of agonistic behaviour and song length. However, relative to the subjects' trills following the high-performance treatment, those following the low-performance treatment were elevated in performance due to trill types with high frequency bandwidth in the third trill of songs. Treatment also affected trill duration through its syllable count in a manner that varied by the song's trill number. Thus, the performance of a signal to which a receiver is exposed drives plasticity in his own performance in sequence-specific manner. Males may showcase their own performance in the presence of lower-performing rivals.

Keywords

bird song, intraspecific competition, Lincoln's sparrow (*Melospiza lincolnii*), sexual signal, simulated territorial intrusion, trill performance.

1. Introduction

Biological signals evolve for the purpose of communication, and therefore they are often conspicuous, at least to the intended receiver. A signal must nevertheless fulfil a set of criteria in order to be distinguished from other non-signalling traits. Among those criteria is the requirement that it elicits a reliable response in receivers that, on average, is adaptive for the signaller and which may be adaptive for the receiver, as well (Maynard Smith & Harper, 1995; Laidre & Johnstone, 2013). Even for known signals, the function may not always be what it seems, as some signals can have multiple purposes and be context dependent (Laidre & Johnstone, 2013).

The songs of male birds are widely considered to be signals of mate attraction (McGregor, 1991; Sockman et al., 2005), aggressive intent (Searcy & Beecher, 2009; Linhart et al., 2013), and identity (Nelson & Poesel, 2007). Depending on species, individuals may vary in their singing effort (e.g., song rate, duration, or amplitude) (Ritschard et al., 2010; Nishida & Takagi, 2018), song complexity (variation within and between songs) (Mountjoy & Lemon, 1991; Leitão et al., 2006), and other aspects of their song, leading to variation in attracting mates (Gentner & Hulse, 2000), in thwarting rivals (de Kort et al., 2009), and in fitness (Andersson, 1994; Catchpole & Slater, 1995). In the last 26 years (Podos, 1997), an aspect of song known as trill performance (also called song performance, vocal performance, or vocal deviation; see explanation below), has come under focus regarding its potential as an honest signal of male quality (e.g., Phillips & Derryberry, 2017b; Nishida & Takagi, 2018) that could drive variation between individuals in reproductive success, possibly even leading to sympatric speciation (Podos, 2001). However, the behavioural significance of trill performance is still not thoroughly understood (Kroodsma, 2017; but see Podos, 2017; Vehrencamp et al., 2017), in part because it is not clear the breadth to which natural variation in trill performance can affect receivers and how the adaptive significance of the signal might vary between different types of receivers.

Trills are rapid repetitions of a single type of song-syllable, and many songbird species produce them. In comparisons both within and between species, the maximum rate of syllable production, often referred to as trill rate and here referred to as syllable rate, is negatively correlated with the trill's frequency bandwidth (Podos, 1997; Wilson et al., 2014). This relationship is thought to result from biomechanical constraints on the vocal tract, in which syllable rate is limited by the degree of conformational change required by the skeletal-muscular components necessary to produce the syllable's frequency range (Podos & Nowicki, 2004; Ballentine, 2009). Trills

that maximize syllable rate relative to bandwidth or that maximize bandwidth relative to syllable rate are said to be high performance (Ballentine et al., 2004; Phillips & Derryberry, 2017a; but see Cardoso et al., 2007; Cardoso, 2017).

Since the initial description of a possible trade-off between syllable rate and frequency bandwidth (Podos, 1997) and the realization of trill performance as a potential honest sexual signal (Dubois et al., 2011), numerous researchers have sought to examine its biological significance. They have done so both through observational investigations of its relationship with measures of the singer's body condition (Ballentine, 2009) and reproductive success (Nishida & Takagi, 2018) and through experimental approaches, in which subjects are measured for their aggressive (Illes et al., 2006; Cramer, 2013; Goodwin & Podos, 2014; Phillips & Derryberry, 2017a) or preference responses (Caro et al., 2010; Lyons et al., 2014) to manipulations of song playbacks that vary in trill performance. But one question that has yet to be satisfactorily addressed is how this vocal signal affects the vocal responses of a receiver and, in particular, whether the trill performance of one individual affects that of another.

Some effects of variation in song signals on receiver vocal output are well established. Probably the most obvious example of this is when songs are learned by the individual singer, a process thought to characterize the roughly 5000 species of oscine songbirds, which produce species-typical songs as a function of their prior song exposure during development (Marler & Peters, 1977; Catchpole & Slater, 1995). Another example is in the case of songtype matching, whereby a male biases his own production of song toward the types produced by his neighbour (Vehrencamp, 2001; Logue & Forstmeier, 2008). In addition, the quality of songs to which an adult is exposed can affect his singing effort in both the short term and longer term (Sockman et al., 2009; Sewall et al., 2010). But little is known with regard to how variation in the trill performance to which an individual is exposed affects his own song (but see Cramer, 2013; Moseley et al., 2013; Liu et al., 2018). In fact, the predicted direction of the receiver's response is not even obvious based on signalling theory (Searcy & Nowicki, 2000; Collins, 2004). Should the elevated threat of a high-performance trill suppress or instead excite the (song) behaviour of territorial receivers? Illes and colleagues (2006) found that male banded wrens (Thryothorus pleurostictus) more rapidly approach songs with high-performance trills than with low-performance trills, and yet they 170

spend less time in close proximity to the high-performance songs when those songs approach the theoretical performance limit. Male swamp sparrows (*Melospiza georgiana*) respond more aggressively to control songs (with intermediate trill performance) than songs with digitally reduced trill performance, but they are also more aggressive, on average, to control songs than songs with digitally enhanced trill performance, a pattern explained, in part, by the subject's own vocal performance (Moseley et al., 2013). In the present article, we report on a study in which we analysed the vocal responses and non-vocal agonistic responses of wild, free-ranging male songbirds briefly exposed to acoustic playbacks of natural, conspecific songs that we experimentally altered in trill performance. This enabled us to address the questions of whether and how the trill performance of a rival (e.g., territory intruder) affects, in the short term, not only the non-vocal agonistic behaviour but also the trill performance and other aspects of song output of the territory owner.

We used wild, free-ranging Lincoln's sparrows (*M. lincolnii*) as our study system. Male Lincoln's sparrows produce warbling, wren-like songs, 1–4 s in duration, with 3–10 syllable-types per song, and usually beginning at a lower sound frequency before a finishing flourish usually at higher sound frequency (Pandolfino et al., 2023). Unlike some bird species that also sing trills (Podos, 1997), Lincoln's sparrows almost always produce multiple, complex trills per song, with each trill of a song comprised of a distinct syllable type (Figure 1). Most songs contain 3–5 trills (Cicero & Benowitz-Fredericks, 2000), but we have recorded songs from free-ranging males with seven, although this is rare.

Individuals sing from one to six song types, with song type based on the composition and order of unique syllable types (Cicero & Benowitz-Fredericks, 2000) (Figure 1). However, even within a single song type, consecutive songs vary in the syllable count of each trill and sometimes by the inclusion or exclusion of single syllable types. Males perform their songs usually from conspicuous perches at a rate between 1 and 10 songs per minute, singing a single song type several times over a period of one to several minutes before singing another song type for one to several minutes (Pandolfino et al., 2023). This may last from a few minutes to more than an hour. At Molas Pass, males sing mostly around dawn, with song-rates dropping dramatically in the first hour after dawn (Beaulieu & Sockman, 2012). They nonetheless continue at lower levels through mid-morning, but afternoon and evening song is infrequent.



Figure 1. Example spectrograms of three digitally manipulated Lincoln's sparrow (*Melospiza lincolnii*) song types. The manipulation involved cutting 15 ms silence from between each syllable of each trill for the high-performance treatment (lower subpanel of each song) and pasting it in the corresponding space of the song's digital copy for the low-performance treatment (upper subpanel of each song). Individual syllables are numbered in the top subpanel, and they are coloured according to syllable type. Syllable 13 is not trilled by our definition (see text).

Trill performance shows substantial between-male variation in our study population (Sockman, 2009), and wild-caught, laboratory-housed females reproductively primed by changes in photoperiod show behavioural preferences toward male songs that are digitally elevated in trill performance over those that are digitally reduced in trill performance (Caro et al., 2010; Lyons et al., 2014). Thus, variation in trill performance has saliency for female receivers. Wild-caught, laboratory-housed male Lincoln's sparrows produce more songs in response to several days of natural song stimuli that tend to be of higher trill performance compared to song stimuli that tend to be of lower trill performance (Sewall et al., 2010), but, in that study, song stimuli differed in other ways, as well. So, whether or not it is trill performance or some combination of other metrics that drives these differences in singing effort is not clear. Here, we asked whether trill performance itself has saliency for male receivers, specifically in terms of rapidly driving change in their vocal and non-vocal agonistic behaviour. We hypothesized that if it is the trill performance per se that matters in males' responses to a territory intruder, the receiver should modulate his non-vocal agonistic behaviours (e.g., time spent close to the speaker) and his song production-rate and quality according to the trill performance of the song stimulus to which they are exposed. In particular, because females prefer songs of high trill performance (Caro et al., 2010; Lyons et al., 2014), we predicted that males will perceive songs of high trill performance as more threatening, and therefore that they will respond more strongly. This assumes that males do not always perform trills and other aspects of song at their maximal levels. Alternatively, female preferences for songs with high trill performance might be driven by the higher overall costs of reproduction for females than males (Williams, 2012). Males might therefore care less than females about other males' trill performance and show little variation in their responses to variation in an intruder's trill performance.

Additionally, as mentioned above, an individual song will almost always contain multiple trills, but it is not clear for this or other species whether and how the individual trills within a song may vary in performance or function. Thus, another purpose of this study was to examine how effects of the playbacks vary between the multiple trills within songs of receivers. We hypothesize that if male Lincoln's sparrows strongly care about an intruder's trill performance, they should elevate the performance of all their trills following the playback of songs in which trills have been digitally increased in performance. Alternatively, all trills may not have the same signalling value, or may not be as costly or difficult to produce, in which case we predict that males will only modulate the performance of some of their trills and not of others, again suggesting that trill performance is not always maximized.

2. Material and methods

2.1. General

We conducted this experiment at Molas Pass, Colorado (37.747°N, 107.697°W; elevation 3250 m), a sub-alpine, open-field, swampy meadow habitat typical of Lincoln's sparrows, which breed at high elevations or high latitudes in the western USA and in Canada. We have described details of the site and population previously (Sockman, 2008, 2009, 2016). Males sing from atop small willow (*Salix wolfii, S. glauca*) shrubs or from low points in spruce (*Picea engelmannii*) trees, principally around dawn and in the first few hours that immediately follow (Beaulieu & Sockman, 2012). Courtship song, mate choice, nest building, and laying occur across the month of June, with hatching and nestling care occurring from late June through July, depending on the breeding pair (Graham et al., 2011).

From 2005–2016 (except 2014), we recorded the songs of male Lincoln's sparrows free-ranging on the study site described above, some of which we used for the construction of the stimuli described below. Using digital audio recorders (PMD 660 and 670, Marantz, Mahwah, NJ, USA) and short-shotgun microphones (ME-66/K6, Sennheiser, Wedemark, Germany), we recorded songs during morning hours (dawn–1000 h) as uncompressed files sampled at 44.1 kHz. We used the software Raven Pro (Cornell Laboratory of Ornithology, Ithaca, NY, USA) for the construction of song stimuli and for generating spectrograms for the measurement of stimulus songs, songs produced by the experimental subjects in response to the playbacks, and songs produced by non-experimental subjects (details below). Spectrograms had a window size of 512 samples, a time grid of 5.80 ms and a frequency grid of 86.1 Hz.

2.2. Construction of song stimuli

The most straightforward way to experimentally alter the trill performance of a Lincoln's sparrow song is to alter its syllable rate by adding or removing periods of silence to or from the gaps between the syllables of a trill, thereby making syllable rate slower or faster, respectively. This is possibly the primary, though not the only means by which some other studies have manipulated trill performance (e.g., Draganoiu et al., 2002; Caro et al., 2010). Altering the other component of trill performance, frequency bandwidth, is not feasible for Lincoln's sparrow songs without spectral alterations that might render them unrecognizable to subjects.

We began by normalizing the peak amplitudes of 4-5 complete songs from each of 10 individual males (hereafter stimulus males), using only recordings with high signal-to-noise ratios and only songs containing exactly four trills, defined for the purposes of our trill manipulations as any song phrase in which the syllable type of that phrase had at least two occurrences separated by a silent period. Lincoln's sparrows also produce non-trilled phrases, as well as occasional trills with no silent portion between syllables, but these types of phrases, when they occurred in our stimulus songs, were not counted among the four manipulated trills and were not altered in any way. The song-types (i.e., the unique combination of syllable types within a song) differed between the 10 stimulus males, but, among the 4-5 songs of a single stimulus male, they were the same. Visualizing the songs as spectrograms in Raven Pro, we cut 15 ms of silence from the space between each syllable of each trill of each song and pasted it into the corresponding intersyllable space of a digital copy of the song, thereby generating two copies of each song that were identical except for the duration of silence between the syllables of their trills and thus the duration of the entire song (see Lyons et al., 2014 for a more detailed description of this technique) (Figure 1). This was the specific manipulation we have done in previous experiments on Lincoln's sparrows and was chosen to provide significant variation in syllable rate but without exceeding natural variation in this parameter (see below). The result was one set of 4–5 songs from each of 10 males with relatively long gaps of silence between each trill's syllables (low-performance treatment) and another set of the same songs from the same males with relatively short gaps of silence between each trill's syllables (high-performance treatment).

We then appended a silent period to the end of each stimulus song (see below), concatenated all of the low-performance stimuli from one stimulus male in random order into a single sound file, and repeated the concatenated set until the file totalled 2.5 min in duration. We then appended 1 min of silence followed by one repetition of the above 2.5 min of song. We made

a second sound file of the stimulus male's corresponding high-performance stimuli in the same way and in the same song order. We then repeated this process for each of the remaining nine stimulus males, thereby producing 10 pairs of 6-min stimulus files (20 files in total) of identical songs in which one member of each pair consisted of the low-performance stimuli and the other of the high-performance stimuli from a single of the 10 stimulus males.

Because the experimental manipulation of syllable rate necessarily made low performance songs slightly longer than high-performance songs, the duration of silence between each song varied but was always approximately 10 s, thereby producing a file with a rate of song production normal for the population. With an average song duration of approximately 2.5 s, a trill count of 4 per song, an average syllable count of 4 per trill, and 15 ms added or removed between each syllable, songs rendered as low performance differed from songs rendered as high performance by approximately 14% in duration, and thus the silent period that followed songs differed by approximately 7%. The members of a stimulus pair were otherwise identical, in that they contained the same songs recorded from the same stimulus male. Compared to other stimulus pairs, they contained the same number of trills and were repeated the same number of times, at the same rate, and at the same peak amplitude.

We measured trill performance in spectrograms Raven Pro produced from the stimulus songs and from the songs of experimental subjects and nonexperimental subjects (recorded from 2005-2010). We used measurements from the songs of non-experimental subjects not only to calculate the upper bound regression for quantifying trill performance but also for comparing with measures of songs from experimental subjects. A trill's performance measure is based on the relationship between its syllable-production rate and its frequency bandwidth, both of which can involve ambiguity in measurement. Syllable-production rate is the syllable count per second. Whole trills are separated from other phrases of a song by periods of silence. Therefore, for the purpose of objectively and reliably measuring syllable-production rate, we delineated a timespan from a distinct acoustic landmark on the first syllable of the trill to the corresponding landmark on the last syllable of the trill, thereby capturing the timespan of all sound and following silent periods of all but one syllable in the trill. We then calculated syllable rate as 1 less than the number of syllables in the trill divided by this timespan in

seconds. We calculated trill duration as the timespan between identical landmarks of the first and last syllable (as mentioned above) plus the ratio of that same timespan and 1 less than the syllable count. This approach of temporally delineating all but one syllable and its following silent period does not necessitate knowing exactly where in the song the trill, which is bracketed by periods of silence, begins and ends and is not affected by sound pressure level or gain, assuming the signal to noise ratio is sufficient for clear visualization of the trill. And, indeed, our measure of trill duration includes an estimated duration of silence following the final syllable. Apparent frequency bandwidth is also affected by gain and can therefore be difficult to quantify objectively. Thus, we simultaneously delineated each trill's high and low frequencies using a software routine that calculates the upper and lower boundaries for the middle 90% of sound energy. Using this subset of sound energy for the calculation of frequency bandwidth enables measures that are not sensitive to sound pressure level and gain settings. It also favours the highest energy frequencies, excluding frequencies that are very low in sound energy and which would otherwise contribute disproportionately to bandwidth. We then calculated trill performance as the perpendicular deviation from an upper bound regression of maximal frequency bandwidth on the centre of syllable-rate bins spanning ranges of three syllables per second (Sockman, 2009), with larger, less negative values representing higher performance. For some species, the use of alternative methods to the upperbound-regression has been recommended (Wilson et al., 2014), but we retain use of it for consistency with other studies on Lincoln's sparrows and because the Lincoln's sparrow was one species for which such an approach supported the presence of a performance trade-off (Wilson et al., 2014).

Syllable rate and therefore trill performance differed significantly between paired treatment levels ($t_{196} = 13.20$, p < 0.001), but they were both within the range that occurs naturally for each trill number of a song in this population (Table 1). As described above, frequency bandwidth was unaltered from the original song recordings and was therefore identical between treatment levels and to natural levels.

2.3. Field procedures

We conducted all stimulus playbacks in 2013 from June 04–18 and initiated them between 0559 and 0750 h local time. We began each procedure by locating a territorial male (hereafter the experimental subject or subject)

Table 1.

Range of syllable rate (syllables/s) and trill performance with respect to the within-song trill number for Lincoln's sparrow (*Melospiza lincolnii*) natural trills and trills that were digitally manipulated for simulated territorial intrusions.

Measurement	Trill 1	Trill 2	Trill 3	Trill 4
Syllable rate				
Natural	3.01 to 19.92	3.14 to 20.65	3.23 to 17.87	3.10 to 21.56
Manipulated	3.39 to 9.58	4.94 to 12.03	4.62 to 14.01	6.49 to 18.96
Trill performance				
Natural Manipulated	-2.48 to -0.18 -2.17 to -0.84	-2.08 to -0.05 -1.43 to -0.50	-2.10 to 0.24 -1.40 to -0.03	-2.21 to 0.31 -1.70 to 0.26

that had been banded previously with a unique combination of colour bands (Sockman, 2009), thereby making him identifiable at the individual level. We then placed a speaker (Legendary 7-100, Pignose, Las Vegas, NV, USA) on the ground facing up centrally in his territory, and connected it to a digital audio player (iPod, Apple, Cupertino, CA, USA), loaded with each of the 20 6-min sound files. After a 10-min period of acclimation, we initiated the 6-min broadcast of a sound file to the subject. Each subject was assigned the stimulus pair from one and only one stimulus male, and we worked our way through the 10 pairs of sound files in random order (without replacement, until using all of them, at which point we re-randomized their order and repeated their use) with each new subject we encountered. We did not measure the sound pressure level of broadcasts, but we maintained identical settings on our audio equipment between playbacks. The initial choice of settings was based on a subjective estimate of the loudness of the normal, natural songs to which we had been exposed thousands of times over previous years on this study site. During the broadcast and for 10 min following it, we collected digital audio recordings of the subject's song, as described above for stimulus males and non-experimental subjects. During the broadcast we also collected measures associated with non-vocal agonistic behaviour, including counting the number of times the subject flew directly over the speaker and moved at least a whole body width (i.e., hops and turns) not including flights over the speaker, and timing the subject's duration spent within 1, 5 and 10 m of the speaker and more than 10 m from the speaker. We also measured the latency of subjects to approach within 1 m of the speaker, but, because some subjects never approached within 1 m, we did not use this measure in analyses.

This was a within-subjects treatment design, with the goal of exposing each subject to both the low- and high-performance levels of its randomly assigned stimulus pair. We began by exposing the first identified subject to the low-performance level and alternated initial treatment levels (and therefore treatment order) between subjects as we located them, such that we exposed the second identified subject first to the high-performance level, etc. Each individual subject was located approximately the same time the day following its initial stimulus exposure and then exposed to the other level in the same manner (6 min exposure and 10 min of observation following exposure). The choice to alternate instead of randomly assign each new subject's treatment level was because treatment order may not be well-distributed due to a small sample size (Hurlbert, 1984) associated with the difficulty of field playbacks with this famously furtive species (Pandolfino et al., 2023).

2.4. Analyses

Due to the potential effects of treatment order on the responses of individual subjects (e.g., Caro et al., 2010), it was important to control for this factor and its interaction with treatment level. Controlling for these two factors by including them as parameters (together with treatment level) would have over-parameterized statistical models, due to the relatively small sample size. So instead, we controlled for their effects by ensuring that all analyses were balanced with respect to treatment order and included only subjects that responded to both treatment levels, ensuring that the interaction between treatment order and treatment level was also balanced.

A total of 19 subjects were present (though not necessarily singing; see below) for both treatment levels. For analyses not involving within-song measurements (i.e., all non-song agonistic behaviour and song counts), we had to exclude one subject (number 19, the final one) to enable the balance described above, yielding a sample size of 18 subjects (36 treatments) for each of these analyses. Only seven subjects sang during both playback treatments, with only six that would have enabled analyses that were balanced. Of these, the songs of two were not analysable due to a low signal-to-noise ratio (see below). Thus, with such a small sample size, we did not perform analyses involving within-song measurements (the duration, trill count, and syllable count of songs and the duration, performance, frequency bandwidth, syllable rate and syllable period (inverse of syllable rate) of trills). For analyses involving measurements of song following playback, we excluded the

178

four subjects that did not sing following either playback treatment (as well as five subjects that sang following only one of the two treatment levels; analyses could not be balanced with the inclusion of any combination of them). This yielded a sample size of 10 subjects (20 treatments) for these analyses. Thus, we conducted all analyses of song measures (as opposed to song count and non-song agonistic behaviour) on the post-playback song only of 10 subjects. With this approach, all analyses were balanced for treatment order and its interaction with treatment level.

We counted the songs of subjects from our digital audio recordings collected during and following playbacks. We could distinguish subject songs from stimulus songs in the audio recordings due to their differences in amplitude and to the unique spectral-temporal properties of their syllable-types. For each subject, we measured up to the first 10 songs. We could not measure some songs due to sound interference (e.g., wind, passing vehicle) or poor signal-to-noise ratio. For each trill, we determined syllable rate (syllables/s), trill frequency bandwidth (kHz), trill performance, trill syllable count, and trill duration (s). We also determined the inverse of syllable rate, syllable period (s/syllable) to facilitate the visual assessment of trill duration, which is simply the product of syllable period and trill syllable count. For each song, we determined its trill count, syllable count and duration (s).

We performed our statistical analyses using the software Stata/IC 15.1 for the Macintosh. In order to reduce dimensionality of our analyses, we performed a principal component analysis (PCA) on the correlation matrix of each of three sets of the above-mentioned behavioural responses. The first of the three sets combined all six non-song agonistic responses (speaker flyover count; count of whole-body (non-flight) movements; and duration spent within 1, 5 and 10 m of the speaker and more than 10 m from the speaker) with song count during playback and song count following playback. Loadings of the first principal component were all similar in magnitude (absolute value 0.29–0.46), except fly-over count (-0.11) and song count during playback (0.08). Thus, we re-conducted the first PCA after removing fly-over count and song count during playback, yielding a first component which explained 66% of the total variance and which produced loadings all of similar magnitude (absolute value 0.29–0.47). The loadings of all except the duration more than 10 m from the speaker were positive. Thus we interpreted the first-axis factor scores from this as agonistic responses.

The remaining two PCA analyses examined various aspects of song behaviour: measures of song length, each of which varied within each treatment; and measures of trills, each of which varied within each song. Thus, the second PCA combined song duration, song syllable count, and song phrase count, which produced all positive loadings of nearly identical magnitude (0.56–0.59) and explained 85% of the total variance. We interpreted its first-axis factor scores as measures of song-length response. Finally, the third PCA combined trill syllable count, syllable rate, frequency bandwidth, trill performance, and the proxy for trill duration described above, which produced all positive loadings of similar magnitude (0.30–0.57) and explained 46% of the total variance. We interpreted its first-axis factor scores as trill responses.

We statistically analysed the effects of the experimental treatment using three general linear mixed models, each corresponding to one of the three PCA-produced responses. Each of the models included the predictor treatment. Analyses of the third (trill response) also included the song's trill number (modelled categorically, with trill number 1 as the reference group) and its interaction with treatment as predictors. Models differed based on the structure of their random effects, which we prescribed following Schielzeth & Forstmeier (2009). For the agonistic response, we nested observation (N = 36) within subject (N = 18) as a random intercept and random coefficient for treatment. For the song-length response, we nested observation (N = 188) within treatment (N = 20) as a random intercept and treatment within subject (N = 10) as a random intercept and random coefficient for treatment. For the trill response, we nested observation (N = 652) within song as a random intercept and random coefficient for trill number, song (N = 183) within treatment as a random intercept, and treatment (N = 20)within subject (N = 10) as a random intercept and random coefficient for treatment.

The analysis of trill response produced significant results (see Results). So, for a post-hoc examination of trill response, we performed six additional mixed-effects models on syllable rate, bandwidth, trill performance, syllable period, trill syllable count, and trill duration, with treatment, trill number, and their interaction as predictors and observation (N = 682 for trill syllable count, N = 652 for all other measures (some syllables that, due to noise, were unanalysable, could still be counted)) within song as a random intercept and random coefficient for trill number, song (N = 188 for trill syllable

180

count, N = 183 for all other measures) within treatment as a random intercept, and treatment (N = 20) within subject (N = 10) as a random intercept and random coefficient for treatment. In figures we depicted marginal means with standard errors estimated for the treatment effect from the above models and thus controlling for random effects. We also conducted post-hoc contrasts of marginal means to compare treatment levels for each trill number.

For each experimental subject, we also used spectrograms in Raven to visually compare the syllable types of trills (trill types) between the two treatments, as well as between those produced by the experimental subjects and those of the stimulus songs, noting whether trill types were the same or different in those comparisons. The conclusions from these comparisons were unambiguous for all compared trills in this study and for nearly all we have ever inspected. As an example, in three different songs (one recorded from each of three different males) (Figure 1), spectral traces readily reveal a total of 16 syllable types (and 12 trill types), particularly under magnification.

From 2005–2010 recordings of non-experimental subjects, we measured the trills of up to the first 10 songs recorded for each positively identified individual per day. Again, individuals had been banded previously with a unique combination of colour bands (Sockman, 2009), thereby making them identifiable at the individual level. Although we did not conduct inferential statistical analyses on the songs of these non-experimental males, we did account for the random effects of individual males ($N \ge 106$) and songs ($N \ge 2398$) in estimating marginal means for their trill ($N \ge 8226$) measures for visual comparison with the trill measures of experimental subjects.

3. Results

We found no treatment effect on either of the first two PCA-generated responses. Treatment did not appear to strongly affect the song and non-song agonistic response (z = 1.34, p = 0.18) or the song-length response (z = -1.11, p > 0.2). However, treatment did affect the trill response in its interaction with trill number (Figure 2). Specifically, the change in trill response from trill 1 to 2 (z = 1.97, p = 0.049), 1 to 3 (z = -2.63, p = 0.009), and possibly 1 to 4 (z = -1.91, p = 0.056) each depended on treatment.

Post-hoc analysis showed no support that treatment as a main effect influenced any of our six trill measures in songs of experimental subjects during the period immediately following the playback treatment (Table 2). However, we found a strong relationship between trill number and each of our



Figure 2. Effects of trill-performance playback treatment on measures of agonistic behaviour and song length and of the playback treatment, within-song trill number, and their interaction on a measure of trill score in Lincoln's sparrows (*Melospiza lincolnii*). Agonistic behaviour, song length and trill score were each derived from the first-axis factor scores of principal component analyses of related variables. Points and error bars are marginal means \pm SE from statistical models described in the text. See text for details.

trill measures, as well as a strong effect of the interaction between treatment and the trill number on trill performance (Table 2). Post-hoc contrasts of marginal means revealed this effect exclusively for trill three (Figure 3). Moreover, the effect of the interaction on trill performance was driven by its effect on frequency bandwidth, as effects on syllable rate were minimal at most (Table 2, Figure 3). Additionally, the interaction between treatment and trill number also affected trill syllable count, with post-hoc contrasts reveal-

Table 2.

Statistical effects of trill-performance playback treatment (0, low; 1, high), trill number (trill 1 as reference), and their interaction (\times) on six trill measures in Lincoln's sparrows (*Melospiza lincolnii*).

Response Predictor	Coefficient	SE	Z	р
Trill performance				
Intercept	-1.399	0.070	-19.890	< 0.001
Treatment	-0.044	0.086	-0.52	>0.2
Trill 2	0.142	0.057	2.480	0.013
Trill 3	0.471	0.059	8.030	< 0.001
Trill 4	0.340	0.062	5.500	< 0.001
Treatment \times trill 2	0.075	0.083	0.900	>0.2
Treatment \times trill 3	-0.301	0.084	-3.570	< 0.001
Treatment \times trill 4	-0.001	0.093	-0.010	>0.2
Syllable rate (syllables/s)				
Intercept	4.572	0.291	15.700	< 0.001
Treatment	0.390	0.416	0.94	>0.2
Trill 2	2.021	0.219	9.240	< 0.001
Trill 3	2.754	0.226	12.190	< 0.001
Trill 4	3.825	0.240	15.930	< 0.001
Treatment \times trill 2	-0.169	0.318	-0.530	>0.2
Treatment \times trill 3	0.152	0.324	0.470	>0.2
Treatment × trill 4	-0.365	0.366	-1.000	>0.2
Frequency bandwidth (kHz)				
Intercept	1.271	0.067	18.830	< 0.001
Treatment	-0.080	0.097	-0.83	>0.2
Trill 2	-0.007	0.053	-0.140	>0.2
Trill 3	0.275	0.055	5.010	< 0.001
Trill 4	0.092	0.058	1.580	0.11
Treatment \times trill 2	0.095	0.077	1.230	>0.2
Treatment \times trill 3	-0.308	0.079	-3.920	< 0.001
Treatment \times trill 4	0.007	0.089	0.080	>0.2
Syllable period (s/syllable)				
Intercept	0.229	0.007	33.120	< 0.001
Treatment	-0.018	0.010	-1.79	0.073
Trill 2	-0.072	0.005	-13.730	< 0.001
Trill 3	-0.085	0.005	-15.590	< 0.001
Trill 4	-0.102	0.006	-17.790	< 0.001
Treatment × trill 2	0.016	0.008	2.120	0.034
Treatment \times trill 3	0.011	0.008	1.430	0.15
Treatment \times trill 4	0.021	0.009	2.390	0.017

Table 2.

(Continued.)

Response Predictor	Coefficient	SE	Z	р
Trill syllable count				
Intercept	2.453	0.178	13.750	< 0.001
Treatment	0.183	0.228	0.80	>0.2
Trill 2	0.745	0.190	3.920	< 0.001
Trill 3	1.003	0.194	5.160	< 0.001
Trill 4	1.023	0.200	5.110	< 0.001
Treatment × trill 2	0.422	0.275	1.540	0.12
Treatment × trill 3	-0.062	0.279	-0.220	>0.2
Treatment \times trill 4	-0.821	0.298	-2.760	0.006
Trill duration (s)				
Intercept	0.543	0.017	31.500	< 0.001
Treatment	-0.008	0.020	-0.37	>0.2
Trill 2	-0.057	0.018	-3.220	0.001
Trill 3	-0.086	0.018	-4.750	< 0.001
Trill 4	-0.112	0.019	-5.850	< 0.001
Treatment × trill 2	0.087	0.026	3.410	0.001
Treatment × trill 3	0.010	0.026	0.380	>0.2
Treatment \times trill 4	-0.065	0.029	-2.230	0.026

Each model nested observation (N = 682 for trill syllable count, N = 652 for all other measures) within subject's song as a random intercept and random coefficient for trill number, subject's song (N = 188 for trill syllable count, N = 183 for all other measures) within playback treatment as a random intercept, and playback treatment (N = 20) within subject (N = 10) as a random intercept and random coefficient for playback treatment.

ing that effect for trills two and four, resulting in a change in the duration of trills two and four (Figure 3).

For nine of the ten experimental subjects used for trill analysis, all trills produced post-playback differed between treatments in their type. One subject produced all the same trill types following the low-performance treatment as he did following the high-performance treatment. Thus, the effect of treatment on the frequency bandwidth of trill three was due to a change between treatments in syllable type, with the syllable type following the high-performance treatment. Following the low-performance treatment. Following the low-performance treatment. Following the low-performance treatment, five, four, and one subjects matched zero, one (12-25% of trill types produced), and two (33%) trill types of the treatment, respectively, but whether or not there was a match did not depend

184



Figure 3. Relationship between within-song trill number and six measures of trills in wild, free-ranging male Lincoln's sparrows (*Melospiza lincolnii*) singing spontaneously (green symbols and lines, $N \ge 106$ individuals) and immediately following a simulated rival's trill-performance (open symbol, dashed line: low-performance, N = 10 individuals; black symbol and solid, black line: high-performance, N = 10 individuals). Points and error bars are marginal means \pm SE from statistical models described in the text. *p*-values from posthoc contrasts between treatment groups are shown above each contrast. Syllable period is the inverse of syllable rate and is shown to facilitate visualizing the products of a trill that determine its duration, syllable period and trill syllable count.

on trill number (i.e., trill 3 was not necessarily the matched trill). Following the high-performance treatment, seven and three subjects matched zero and one (9%) trill types of the treatment, respectively, and again, whether or not there was a match did not depend on trill number. Given the high proportion of zero matches, we analysed this difference using a logit model, in which the response was the dichotomous any matches (1) or no matches (0). A full, mixed model accounting for observation nested within subject would not converge on a solution, so we removed the nested random effect and conducted a general linear model, which showed no strong effect of treatment on the probability that any syllable produced following playback was matched (z = -0.85, p > 0.2).

4. Discussion

The trill performance of a simulated rival has effects on the vocal output of wild, free-ranging male receivers, specifically on the bandwidth of their trills, which drive a change in their trill performance. If maximal trill performance is biomechanically constrained as we suggest above, then this plasticity in performance indicates males are not always performing at their maximum (Podos, 2017). These findings show that ecologically relevant exposure to natural variation in trill performance affects vocal behaviour relevant to sexual signalling (Caro et al., 2010; Lyons et al., 2014).

Some species are known to modulate trills such that their performance changes (e.g., DuBois et al., 2009), but, to our knowledge, ours is the first report that a facultative change in trill performance, in this case driven by facultative change in syllable type, can be driven by a simulated rival's trill performance. Additionally, although this is not the first study that explicitly addresses the question of how song of a receiver is affected by the trill performance to which it was recently exposed, it may be the first to show this effect using trill performance manipulated entirely within the natural range of variation over not only the song as a whole but also for each of multiple trills within the song. Liu and colleagues (2018) tested whether swamp sparrows are more likely to song-type match when exposed to songs of relatively low trill performance or to songs of relatively high trill performance, but they did not find a difference between treatments. Cramer (2013) similarly reported no effect of trill performance on song responses in house wrens (Troglodytes aedon). Moseley and colleagues (2013) found that the rate of soft song production in swamp sparrows was reduced by the digital reduction of the trill

187

performance of playbacks in comparison to control, non-manipulated songs. However, several trills appear to have been manipulated such that they were outside the range of performance that is known to occur naturally for the species (see Figure 1 in Moseley et al., 2013), so it is not clear whether the findings would apply to only natural variation in trill performance. Our findings here show that variation in trill performance within a natural range to which a male receiver is exposed can significantly influence his own song behaviour.

Additionally, our study reveals that responses of experimental subjects are specific to the trill-number within the song, with the change in bandwidth and consequently trill-performance restricted to trill three of songs and the change in trill syllable-count and consequently trill duration occurring for both trill two and trill four (Figure 3). We are not aware of studies that have examined variation in trill performance as a function of trill number, but the results we present here raise questions regarding the function not only of trills in general but of each trill one is very low compared to other trills in the song (Figure 3). Perhaps trill one serves as a vocal warm-up for or contrast to subsequent trills males may attempt to maximize or serves some other purpose entirely. Indeed the function of any of the trills is not clear, but below we discuss some possibilities.

Individuals did not appear to adjust their non-vocal agonistic behaviour according to playback treatment. Perhaps a comparison between high- and low-quality receivers would have revealed a dichotomous response to highperformance playbacks, as Mosely and colleagues (2013) discovered for swamp sparrows. Unfortunately, we have no measures of the quality of our subjects, except, perhaps, in relation to their own trill performance. However, the effect of treatment on the subjects' trill performance precludes use of the subjects' trill performance as an independent measure of their quality. Regardless, we do not have a strong reason to conclude that the treatment-induced difference in vocal response (Figure 3) is a signal of aggressive intent (Searcy & Beecher, 2009; Hof & Podos, 2013). In other words, because they did not appear to flee or even distance themselves from high-performance playback relative to low-performance playback, we cannot conclude that their own high-performance trills produced in response to the low-performance treatment were intended to drive away the putative intruder. Moreover, song rate was lower during than following playback (data not shown). If songs were intended to repel an intruder, we would have expected the opposite. Thus, perhaps subjects were signalling something else, such as their quality to female eavesdroppers relative to the intruder (Logue & Forstmeier, 2008). In fact, female Lincoln's sparrows are known to form long-term memories of the songs of individual males (Beaulieu & Sockman, 2012) and of songs based explicitly on trill performance (Lyons et al., 2014), after which they express song preferences in a mate-choice context based on those memories. This might explain the high-performance response to the low-performance treatment. The high-performance treatment may have exceeded the average subject's own maximal performance, hypothetically resulting in an alternative, perhaps more feasible counter-response in the form of some behaviour we did not assess. Experimental subjects did respond to high-performance trills by elevating the syllable count of their second trills relative to their response to low-performance trills (Figure 3). Moreover, fourth trills were much lower in syllable count following exposure to high-performance trills than those of non-experimental subjects or those following low-performance trills (Figure 3). In short, we do not know the intent of the subjects nor the function of their responses, and thus further studies will be necessary to address these questions.

The treatment-induced change in trill performance was driven by the use of a syllable-type based on its frequency bandwidth rather than by a change in syllable rate, the component of trill performance that we manipulated in our playback stimuli and which, at first pass, would seem more amenable to flexible modulation by the subjects. Lincoln's sparrows clearly vary their syllable rate between the trills of individual songs (Figure 3), however we found no evidence that they varied it in response to treatment. Also, because all but one produced different syllable types between treatments, experimental subjects did not vary frequency bandwidth within a trill type but instead changed the trill type itself, producing types of greater bandwidth following exposure to low-performance trills and types of lower bandwidth following exposure to high-performance trills. Our study population produces more than 100 syllable types (including those in trills and non-trills) (Reinhardt & Sockman unpublished data), in line with the more than 200 types shown for several populations of Lincoln's sparrows across the state of California (Cicero & Benowitz-Fredericks, 2000). Given the high number of syllable types, a detailed understanding of the degree to which any single syllable type might vary in either syllable rate or frequency bandwidth is beyond the

scope of this article. Nevertheless, despite some focused assessment, we have never found a syllable type to vary in bandwidth, although they can vary in centre frequency (Reinhardt and Sockman unpublished data). Therefore, the treatment responses that we observed make sense as a means of facultatively changing the bandwidth of a trill, in that, for the spectrally complex trills of Lincoln's sparrows, varying bandwidth within a trill type could render the trill unrecognizable. Individual Lincoln's sparrows have a repertoire of 20 syllable types, on average (Cicero & Benowitz-Fredericks, 2000), greatly exceeding the number of syllable types produced in a song (Figure 1) or even in a single bout of singing. Thus the lack of similarity in trill types between treatments could be due, in part, to temporal variation and the probability that in comparing any two days there is high likelihood of change. But that does not explain the strong tendency to vary trill type directionally with respect to its frequency bandwidth in response to the treatment, and thus, it appears that some syllable types are chosen over others, depending on the quality of the song competition.

We do not know the ultimate reason for the change in trill type that led to a change in frequency bandwidth in response to song exposure that varies with trill performance, although one possibility is that there is no ultimate reason for the change other than as a mechanism to change frequency bandwidth and thus trill performance. Theoretical modelling supports the hypothesis that a singer should match the song type of a rival when the singer can produce the song type at a higher level of performance (Logue & Forstmeier, 2008; but see Liu et al., 2018 regarding empirical work). Although there is no evidence of song-type matching in Lincoln's sparrows (Cicero & Benowitz-Fredericks, 2000), syllable-type matching might occur, raising the question of whether variation between treatments in the use of specific syllable types was due to matching following exposure to the low-performance treatment and matching avoidance following the high-performance treatment. However, in our study, very few of the subject's syllable types matched those in playbacks, regardless of treatment, and we found no statistical support for elevated syllable-type matching following the low-performance treatment over that following the high-performance treatment. So, although that does not refute the possibility that Lincoln's sparrows engage in some degree of syllable-type matching, syllable-type matching is unlikely to have driven our results. Further examination using playback stimuli constructed from syllable types known to be within the subjects' repertoires would be worthwhile.

Nonetheless, our results suggest Lincoln's sparrows facultatively choose between syllable types that vary in frequency bandwidth, but additional studies are needed to further elucidate the vocal mechanisms for plasticity in trill performance.

In his critique of vocal performance studies in songbirds, Kroodsma (2017) proposed that a singer's combination of syllable rate and frequency bandwidth was primarily a result of the particular song type he had learned and not a product of a biomechanical constraint on vocal performance. In a response, Podos (2017) noted that the specific individuals beside which a young bird settles and therefore the specific songs he learns may, in part, be a function of the two birds' vocal proficiency relative to one another, suggesting that the learning of trills does not itself preclude the possibility that current measures of vocal deviation are indicators of male quality. As mentioned above, Lincoln's sparrows are not known to engage in song-type matching. However, if they engage in syllable-type matching, then perhaps the subset of the syllables to which they are exposed during a critical learning period and that they eventually produce is restricted by their vocal ability.

As stated above, we observed no effect of our treatment on non-vocal agonistic responses, and yet, several other studies have reported such effects, as mentioned in the introduction (Illes et al., 2006; Cramer, 2013; Goodwin & Podos, 2014; Phillips & Derryberry, 2017a). The reason for this disparity is not obvious. Trill performance in Lincoln's sparrows may primarily target females. Still, our experiment used only acoustic stimuli, and was designed to address variation in rival trill performance only. Perhaps the addition of a visual stimulus, such as a live decoy or a taxidermic mount, would have produced different results. The mere presence of playbacks itself did seem to stimulate agonistic responses, but this is speculation, given that we lacked a non-playback comparison group.

Finally, it is worth recalling that our digital manipulation of trill performance was, at its essence, a manipulation of syllable rate, which, in turn, changed the trill performance and also the trill duration and thus song duration of playback stimuli. In light of this present study alone, we cannot know whether subjects were affected by one of these parameters but not another, by a subset of them, or by all of them, and thus it remains possible that the driving factor was something other than trill performance, such as song duration. However, previous studies on Lincoln's sparrows showing elevated song rates in response to songs of not only higher trill performance but also

190

of greater duration (Sewall et al., 2010) suggest that it was the trill performance or syllable rate in the present study to which they preferentially attended and not song duration. This interpretation is consistent with the hypothesis that trill performance (or at least syllable rate) can serve as a biologically relevant signal and that this relevance is likely despite, not because of any difference in the duration of the signal.

Over approximately a quarter century since the first analyses of a tradeoff between syllable rate and frequency bandwidth (Podos, 1997), many studies have addressed the hypothesis that the trills of bird songs are biologically meaningful measures of performance and possibly honest signals of male quality. As mentioned in the Introduction, these studies have spanned a range, from observational field studies of correlations between fitness (e.g., Nishida & Takagi, 2018) or condition (e.g., Ballentine, 2009) and trill performance to manipulative experiments in both the field and laboratory aimed at identifying causal relationships between trill performance and receiver responses (e.g., Illes et al., 2006; Lyons et al., 2014; Phillips & Derryberry, 2017a). Using stimuli that were experimentally manipulated within the range of natural variation and using an ecologically relevant field setting, we have shown in this study that trill performance of a simulated rival drives significant variation in receiver trill performance, not only fulfilling an important criterion in defining trill performance as a biologically meaningful signal (Laidre & Johnstone, 2013) but also expanding our understanding of this signal in its role in driving variation in sexual counter-signalling. We anticipate that future studies continue to elucidate the roles of trill performance and the mechanisms through which it affects receiver responses and ultimately lead to a better understanding of trill performance as an important communication signal in animals.

Acknowledgements

We thank numerous UNC undergraduate students and others for their help collecting data. Support came from the National Institute of Neurological Disorders and Stroke NS055125 to K.W.S.; by the Léon Speeckaert Fund fellowship from the King Baudouin Foundation and the Belgian American Educational Foundation to S.P.C.; and by the American Ornithologist Union, Cooper Ornithological Society, and Animal Behavior Society to S.M.L.

References

Andersson, M. (1994). Sexual selection. — Princeton University Press, Princeton, NJ.

- Ballentine, B. (2009). The ability to perform physically challenging songs predicts age and size in male swamp sparrows, *Melospiza georgiana*. Anim. Behav. 77: 973-978.
- Ballentine, B., Hyman, J. & Nowicki, S. (2004). Vocal performance influences female response to male bird song: an experimental test. — Behav. Ecol. 15: 163-168.
- Beaulieu, M. & Sockman, K.W. (2012). Song in the cold is 'hot': memory of and preference for sexual signals perceived under thermal challenge. — Biol. Lett. 8: 751-753.
- Cardoso, G.C. (2017). Advancing the inference of performance in birdsong. Anim. Behav. 125: e29-e32.
- Cardoso, G.C., Atwell, J.W., Ketterson, E.D. & Price, T.D. (2007). Inferring performance in the songs of dark-eyed juncos (*Junco hyemalis*). — Behav. Ecol. 18: 1051-1057.
- Caro, S.P., Sewall, K.B., Salvante, K.G. & Sockman, K.W. (2010). Female Lincoln's sparrows modulate their behavior in response to variation in male song quality. — Behav. Ecol. 21: 562-569.
- Catchpole, C.K. & Slater, P.J.B. (1995). Bird song. Cambridge University Press, Cambridge.
- Cicero, C. & Benowitz-Fredericks, Z.M. (2000). Song types and variation in insular populations of Lincoln's sparrow (*Melospiza lincolnii*), and comparisons with other *Melospiza*. — Auk 117: 52-64.
- Collins, S. (2004). Vocal fighting and flirting: the functions of birdsong. In: Nature's music: the science of bird song (Marler, P. & Slabbekoorn, H., eds). Elsevier Academic Press, San Diego, CA, p. 39-79.
- Cramer, E.R.A. (2013). Vocal deviation and trill consistency do not affect male response to playback in house wrens. — Behav. Ecol. 24: 412-420.
- De Kort, S.R., Eldermire, E.R.B., Cramer, E.R.A. & Vehrencamp, S.L. (2009). The deterrent effect of bird song in territory defense. Behav. Ecol. 20: 200-206.
- Draganoiu, T.I., Nagle, L. & Kreutzer, M. (2002). Directional female preference for an exaggerated male trait in Canary (*Serinus canaria*) song. — Proc. Roy. Soc. Lond. B: Biol. Sci. 269: 2525-2531.
- Dubois, A.L., Nowicki, S. & Searcy, W.A. (2009). Swamp sparrows modulate vocal performance in an aggressive context. — Biol. Lett. 5: 163-165.
- Dubois, A.L., Nowicki, S. & Searcy, W.A. (2011). Discrimination of vocal performance by male swamp sparrows. — Behav. Ecol. Sociobiol. 65: 717-726.
- Gentner, T.Q. & Hulse, S.H. (2000). Female European starling preference and choice for variation in conspecific male song. — Anim. Behav. 59: 443-458.
- Goodwin, S.E. & Podos, J. (2014). Team of rivals: alliance formation in territorial songbirds is predicted by vocal signal structure. — Biol. Lett. 10: 20131083.
- Graham, E.B., Caro, S.P. & Sockman, K.W. (2011). Change in offspring sex ratio over a very short season in Lincoln's sparrows: the potential role of bill development. — J. Field Ornithol. 82: 44-51.
- Hof, D. & Podos, J. (2013). Escalation of aggressive vocal signals: a sequential playback study. — Proc. Roy. Soc. Lond. B: Biol. Sci. 280: 20131553.

- Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. Ecol. Monogr. 54: 187-211.
- Illes, A.E., Hall, M.L. & Vehrencamp, S.L. (2006). Vocal performance influences male receiver response in the banded wren. — Proc. Roy. Soc. Lond. B: Biol. Sci. 273: 1907-1912.
- Kroodsma, D. (2017). Birdsong performance studies: a contrary view. Anim. Behav. 125: e1-e16.
- Laidre, M.E. & Johnstone, R.A. (2013). Animal signals. Curr. Biol. 23: R829-R833.
- Leitão, A., Ten Cate, C. & Riebel, K. (2006). Within-song complexity in a songbird is meaningful to both male and female receivers. — Anim. Behav. 71: 1289-1296.
- Linhart, P., Jaška, P., Petrusková, T., Petrusek, A. & Fuchs, R. (2013). Being angry, singing fast? Signalling of aggressive motivation by syllable rate in a songbird with slow song. — Behav. Process. 100: 139-145.
- Liu, I.A., Soha, J.A. & Nowicki, S. (2018). Song type matching and vocal performance in territorial signalling by male swamp sparrows. — Anim. Behav. 139: 117-125.
- Logue, D.M. & Forstmeier, W. (2008). Constrained performance in a communication network: implications for the function of song-type matching and for the evolution of multiple ornaments. — Am. Nat. 172: 34-41.
- Lyons, S.M., Beaulieu, M. & Sockman, K.W. (2014). Contrast influences female attraction to performance-based sexual signals in a songbird. — Biol. Lett. 10: 20140588.
- Marler, P. & Peters, S. (1977). Selective vocal learning in a sparrow. Science 198: 519-521.
- Maynard Smith, J. & Harper, D.G.C. (1995). Animal signals: models and terminology. J. Theor. Biol. 177: 305-311.
- McGregor, P.K. (1991). The singer and the song: on the receiving end of bird song. Biol. Rev. 66: 57-81.
- Moseley, D.L., Lahti, D.C. & Podos, J. (2013). Responses to song playback vary with the vocal performance of both signal senders and receivers. — Proc. Roy. Soc. Lond. B: Biol. Sci. 280: 20131401.
- Mountjoy, D.J. & Lemon, R.E. (1991). Song as an attractant for male and female European starlings, and the influence of song complexity on their response. — Behav. Ecol. Sociobiol. 28: 97-100.
- Nelson, D.A. & Poesel, A. (2007). Segregation of information in a complex acoustic signal: individual and dialect identity in white-crowned sparrow song. — Anim. Behav. 74: 1073-1084.
- Nishida, Y. & Takagi, M. (2018). Song performance is a condition-dependent dynamic trait honestly indicating the quality of paternal care in the bull-headed shrike. — J. Avian Biol. 2018: e01794.
- Pandolfino, E.R., Ammon, E.M. & Sockman, K.W. (2023). Lincoln's sparrow (*Melospiza lincolnii*), version 2.0. In: Birds of the world (Rodewald, P.G., ed.). Cornell Lab of Ornithology, Ithaca, NY.
- Phillips, J.N. & Derryberry, E.P. (2017a). Equivalent effects of bandwidth and trill rate: support for a performance constraint as a competitive signal. — Anim. Behav. 132: 209-215.

- Phillips, J.N. & Derryberry, E.P. (2017b). Vocal performance is a salient signal for male-male competition in white-crowned sparrows. — Auk 134: 564-574.
- 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.
- Podos, J. (2017). Birdsong performance studies: reports of their death have been greatly exaggerated. — Anim. Behav. 125: e17-e24.
- Podos, J. & Nowicki, S. (2004). Performance limits on birdsong production. In: Nature's music: the science of bird song (Marler, P. & Slabbekoorn, H., eds). Academic Press, New York, NY, p. 318-342.
- Ritschard, M., Riebel, K. & Brumm, H. (2010). Female zebra finches prefer high-amplitude song. — Anim. Behav. 79: 877-883.
- Schielzeth, H. & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. — Behav. Ecol. 20: 416-420.
- Searcy, W.A. & Beecher, M.D. (2009). Song as an aggressive signal in songbirds. Anim. Behav. 78: 1281-1292.
- Searcy, W.A. & Nowicki, S. (2000). Male-male competition and female choice in the evolution of vocal signaling. — In: Animal signals: signalling and signal design in animal communication (Espmark, Y., Amundsen, T. & Rosenqvist, G., eds). Tapir Academic Press, Trondheim, Norway, p. 301-315.
- Sewall, K.B., Dankoski, E.C. & Sockman, K.W. (2010). Song environment affects singing effort and vasotocin immunoreactivity in the forebrain of male Lincoln's sparrows. — Horm. Behav. 58: 544-553.
- Sockman, K.W. (2008). Ovulation order mediates a trade-off between pre-hatching and posthatching viability in an altricial bird. — PLoS ONE 3: e1785.
- Sockman, K.W. (2009). Annual variation in vocal performance and its relationship with bill morphology in Lincoln's sparrows, *Melospiza lincolnii*. — Anim. Behav. 77: 663-671.
- Sockman, K.W. (2016). The regulation of behavioral plasticity by performance-based feedback and an experimental test with avian egg production. — Am. Nat. 187: 564-575.
- Sockman, K.W., Salvante, K.G., Racke, D.M., Campbell, C.R. & Whitman, B.A. (2009). Song competition changes the brain and behavior of a male songbird. — J. Exp. Biol. 212: 2411-2418.
- Sockman, K.W., Sewall, K.B., Ball, G.F. & Hahn, T.P. (2005). Economy of mate attraction in the Cassin's finch. — Biol. Lett. 1: 34-37.
- Vehrencamp, S.L. (2001). Is song-type matching a conventional signal of aggressive intentions? — Proc. Roy. Soc. Lond. B: Biol. Sci. 268: 1637-1642.
- Vehrencamp, S.L., De Kort, S.R. & Illes, A.E. (2017). Response to Kroodsma's critique of banded wren song performance research. — Anim. Behav. 125: e25-e28.
- Williams, T.D. (2012). Physiological adaptations for breeding in birds. Princeton University Press, Princeton, NJ.
- Wilson, D.R., Bitton, P.-P., Podos, J. & Mennill, D.J. (2014). Uneven sampling and the analysis of vocal performance constraints. — Am. Nat. 183: 214-228.