



**HAL**  
open science

## Vocal performance of one affects that of another

Keith W. Sockman, Susan M. Lyons, Samuel Caro

► **To cite this version:**

Keith W. Sockman, Susan M. Lyons, Samuel Caro. Vocal performance of one affects that of another. Behaviour, 2024, 161 (2), pp.167-194. 10.1163/1568539X-bja10256 . hal-04756922

**HAL Id: hal-04756922**

**<https://hal.science/hal-04756922v1>**

Submitted on 28 Oct 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Vocal performance of one affects that of another**

2

3 **Keith W. Sockman\***, Susan M. Lyons<sup>1</sup>, & Samuel P. Caro<sup>2</sup>

4 Department of Biology, University of North Carolina at Chapel Hill, U.S.A.

5

6

7

8

9 \*Corresponding author's e-mail address: kws@unc.edu

10

11

12 Current Affiliation:

13 <sup>1</sup>Great Minds PBC, 55 M Street SE, Suite 340, Washington, D.C. 20003

14 <sup>2</sup>CEFE, University of Montpellier, CNRS, EPHE, IRD, Montpellier, France

15

16

17 Short Title: Vocal performance in Lincoln's sparrows

18 **Summary**

19

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

31

32 **Keywords:** bird song, intraspecific competition, Lincoln's sparrow (*Melospiza lincolni*), sexual  
33 signal, simulated territorial intrusion, trill performance

## 34 Introduction

35

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

44 The songs of male birds are widely considered to be signals of mate attraction  
45 (McGregor, 1991; Sockman et al., 2005), aggressive intent (Searcy & Beecher, 2009; Linhart et  
46 al., 2013), and identity (Nelson & Poesel, 2007). Depending on species, individuals may vary in  
47 their singing effort (e.g., song rate, duration, or amplitude) (Ritschard et al., 2010; Nishida &  
48 Takagi, 2018), song complexity (variation within and between songs) (Mountjoy & Lemon,  
49 1991; Leitão et al., 2006), and other aspects of their song, leading to variation in attracting  
50 mates (Gentner & Hulse, 2000), in thwarting rivals (de Kort et al., 2009), and in fitness  
51 (Andersson, 1994; Catchpole & Slater, 1995). In the last 26 years (Podos, 1997), an aspect of  
52 song known as trill performance (also called song performance, vocal performance, or vocal  
53 deviation; see explanation below), has come under focus regarding its potential as an honest  
54 signal of male quality (e.g., Phillips & Derryberry, 2017b; Nishida & Takagi, 2018) that could  
55 drive variation between individuals in reproductive success, possibly even leading to  
56 sympatric speciation (Podos, 2001). However, the behavioral significance of trill performance  
57 is still not thoroughly understood (Kroodsma, 2017; but see Podos, 2017; Vehrencamp et al.,  
58 2017), in part because it is not clear the breadth to which natural variation in trill performance  
59 can affect receivers and how the adaptive significance of the signal might vary between  
60 different types of receivers.

61 Trills are rapid repetitions of a single type of song-syllable, and many songbird species  
62 produce them. In comparisons both within and between species, the maximum rate of syllable

63 production, often referred to as trill rate and here referred to as syllable rate, is negatively  
64 correlated with the trill's frequency bandwidth (Podos, 1997; Wilson et al., 2014). This  
65 relationship is thought to result from biomechanical constraints on the vocal tract, in which  
66 syllable rate is limited by the degree of conformational change required by the skeletal-  
67 muscular components necessary to produce the syllable's frequency range (Podos & Nowicki,  
68 2004; Ballentine, 2009). Trills that maximize syllable rate relative to bandwidth or that maximize  
69 bandwidth relative to syllable rate are said to be high performance (Ballentine et al., 2004;  
70 Phillips & Derryberry, 2017a; but see Cardoso, 2017; Cardoso, Atwell, Ketterson, & Price,  
71 2007).

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

83         Some effects of variation in song signals on receiver vocal output are well established.  
84 Probably the most obvious example of this is when songs are learned by the individual singer,  
85 a process thought to characterize the roughly 5000 species of oscine songbirds, which  
86 produce species-typical songs as a function of their prior song exposure during development  
87 (Marler & Peters, 1977; Catchpole & Slater, 1995). Another example is in the case of song-type  
88 matching, whereby a male biases his own production of song toward the types produced by  
89 his neighbor (Vehrencamp, 2001; Logue & Forstmeier, 2008). In addition, the quality of songs  
90 to which an adult is exposed can affect his singing effort in both the short term and longer term  
91 (Sockman et al., 2009; Sewall et al., 2010). But little is known with regard to how variation in the

92 trill performance to which an individual is exposed affects his own song (but see Cramer, 2013;  
93 Moseley et al., 2013; Liu et al., 2018). In fact, the predicted direction of the receiver's response  
94 is not even obvious based on signaling theory (Searcy & Nowicki, 2000; Collins, 2004). Should  
95 the elevated threat of a high-performance trill suppress or instead excite the (song) behavior of  
96 territorial receivers? Illes and colleagues (2006) found that male banded wrens (*Thryothorus*  
97 *pleurostictus*) more rapidly approach songs with high-performance trills than with low-  
98 performance trills, and yet they spend less time in close proximity to the high-performance  
99 songs when those songs approach the theoretical performance limit. Male swamp sparrows  
100 (*Melospiza georgiana*) respond more aggressively to control songs (with intermediate trill  
101 performance) than songs with digitally reduced trill performance, but they are also more  
102 aggressive, on average, to control songs than songs with digitally enhanced trill performance,  
103 a pattern explained, in part, by the subject's own vocal performance (Moseley et al., 2013). In  
104 the present article, we report on a study in which we analyzed the vocal responses and non-  
105 vocal agonistic responses of wild, free-ranging male songbirds briefly exposed to acoustic  
106 playbacks of natural, conspecific songs that we experimentally altered in trill performance. This  
107 enabled us to address the questions of whether and how the trill performance of a rival (e.g.,  
108 territory intruder) affects, in the short term, not only the non-vocal agonistic behavior but also  
109 the trill performance and other aspects of song output of the territory owner.

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

118 Individuals sing from one to six song types, with song type based on the composition  
119 and order of unique syllable types (Cicero & Benowitz-Fredericks, 2000) (Figure 1). However,  
120 even within a single song type, consecutive songs vary in the syllable count of each trill and

121 sometimes by the inclusion or exclusion of single syllable types. Males perform their songs  
122 usually from conspicuous perches at a rate between 1 and 10 songs per minute, singing a  
123 single song type several times over a period of one to several minutes before singing another  
124 song type for one to several minutes (Pandolfino et al., 2023). This may last from a few minutes  
125 to more than an hour. At Molas Pass, males sing mostly around dawn, with song-rates  
126 dropping dramatically in the first hour after dawn (Beaulieu & Sockman, 2012). They  
127 nonetheless continue at lower levels through mid-morning, but afternoon and evening song is  
128 infrequent.

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

150 males (Williams, 2012). Males might therefore care less than females about other males' trill  
151 performance and show little variation in their responses to variation in an intruder's trill  
152 performance.

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

163

## 164 **Material and methods**

165

### 166 *General*

167

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

177         From 2005-2016 (except 2014), we recorded the songs of male Lincoln's sparrows free-  
178 ranging on the study site described above, some of which we used for the construction of the



179 stimuli described below. Using digital audio recorders (Marantz PMD 660 and 670, Mahwah,  
180 NJ, U.S.A.) and short-shotgun microphones (Sennheiser ME-66/K6, Wedemark, Germany), we  
181 recorded songs during morning hours (dawn-1000 hours) as uncompressed files sampled at  
182 44.1 kHz. We used the software Raven Pro (Cornell Laboratory of Ornithology, Ithaca, NY,  
183 U.S.A.) for the construction of song stimuli and for generating spectrograms for the  
184 measurement of stimulus songs, songs produced by the experimental subjects in response to  
185 the playbacks, and songs produced by non-experimental subjects (details below).  
186 Spectrograms had a window size of 512 samples, a time grid of 5.80 ms, and a frequency grid  
187 of 86.1 Hz.

188

### 189 *Construction of Song Stimuli*

190

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

198 We began by normalizing the peak amplitudes of 4-5 complete songs from each of 10  
199 individual males (hereafter stimulus males), using only recordings with high signal-to-noise  
200 ratios and only songs containing exactly four trills, defined for the purposes of our trill  
201 manipulations as any song phrase in which the syllable type of that phrase had at least two  
202 occurrences separated by a silent period. Lincoln's sparrows also produce non-trilled phrases,  
203 as well as occasional trills with no silent portion between syllables, but these types of phrases,  
204 when they occurred in our stimulus songs, were not counted among the four manipulated trills  
205 and were not altered in any way. The song-types (i.e., the unique combination of syllable types  
206 within a song) differed between the 10 stimulus males, but, among the 4-5 songs of a single  
207 stimulus male, they were the same. Visualizing the songs as spectrograms in Raven Pro, we

208 cut 15 ms of silence from the space between each syllable of each trill of each song and  
209 pasted it into the corresponding inter-syllable space of a digital copy of the song, thereby  
210 generating two copies of each song that were identical except for the duration of silence  
211 between the syllables of their trills and thus the duration of the entire song (see Lyons et al.,  
212 2014 for a more detailed description of this technique) (Figure 1). This was the specific  
213 manipulation we have done in previous experiments on Lincoln's sparrows and was chosen to  
214 provide significant variation in syllable rate but without exceeding natural variation in this  
215 parameter (see below). The result was one set of 4-5 songs from each of 10 males with  
216 relatively long gaps of silence between each trill's syllables (low-performance treatment) and  
217 another set of the same songs from the same males with relatively short gaps of silence  
218 between each trill's syllables (high-performance treatment).

219 We then appended a silent period to the end of each stimulus song (see below),  
220 concatenated all of the low-performance stimuli from one stimulus male in random order into a  
221 single sound file, and repeated the concatenated set until the file totaled 2.5 min in duration.  
222 We then appended 1 min of silence followed by one repetition of the above 2.5 min of song.  
223 We made a second sound file of the stimulus male's corresponding high-performance stimuli in  
224 the same way and in the same song order. We then repeated this process for each of the  
225 remaining nine stimulus males, thereby producing 10 pairs of 6-min stimulus files (20 files in  
226 total) of identical songs in which one member of each pair consisted of the low-performance  
227 stimuli and the other of the high-performance stimuli from a single of the 10 stimulus males.

228 Because the experimental manipulation of syllable rate necessarily made low  
229 performance songs slightly longer than high-performance songs, the duration of silence  
230 between each song varied but was always approximately 10 sec, thereby producing a file with  
231 a rate of song production normal for the population. With an average song duration of  
232 approximately 2.5 sec, a trill count of 4 per song, an average syllable count of 4 per trill, and  
233 15 ms added or removed between each syllable, songs rendered as low performance differed  
234 from songs rendered as high performance by approximately 14% in duration, and thus the  
235 silent period that followed songs differed by approximately 7%. The members of a stimulus pair  
236 were otherwise identical, in that they contained the same songs recorded from the same

237 stimulus male. Compared to other stimulus pairs, they contained the same number of trills and  
238 were repeated the same number of times, at the same rate, and at the same peak amplitude.

239 We measured trill performance in spectrograms Raven Pro produced from the stimulus  
240 songs and from the songs of experimental subjects and non-experimental subjects (recorded  
241 from 2005-2010). We used measurements from the songs of non-experimental subjects not  
242 only to calculate the upper bound regression for quantifying trill performance but also for  
243 comparing with measures of songs from experimental subjects. A trill's performance measure  
244 is based on the relationship between its syllable-production rate and its frequency bandwidth,  
245 both of which can involve ambiguity in measurement. Syllable-production rate is the syllable  
246 count per second. Whole trills are separated from other phrases of a song by periods of  
247 silence. Therefore, for the purpose of objectively and reliably measuring syllable-production  
248 rate, we delineated a timespan from a distinct acoustic landmark on the first syllable of the trill  
249 to the corresponding landmark on the last syllable of the trill, thereby capturing the timespan of  
250 all sound and following silent periods of all but one syllable in the trill. We then calculated  
251 syllable rate as 1 less than the number of syllables in the trill divided by this timespan in  
252 seconds. We calculated trill duration as the timespan between identical landmarks of the first  
253 and last syllable (as mentioned above) plus the ratio of that same timespan and 1 less than the  
254 syllable count. This approach of temporally delineating all but one syllable and its following  
255 silent period does not necessitate knowing exactly where in the song the trill, which is  
256 bracketed by periods of silence, begins and ends and is not affected by sound pressure level  
257 or gain, assuming the signal to noise ratio is sufficient for clear visualization of the trill. And,  
258 indeed, our measure of trill duration includes an estimated duration of silence following the final  
259 syllable. Apparent frequency bandwidth is also affected by gain and can therefore be difficult  
260 to quantify objectively. Thus, we simultaneously delineated each trill's high and low frequencies  
261 using a software routine that calculates the upper and lower boundaries for the middle 90% of  
262 sound energy. Using this subset of sound energy for the calculation of frequency bandwidth  
263 enables measures that are not sensitive to sound pressure level and gain settings. It also  
264 favors the highest energy frequencies, excluding frequencies that are very low in sound energy  
265 and which would otherwise contribute disproportionately to bandwidth. We then calculated trill

266 performance as the perpendicular deviation from an upper bound regression of maximal  
267 frequency bandwidth on the center of syllable-rate bins spanning ranges of three syllables per  
268 second (Sockman, 2009), with larger, less negative values representing higher performance.  
269 For some species, the use of alternative methods to the upper-bound-regression has been  
270 recommended (Wilson et al., 2014), but we retain use of it for consistency with other studies on  
271 Lincoln's sparrows and because the Lincoln's sparrow was one species for which such an  
272 approach supported the presence of a performance trade-off (Wilson et al., 2014).

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

278

#### 279 *Field Procedures*

280

281 We conducted all stimulus playbacks in 2013 from June 04-18 and initiated them between  
282 0559 and 0750 hours local time. We began each procedure by locating a territorial male  
283 (hereafter the experimental subject or subject) that had been banded previously with a unique  
284 combination of color bands (Sockman, 2009), thereby making him identifiable at the individual  
285 level. We then placed a speaker (Pignose, Legendary 7-100, Las Vegas, NV) on the ground  
286 facing up centrally in his territory, and connected it to a digital audio player (Apple iPod,  
287 Cupertino, CA), loaded with each of the 20 6-min sound files. After a 10-min period of  
288 acclimation, we initiated the 6-min broadcast of a sound file to the subject. Each subject was  
289 assigned the stimulus pair from one and only one stimulus male, and we worked our way  
290 through the 10 pairs of sound files in random order (without replacement, until using all of  
291 them, at which point we re-randomized their order and repeated their use) with each new  
292 subject we encountered. We did not measure the sound pressure level of broadcasts, but we  
293 maintained identical settings on our audio equipment between playbacks. The initial choice of  
294 settings was based on a subjective estimate of the loudness of the normal, natural songs to

295 which we had been exposed thousands of times over previous years on this study site. During  
296 the broadcast and for 10 min following it, we collected digital audio recordings of the subject's  
297 song, as described above for stimulus males and non-experimental subjects. During the  
298 broadcast we also collected measures associated with non-vocal agonistic behavior, including  
299 counting the number of times the subject flew directly over the speaker and moved at least a  
300 whole body width (i.e., hops and turns) not including flights over the speaker, and timing the  
301 subject's duration spent within 1, 5, and 10 m of the speaker and more than 10 m from the  
302 speaker. We also measured the latency of subjects to approach within 1 m of the speaker, but,  
303 because some subjects never approached within 1 m, we did not use this measure in  
304 analyses.

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

316

### 317 *Analyses*

318

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

324 balanced with respect to treatment order and included only subjects that responded to both  
325 treatment levels, ensuring that the interaction between treatment order and treatment level was  
326 also balanced.

327         A total of 19 subjects were present (though not necessarily singing; see below) for both  
328 treatment levels. For analyses not involving within-song measurements (i.e., all non-song  
329 agonistic behavior and song counts), we had to exclude one subject (number 19, the final one)  
330 to enable the balance described above, yielding a sample size of 18 subjects (36 treatments)  
331 for each of these analyses. Only seven subjects sang *during* both playback treatments, with  
332 only six that would have enabled analyses that were balanced. Of these, the songs of two were  
333 not analyzable due to a low signal-to-noise ratio (see below). Thus, with such a small sample  
334 size, we did not perform analyses involving within-song measurements (the duration, trill count,  
335 and syllable count of songs and the duration, performance, frequency bandwidth, syllable rate,  
336 and syllable period [inverse of syllable rate] of trills). For analyses involving measurements of  
337 song *following* playback, we excluded the four subjects that did not sing following either  
338 playback treatment (as well as five subjects that sang following only one of the two treatment  
339 levels; analyses could not be balanced with the inclusion of any combination of them). This  
340 yielded a sample size of 10 subjects (20 treatments) for these analyses. Thus, we conducted  
341 all analyses of song measures (as opposed to song count and non-song agonistic behavior)  
342 on the post-playback song only of 10 subjects. With this approach, all analyses were balanced  
343 for treatment order and its interaction with treatment level.

344         We counted the songs of subjects from our digital audio recordings collected during  
345 and following playbacks. We could distinguish subject songs from stimulus songs in the audio  
346 recordings due to their differences in amplitude and to the unique spectral-temporal properties  
347 of their syllable-types. For each subject, we measured up to the first 10 songs. We could not  
348 measure some songs due to sound interference (e.g., wind, passing vehicle) or poor signal-to-  
349 noise ratio. For each trill, we determined syllable rate (syllables sec<sup>-1</sup>), trill frequency bandwidth  
350 (kHz), trill performance, trill syllable count, and trill duration (sec). We also determined the  
351 inverse of syllable rate, syllable period (sec syllable<sup>-1</sup>) to facilitate the visual assessment of trill

352 duration, which is simply the product of syllable period and trill syllable count. For each song,  
353 we determined its trill count, syllable count, and duration (sec).

354 We performed our statistical analyses using the software Stata/IC 15.1 for the  
355 Macintosh. In order to reduce dimensionality of our analyses, we performed a principal  
356 component analysis (PCA) on the correlation matrix of each of three sets of the above-  
357 mentioned behavioral responses. The first of the three sets combined all six non-song  
358 agonistic responses (speaker fly-over count; count of whole-body (non-flight) movements; and  
359 duration spent within 1, 5, and 10 m of the speaker and more than 10 m from the speaker) with  
360 song count during playback and song count following playback. Loadings of the first principal  
361 component were all similar in magnitude (absolute value 0.29-0.46), except fly-over count  
362 (-0.11) and song count during playback (0.08). Thus, we re-conducted the first PCA after  
363 removing fly-over count and song count during playback, yielding a first component which  
364 explained 66% of the total variance and which produced loadings all of similar magnitude  
365 (absolute value 0.29-0.47). The loadings of all except the duration more than 10 m from the  
366 speaker were positive. Thus we interpreted the first-axis factor scores from this as agonistic  
367 responses.

368 The remaining two PCA analyses examined various aspects of song behavior:  
369 measures of song length, each of which varied within each treatment; and measures of trills,  
370 each of which varied within each song. Thus, the second PCA combined song duration, song  
371 syllable count, and song phrase count, which produced all positive loadings of nearly identical  
372 magnitude (0.56-0.59) and explained 85% of the total variance. We interpreted its first-axis  
373 factor scores as measures of song-length response. Finally, the third PCA combined trill  
374 syllable count, syllable rate, frequency bandwidth, trill performance, and the proxy for trill  
375 duration described above, which produced all positive loadings of similar magnitude (0.30-  
376 0.57) and explained 46% of the total variance. We interpreted its first-axis factor scores as trill  
377 responses.

378 We statistically analyzed the effects of the experimental treatment using three general  
379 linear mixed models, each corresponding to one of the three PCA-produced response. Each of  
380 the models included the predictor treatment. Analyses of the third (trill response) also included

381 the song's trill number (modeled categorically, with trill number 1 as the reference group) and  
382 its interaction with treatment as predictors. Models differed based on the structure of their  
383 random effects, which we prescribed following Schielzeth and Forstmeier (2009). For the  
384 agonistic response, we nested observation ( $n = 36$ ) within subject ( $n = 18$ ) as a random  
385 intercept and random coefficient for treatment. For the song-length response, we nested  
386 observation ( $n = 188$ ) within treatment ( $n = 20$ ) as a random intercept and treatment within  
387 subject ( $n = 10$ ) as a random intercept and random coefficient for treatment. For the trill  
388 response, we nested observation ( $n = 652$ ) within song as a random intercept and random  
389 coefficient for trill number, song ( $n = 183$ ) within treatment as a random intercept, and  
390 treatment ( $n = 20$ ) within subject ( $n = 10$ ) as a random intercept and random coefficient for  
391 treatment.

392         The analysis of trill response produced significant results (see Results). So, for a post-  
393 hoc examination of trill response, we performed six additional mixed-effects models on syllable  
394 rate, bandwidth, trill performance, syllable period, trill syllable count, and trill duration, with  
395 treatment, trill number, and their interaction as predictors and observation ( $n = 682$  for trill  
396 syllable count,  $n = 652$  for all other measures [some syllables that, due to noise, were  
397 unanalyzable, could still be counted]) within song as a random intercept and random  
398 coefficient for trill number, song ( $n = 188$  for trill syllable count,  $n = 183$  for all other measures)  
399 within treatment as a random intercept, and treatment ( $n = 20$ ) within subject ( $n = 10$ ) as a  
400 random intercept and random coefficient for treatment. In figures we depicted marginal means  
401 with standard errors estimated for the treatment effect from the above models and thus  
402 controlling for random effects. We also conducted post-hoc contrasts of marginal means to  
403 compare treatment levels for each trill number.

404         For each experimental subject, we also used spectrograms in Raven to visually  
405 compare the syllable types of trills (trill types) between the two treatments, as well as between  
406 those produced by the experimental subjects and those of the stimulus songs, noting whether  
407 trill types were the same or different in those comparisons. The conclusions from these  
408 comparisons were unambiguous for all compared trills in this study and for nearly all we have  
409 ever inspected. As an example, in three different songs (one recorded from each of three



410 different males) (Figure 1), spectral traces readily reveal a total of 16 syllable types (and 12 trill  
411 types), particularly under magnification.

412 From 2005-2010 recordings of non-experimental subjects, we measured the trills of up  
413 to the first 10 songs recorded for each positively identified individual per day. Again,  
414 individuals had been banded previously with a unique combination of color bands (Sockman,  
415 2009), thereby making them identifiable at the individual level. Although we did not conduct  
416 inferential statistical analyses on the songs of these non-experimental males, we did account  
417 for the random effects of individual males ( $n \geq 106$ ) and songs ( $n \geq 2398$ ) in estimating  
418 marginal means for their trill ( $n \geq 8226$ ) measures for visual comparison with the trill measures  
419 of experimental subjects.

420

## 421 Results

422

423 We found no treatment effect on either of the first two PCA-generated responses. Treatment  
424 did not appear to strongly affect the song and non-song agonistic response ( $z = 1.34$ ,  $P =$   
425  $0.18$ ) or the song-length response ( $z = -1.11$ ,  $p > 0.2$ ). However, treatment did affect the trill  
426 response in its interaction with trill number (Figure 2). Specifically, the change in trill response  
427 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 ( $z =$   
428  $-1.91$ ,  $p = 0.056$ ) each depended on treatment.

429 Post-hoc analysis showed no support that treatment as a main effect influenced any of  
430 our six trill measures in songs of experimental subjects during the period immediately following  
431 the playback treatment (Table 2). However, we found a strong relationship between trill number  
432 and each of our trill measures, as well as a strong effect of the interaction between treatment  
433 and the trill number on trill performance (Table 2). Post-hoc contrasts of marginal means  
434 revealed this effect exclusively for trill three (Figure 3). Moreover, the effect of the interaction on  
435 trill performance was driven by its effect on frequency bandwidth, as effects on syllable rate  
436 were minimal at most (Table 2, Figure 3). Additionally, the interaction between treatment and  
437 trill number also affected trill syllable count, with post-hoc contrasts revealing that effect for  
438 trills two and four, resulting in a drop in the duration of trills two and four (Figure 3).

439 For nine of the ten experimental subjects used for trill analysis, all trills produced post-  
440 playback differed between treatments in their type. One subject produced all the same trill  
441 types following the low-performance treatment as he did following the high-performance  
442 treatment. Thus, the effect of treatment on the frequency bandwidth of trill three was due to a  
443 change between treatments in syllable type, with the syllable type following the high-  
444 performance treatment having lower frequency bandwidth than the syllable type following the  
445 low-performance treatment. Following the low-performance treatment, five, four, and one  
446 subjects matched zero, one (12-25% of trill types produced), and two (33%) trill types of the  
447 treatment, respectively, but whether or not there was a match did not depend on trill number  
448 (i.e., trill 3 was not necessarily the matched trill). Following the high-performance treatment,  
449 seven and three subjects matched zero and one (9%) trill types of the treatment, respectively,  
450 and again, whether or not there was a match did not depend on trill number. Given the high  
451 proportion of zero matches, we analyzed this difference using a logit model, in which the  
452 response was the dichotomous any matches (1) or no matches (0). A full, mixed model  
453 accounting for observation nested within subject would not converge on a solution, so we  
454 removed the nested random effect and conducted a general linear model, which showed no  
455 strong effect of treatment on the probability that any syllable produced following playback was  
456 matched ( $z = -0.85, p > 0.2$ ).

457

## 458 Discussion

459

460 The trill performance of a simulated rival has effects on the vocal output of wild, free-ranging  
461 male receivers, specifically on the bandwidth of their trills, which drive a change in their trill  
462 performance. If maximal trill performance is biomechanically constrained as we suggest  
463 above, then this plasticity in performance indicates males are not always performing at their  
464 maximum (Podos, 2017). These findings show that ecologically relevant exposure to natural  
465 variation in trill performance affects vocal behavior relevant to sexual signaling (Caro et al.,  
466 2010; Lyons et al., 2014).

467           Some species are known to modulate trills such that their performance changes (e.g.,  
468 DuBois et al., 2009), but, to our knowledge, ours is the first report that a facultative change in  
469 trill performance, in this case driven by facultative change in syllable type, can be driven by a  
470 simulated rival's trill performance. Additionally, although this is not the first study that explicitly  
471 addresses the question of how song of a receiver is affected by the trill performance to which it  
472 was recently exposed, it may be the first to show this effect using trill performance manipulated  
473 entirely within the natural range of variation over not only the song as a whole but also for each  
474 of multiple trills within the song. Liu and colleagues (2018) tested whether swamp sparrows are  
475 more likely to song-type match when exposed to songs of relatively low trill performance or to  
476 songs of relatively high trill performance, but they did not find a difference between treatments.  
477 Cramer (2013) similarly reported no effect of trill performance on song responses in house  
478 wrens (*Troglodytes aedon*). Moseley and colleagues (2013) found that the rate of soft song  
479 production in swamp sparrows was reduced by the digital reduction of the trill performance of  
480 playbacks in comparison to control, non-manipulated songs. However, several trills appear to  
481 have been manipulated such that they were outside the range of performance that is known to  
482 occur naturally for the species (see figure 1 in Moseley et al., 2013), so it is not clear whether  
483 the findings would apply to only natural variation in trill performance. Our findings here show  
484 that variation in trill performance within a natural range to which a male receiver is exposed can  
485 significantly influence his own song behavior.

486           Additionally, our study reveals that responses of experimental subjects are specific to  
487 the trill-number within the song, with the change in bandwidth and consequently trill-  
488 performance restricted to trill three of songs and the change in trill syllable-count and  
489 consequently trill duration occurring for both trill two and trill four (Figure 3). We are not aware  
490 of studies that have examined variation in trill performance as a function of trill number, but the  
491 results we present here raise questions regarding the function not only of trills in general but of  
492 each trill in multi-trill songs. Toward that end, it is interesting that performance in trill one is very  
493 low compared to other trills in the song (Figure 3). Perhaps trill one serves as a vocal warm-up  
494 for or contrast to subsequent trills males may attempt to maximize or serves some other

495 purpose entirely. Indeed the function of any of the trills is not clear, but below we discuss some  
496 possibilities.

497         Individuals did not appear to adjust their non-vocal agonistic behavior according to  
498 playback treatment. Perhaps a comparison between high- and low-quality receivers would  
499 have revealed a dichotomous response to high-performance playbacks, as Mosely and  
500 colleagues (2013) discovered for swamp sparrows. Unfortunately, we have no measures of the  
501 quality of our subjects, except, perhaps, in relation to their own trill performance. However, the  
502 effect of treatment on the subjects' trill performance precludes use of the subjects' trill  
503 performance as an independent measure of their quality. Regardless, we do not have a strong  
504 reason to conclude that the treatment-induced difference in vocal response (Figure 3) is a  
505 signal of aggressive intent (Searcy & Beecher, 2009; Hof & Podos, 2013). In other words,  
506 because they did not appear to flee or even distance themselves from high-performance  
507 playback relative to low-performance playback, we cannot conclude that their own high-  
508 performance trills produced in response to the low-performance treatment were intended to  
509 drive away the putative intruder. Moreover, song rate was lower during than following playback  
510 (data not shown). If songs were intended to repel an intruder, we would have expected the  
511 opposite. Thus, perhaps subjects were signaling something else, such as their quality to  
512 female eavesdroppers relative to the intruder (Logue & Forstmeier, 2008). In fact, female  
513 Lincoln's sparrows are known to form long-term memories of the songs of individual males  
514 (Beaulieu & Sockman, 2012) and of songs based explicitly on trill performance (Lyons et al.,  
515 2014), after which they express song preferences in a mate-choice context based on those  
516 memories. This might explain the high-performance response to the low-performance  
517 treatment. The high-performance treatment may have exceeded the average subject's own  
518 maximal performance, hypothetically resulting in an alternative, perhaps more feasible counter-  
519 response in the form of some behavior we did not assess. Experimental subjects did respond  
520 to high-performance trills by elevating the syllable count of their second trills relative to their  
521 response to low-performance trills (Figure 3). Moreover, fourth trills were much lower in syllable  
522 count following exposure to high-performance trills than those of non-experimental subjects or  
523 those following low-performance trills (Figure 3). In short, we do not know the intent of the

524 subjects nor the function of their responses, and thus further studies will be necessary to  
525 address these questions.

526         The treatment-induced change in trill performance was driven by the use of a syllable-  
527 type based on its frequency bandwidth rather than by a change in syllable rate, the component  
528 of trill performance that we manipulated in our playback stimuli and which, at first pass, would  
529 seem more amenable to flexible modulation by the subjects. Lincoln's sparrows clearly vary  
530 their syllable rate between the trills of individual songs (Figure 3), however we found no  
531 evidence that they varied it in response to treatment. Also, because all but one produced  
532 different syllable types between treatments, experimental subjects did not vary frequency  
533 bandwidth within a trill type but instead changed the trill type itself, producing types of greater  
534 bandwidth following exposure to low-performance trills and types of lower bandwidth following  
535 exposure to high-performance trills. Our study population produces more than 100 syllable  
536 types (including those in trills and non-trills) (Reinhardt and Sockman unpublished data), in line  
537 with the more than 200 types shown for several populations of Lincoln's sparrows across the  
538 state of California (Cicero & Benowitz-Fredericks, 2000). Given the high number of syllable  
539 types, a detailed understanding of the degree to which any single syllable type might vary in  
540 either syllable rate or frequency bandwidth is beyond the scope of this article. Nevertheless,  
541 despite some focused assessment, we have never found a syllable type to vary in bandwidth,  
542 although they can vary in center frequency (Reinhardt and Sockman unpublished data).  
543 Therefore, the treatment responses that we observed make sense as a means of facultatively  
544 changing the bandwidth of a trill, in that, for the spectrally complex trills of Lincoln's sparrows,  
545 varying bandwidth within a trill type could render the trill unrecognizable. Individual Lincoln's  
546 sparrows have a repertoire of 20 syllable types, on average (Cicero & Benowitz-Fredericks,  
547 2000), greatly exceeding the number of syllable types produced in a song (Figure 1) or even in  
548 a single bout of singing. Thus the lack of similarity in trill types between treatments could be  
549 due, in part, to temporal variation and the probability that in comparing any two days there is  
550 high likelihood of change. But that does not explain the strong tendency to vary trill type  
551 directionally with respect to its frequency bandwidth in response to the treatment, and thus, it

552 appears that some syllable types are chosen over others, depending on the quality of the song  
553 competition.

554         We do not know the ultimate reason for the change in trill type that led to a change in  
555 frequency bandwidth in response to song exposure that varies with trill performance, although  
556 one possibility is that there is no ultimate reason for the change other than as a mechanism to  
557 change frequency bandwidth and thus trill performance. Theoretical modeling supports the  
558 hypothesis that a singer should match the song type of a rival when the singer can produce the  
559 song type at a higher level of performance (Logue & Forstmeier, 2008; but see Liu et al., 2018  
560 regarding empirical work). Although there is no evidence of song-type matching in Lincoln's  
561 sparrows (Cicero & Benowitz-Fredericks, 2000), syllable-type matching might occur, raising  
562 the question of whether variation between treatments in the use of specific syllable types was  
563 due to matching following exposure to the low-performance treatment and matching avoidance  
564 following the high-performance treatment. However, in our study, very few of the subject's  
565 syllable types matched those in playbacks, regardless of treatment, and we found no statistical  
566 support for elevated syllable-type matching following the low-performance treatment over that  
567 following the high-performance treatment. So, although that does not refute the possibility that  
568 Lincoln's sparrows engage in some degree of syllable-type matching, syllable-type matching is  
569 unlikely to have driven our results. Further examination using playback stimuli constructed from  
570 syllable types known to be within the subjects' repertoires would be worthwhile. Nonetheless,  
571 our results suggest Lincoln's sparrows facultatively choose between syllable types that vary in  
572 frequency bandwidth, but additional studies are needed to further elucidate the vocal  
573 mechanisms for plasticity in trill performance.

574         In his critique of vocal performance studies in songbirds, Kroodsma (2017) proposed  
575 that a singer's combination of syllable rate and frequency bandwidth was primarily a result of  
576 the particular song type he had learned and not a product of a biomechanical constraint on  
577 vocal performance. In a response, Podos (2017) noted that the specific individuals beside  
578 which a young bird settles and therefore the specific songs he learns may, in part, be a  
579 function of the two birds' vocal proficiency relative to one another, suggesting that the learning  
580 of trills does not itself preclude the possibility that current measures of vocal deviation are

581 indicators of male quality. As mentioned above, Lincoln's sparrows are not known to engage in  
582 song-type matching. However, if they engage in syllable-type matching, then perhaps the  
583 subset of the syllables to which they are exposed during a critical learning period and that they  
584 eventually produce is restricted by their vocal ability.

585         As stated above, we observed no effect of our treatment on non-vocal agonistic  
586 responses, and yet, several other studies have reported such effects, as mentioned in the  
587 introduction (Illes et al., 2006; Cramer, 2013; Goodwin & Podos, 2014; Phillips & Derryberry,  
588 2017a). The reason for this disparity is not obvious. Trill performance in Lincoln's sparrows may  
589 primarily target females. Still, our experiment used only acoustic stimuli, and was designed to  
590 address variation in rival trill performance only. Perhaps the addition of a visual stimulus, such  
591 as a live decoy or a taxidermic mount, would have produced different results. The mere  
592 presence of playbacks itself did seem to stimulate agonistic responses, but this is speculation,  
593 given that we lacked a non-playback comparison group.

594         Finally, it is worth recalling that our digital manipulation of trill performance was, at its  
595 essence, a manipulation of syllable rate, which, in turn, changed the trill performance and also  
596 the trill duration and thus song duration of playback stimuli. In light of this present study alone,  
597 we cannot know whether subjects were affected by one of these parameters but not another,  
598 by a subset of them, or by all of them, and thus it remains possible that the driving factor was  
599 something other than trill performance, such as song duration. However, previous studies on  
600 Lincoln's sparrows showing elevated song rates in response to songs of not only higher trill  
601 performance but also of greater duration (Sewall et al., 2010) suggest that it was the trill  
602 performance or syllable rate in the present study to which they preferentially attended and not  
603 song duration. This interpretation is consistent with the hypothesis that trill performance (or at  
604 least syllable rate) can serve as a biologically relevant signal and that this relevance is likely  
605 despite, not because of any difference in the duration of the signal.

606         Over approximately a quarter century since the first analyses of a trade-off between  
607 syllable rate and frequency bandwidth (Podos, 1997), many studies have addressed the  
608 hypothesis that the trills of bird songs are biologically meaningful measures of performance  
609 and possibly honest signals of male quality. As mentioned in the Introduction, these studies

610 have spanned a range, from observational field studies of correlations between fitness (e.g.,  
611 Nishida & Takagi, 2018) or condition (e.g., Ballentine, 2009) and trill performance to  
612 manipulative experiments in both the field and laboratory aimed at identifying causal  
613 relationships between trill performance and receiver responses (e.g., Illes et al., 2006; Lyons et  
614 al., 2014; Phillips & Derryberry, 2017a). Using stimuli that were experimentally manipulated  
615 within the range of natural variation and using an ecologically relevant field setting, we have  
616 shown in this study that trill performance of a simulated rival drives significant variation in  
617 receiver trill performance, not only fulfilling an important criterion in defining trill performance as  
618 a biologically meaningful signal (Laidre & Johnstone, 2013) but also expanding our  
619 understanding of this signal in its role in driving variation in sexual counter-signaling. We  
620 anticipate that future studies continue to elucidate the roles of trill performance and the  
621 mechanisms through which it affects receiver responses and ultimately lead to a better  
622 understanding of trill performance as an important communication signal in animals.

623

## 624 Acknowledgments

625

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

631

## 632 References

633

- 634 Andersson, M. (1994). *Sexual selection*. — Princeton University Press, Princeton, NJ.
- 635 Ballentine, B. (2009). The ability to perform physically challenging songs predicts age and size  
636 in male swamp sparrows, *Melospiza georgiana*. — *Anim. Behav.* 77: 973-978.
- 637 Ballentine, B., Hyman, J. & Nowicki, S. (2004). Vocal performance influences female response  
638 to male bird song: an experimental test. — *Behav. Ecol.* 15: 163-168.



- 639 Beaulieu, M. & Sockman, K. W. (2012). Song in the cold is 'hot': memory of and preference for  
640 sexual signals perceived under thermal challenge. — *Biol. Lett.* 8: 751-753.
- 641 Cardoso, G. C. (2017). Advancing the inference of performance in birdsong. — *Anim. Behav.*  
642 125: e29-e32.
- 643 Cardoso, G. C., Atwell, J. W., Ketterson, E. D. & Price, T. D. (2007). Inferring performance in  
644 the songs of dark-eyed juncos (*Junco hyemalis*). — *Behav. Ecol.* 18: 1051-1057.
- 645 Caro, S. P., Sewall, K. B., Salvante, K. G. & Sockman, K. W. (2010). Female Lincoln's sparrows  
646 modulate their behavior in response to variation in male song quality. — *Behav. Ecol.*  
647 21: 562-569.
- 648 Catchpole, C. K. & Slater, P. J. B. (1995). *Bird song*. — Cambridge University Press,  
649 Cambridge.
- 650 Cicero, C. & Benowitz-Fredericks, Z. M. (2000). Song types and variation in insular populations  
651 of Lincoln's sparrow (*Melospiza lincolni*), and comparisons with other *Melospiza*. — *Auk*  
652 117: 52-64.
- 653 Collins, S. (2004). Vocal fighting and flirting: the functions of birdsong. — In: *Nature's Music*  
654 (Marler, P. & Slabbekoorn, H., eds). Elsevier Academic Press, San Diego, CA, p. 39-79.
- 655 Cramer, E. R. A. (2013). Vocal deviation and trill consistency do not affect male response to  
656 playback in house wrens. — *Behav. Ecol.* 24: 412-420.
- 657 De Kort, S. R., Eldermire, E. R. B., Cramer, E. R. A. & Vehrencamp, S. L. (2009). The deterrent  
658 effect of bird song in territory defense. — *Behav. Ecol.* 20: 200-206.
- 659 Draganoiu, T. I., Nagle, L. & Kreutzer, M. (2002). Directional female preference for an  
660 exaggerated male trait in canary (*Serinus canaria*) song. — *Proc. Roy. Soc. Lond. B*  
661 269: 2525-2531.
- 662 Dubois, A. L., Nowicki, S. & Searcy, W. A. (2009). Swamp sparrows modulate vocal  
663 performance in an aggressive context. — *Biol. Lett.* 5: 163-165.
- 664 —. (2011). Discrimination of vocal performance by male swamp sparrows. — *Behav. Ecol.*  
665 *Sociobiol.* 65: 717-726.
- 666 Gentner, T. Q. & Hulse, S. H. (2000). Female European starling preference and choice for  
667 variation in conspecific male song. — *Anim. Behav.* 59: 443-458.
- 668 Goodwin, S. E. & Podos, J. (2014). Team of rivals: alliance formation in territorial songbirds is  
669 predicted by vocal signal structure. — *Biol. Lett.* 10: 20131083.

- 670 Graham, E. B., Caro, S. P. & Sockman, K. W. (2011). Change in offspring sex ratio over a very  
671 short season in Lincoln's sparrows: the potential role of bill development. — *J. Field*  
672 *Ornithol.* 82: 44-51.
- 673 Hof, D. & Podos, J. (2013). Escalation of aggressive vocal signals: a sequential playback study.  
674 — *Proc. Roy. Soc. Lond. B* 280: 20131553.
- 675 Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. —  
676 *Ecol. Monogr.* 54: 187-211.
- 677 Illes, A. E., Hall, M. L. & Vehrencamp, S. L. (2006). Vocal performance influences male receiver  
678 response in the banded wren. — *Proc. Roy. Soc. Lond. B* 273: 1907-1912.
- 679 Kroodsma, D. (2017). Birdsong performance studies: a contrary view. — *Anim. Behav.* 125: e1-  
680 e16.
- 681 Laidre, M. E. & Johnstone, R. A. (2013). Animal signals. — *Curr. Biol.* 23: R829-R833.
- 682 Leitão, A., Ten Cate, C. & Riebel, K. (2006). Within-song complexity in a songbird is meaningful  
683 to both male and female receivers. — *Anim. Behav.* 71: 1289-1296.
- 684 Linhart, P., Jaška, P., Petrusková, T., Petrusek, A. & Fuchs, R. (2013). Being angry, singing  
685 fast? Signalling of aggressive motivation by syllable rate in a songbird with slow song. —  
686 *Behav. Processes* 100: 139-145.
- 687 Liu, I. A., Soha, J. A. & Nowicki, S. (2018). Song type matching and vocal performance in  
688 territorial signalling by male swamp sparrows. — *Anim. Behav.* 139: 117-125.
- 689 Logue, D. M. & Forstmeier, W. (2008). Constrained performance in a communication network:  
690 implications for the function of song-type matching and for the evolution of multiple  
691 ornaments. — *Am. Nat.* 172: 34-41.
- 692 Lyons, S. M., Beaulieu, M. & Sockman, K. W. (2014). Contrast influences female attraction to  
693 performance-based sexual signals in a songbird. — *Biol. Lett.* 10: 20140588.
- 694 Marler, P. & Peters, S. (1977). Selective vocal learning in a sparrow. — *Science* 198: 519-521.
- 695 Maynard Smith, J. & Harper, D. G. C. (1995). Animal signals: models and terminology. — *J.*  
696 *Theor. Biol.* 177: 305-311.
- 697 Mcgregor, P. K. (1991). The singer and the song: on the receiving end of bird song. — *Biol.*  
698 *Rev.* 66: 57-81.
- 699 Moseley, D. L., Lahti, D. C. & Podos, J. (2013). Responses to song playback vary with the vocal  
700 performance of both signal senders and receivers. — *Proc. Roy. Soc. Lond. B* 280:  
701 20131401.

- 702 Mountjoy, D. J. & Lemon, R. E. (1991). Song as an attractant for male and female European  
703 starlings, and the influence of song complexity on their response. — *Behav. Ecol.*  
704 *Sociobiol.* 28: 97-100.
- 705 Nelson, D. A. & Poesel, A. (2007). Segregation of information in a complex acoustic signal:  
706 individual and dialect identity in White-crowned sparrow song. — *Anim. Behav.* 74:  
707 1073-1084.
- 708 Nishida, Y. & Takagi, M. (2018). Song performance is a condition-dependent dynamic trait  
709 honestly indicating the quality of paternal care in the bull-headed shrike. — *J. Avian Biol.*  
710 2018: e01794.
- 711 Pandolfino, E. R., Ammon, E. M. & Sockman, K. W. (2023). Lincoln's sparrow (*Melospiza*  
712 *lincolni*), version 2.0. — In: *Birds of the World* (Rodewald, P. G., ed). Cornell Lab of  
713 Ornithology, Ithaca, NY.
- 714 Phillips, J. N. & Derryberry, E. P. (2017a). Equivalent effects of bandwidth and trill rate: support  
715 for a performance constraint as a competitive signal. — *Anim. Behav.* 132: 209-215.  
716 —. (2017b). Vocal performance is a salient signal for male-male competition in white-crowned  
717 sparrows. — *Auk* 134: 564-574.
- 718 Podos, J. (1997). A performance constraint on the evolution of trilled vocalizations in a songbird  
719 family (Passeriformes: Emberizidae). — *Evolution* 51: 537-551.
- 720 —. (2001). Correlated evolution of morphology and vocal signal structure in Darwin's finches. —  
721 *Nature* 409: 185-188.
- 722 Podos, J. (2017). Birdsong performance studies: reports of their death have been greatly  
723 exaggerated. — *Anim. Behav.* 125: e17-e24.
- 724 Podos, J. & Nowicki, S. (2004). Performance limits on birdsong production. — In: *Nature's*  
725 *Music: The Science of Bird Song* (Marler, P. & Slabbekoorn, H., eds). Academic Press,  
726 New York, p. 318-342.
- 727 Ritschard, M., Riebel, K. & Brumm, H. (2010). Female zebra finches prefer high-amplitude song.  
728 — *Anim. Behav.* 79: 877-883.
- 729 Schielzeth, H. & Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in  
730 mixed models. — *Behav. Ecol.* 20: 416-420.
- 731 Searcy, W. A. & Beecher, M. D. (2009). Song as an aggressive signal in songbirds. — *Anim.*  
732 *Behav.* 78: 1281-1292.
- 733 Searcy, W. A. & Nowicki, S. (2000). Male-male competition and female choice in the evolution  
734 of vocal signaling. — In: *Animal Signals: Signalling and Signal Design in Animal*

- 735           Communication (Espmark, Y., Amundsen, T. & Rosenqvist, G., eds). Tapir Academic  
736           Press, Trondheim, Norway, p. 301-315.
- 737 Sewall, K. B., Dankoski, E. C. & Sockman, K. W. (2010). Song environment affects singing  
738           effort and vasotocin immunoreactivity in the forebrain of male Lincoln's sparrows. —  
739           Horm. Behav. 58: 544-53.
- 740 Sockman, K. W. (2008). Ovulation order mediates a trade-off between pre-hatching and post-  
741           hatching viability in an altricial bird. — PLOS One 3: e1785.
- 742 —. (2009). Annual variation in vocal performance and its relationship with bill morphology in  
743           Lincoln's sparrows, *Melospiza lincolnii*. — Anim. Behav. 77: 663-671.
- 744 —. (2016). The regulation of behavioral plasticity by performance-based feedback and an  
745           experimental test with avian egg production. — Am. Nat. 187: 564-575.
- 746 Sockman, K. W., Salvante, K. G., Racke, D. M., Campbell, C. R. & Whitman, B. A. (2009). Song  
747           competition changes the brain and behavior of a male songbird. — J. Exp. Biol. 212:  
748           2411-2418.
- 749 Sockman, K. W., Sewall, K. B., Ball, G. F. & Hahn, T. P. (2005). Economy of mate attraction in  
750           the Cassin's finch. — Biol. Lett. 1: 34-37.
- 751 Vehrencamp, S. L. (2001). Is song-type matching a conventional signal of aggressive  
752           intentions? — Proc. Roy. Soc. Lond. B 268: 1637-1642.
- 753 Vehrencamp, S. L., De Kort, S. R. & Illes, A. E. (2017). Response to Kroodsma's critique of  
754           banded wren song performance research. — Anim. Behav. 125: e25-e28.
- 755 Williams, T. D. (2012). Physiological Adaptations for Breeding in Birds. — Princeton University  
756           Press, Princeton, NJ.
- 757 Wilson, D. R., Bitton, P.-P., Podos, J. & Mennill, D. J. (2014). Uneven sampling and the analysis  
758           of vocal performance constraints. — Am. Nat. 183: 214-228.
- 759
- 760

**Table 1.** Range of syllable rate (syllables sec<sup>-1</sup>) and trill performance with respect to the within-song trill number for Lincoln's sparrow (*Melospiza lincolni*) natural trills and trills that were digitally manipulated for simulated territorial intrusions.

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

**Table 2.** Statistical effects of trill-performance playback treatment (0: low, 1: high), trill number (trill 1 as reference), and their interaction (X) on six trill measures in Lincoln's sparrows (*Melospiza lincolni*). 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.

Response					Response				
Predictor	Coefficient	SE	$z$	$p$	Predictor	Coefficient	SE	$z$	$p$
<b>Trill Performance</b>					<b>Syllable Period (sec syllable<sup>-1</sup>)</b>				
intercept	-1.399	0.070	-19.890	< 0.001	intercept	0.229	0.007	33.120	< 0.001
treatment	-0.044	0.086	-0.52	> 0.2	treatment	-0.018	0.010	-1.79	0.073
trill 2	0.142	0.057	2.480	0.013	trill 2	-0.072	0.005	-13.730	< 0.001
trill 3	0.471	0.059	8.030	< 0.001	trill 3	-0.085	0.005	-15.590	< 0.001
trill 4	0.340	0.062	5.500	< 0.001	trill 4	-0.102	0.006	-17.790	< 0.001
treatment X trill 2	0.075	0.083	0.900	> 0.2	treatment X trill 2	0.016	0.008	2.120	0.034
treatment X trill 3	-0.301	0.084	-3.570	< 0.001	treatment X trill 3	0.011	0.008	1.430	0.15
treatment X trill 4	-0.001	0.093	-0.010	> 0.2	treatment X trill 4	0.021	0.009	2.390	0.017
<b>Syllable Rate (syllables sec<sup>-1</sup>)</b>					<b>Trill Syllable Count</b>				
intercept	4.572	0.291	15.700	< 0.001	intercept	2.453	0.178	13.750	< 0.001
treatment	0.390	0.416	0.94	> 0.2	treatment	0.183	0.228	0.80	> 0.2
trill 2	2.021	0.219	9.240	< 0.001	trill 2	0.745	0.190	3.920	< 0.001
trill 3	2.754	0.226	12.190	< 0.001	trill 3	1.003	0.194	5.160	< 0.001
trill 4	3.825	0.240	15.930	< 0.001	trill 4	1.023	0.200	5.110	< 0.001
treatment X trill 2	-0.169	0.318	-0.530	> 0.2	treatment X trill 2	0.422	0.275	1.540	0.12
treatment X trill 3	0.152	0.324	0.470	> 0.2	treatment X trill 3	-0.062	0.279	-0.220	> 0.2
treatment X trill 4	-0.365	0.366	-1.000	> 0.2	treatment X trill 4	-0.821	0.298	-2.760	0.006
<b>Frequency Bandwidth (kHz)</b>					<b>Trill Duration (sec)</b>				
intercept	1.271	0.067	18.830	< 0.001	intercept	0.543	0.017	31.500	< 0.001
treatment	-0.080	0.097	-0.83	> 0.2	treatment	-0.008	0.020	-0.37	> 0.2
trill 2	-0.007	0.053	-0.140	> 0.2	trill 2	-0.057	0.018	-3.220	0.001
trill 3	0.275	0.055	5.010	< 0.001	trill 3	-0.086	0.018	-4.750	< 0.001
trill 4	0.092	0.058	1.580	0.11	trill 4	-0.112	0.019	-5.850	< 0.001
treatment X trill 2	0.095	0.077	1.230	> 0.2	treatment X trill 2	0.087	0.026	3.410	0.001
treatment X trill 3	-0.308	0.079	-3.920	< 0.001	treatment X trill 3	0.010	0.026	0.380	> 0.2
treatment X trill 4	0.007	0.089	0.080	> 0.2	treatment X trill 4	-0.065	0.029	-2.230	0.026

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

771

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

778

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

788

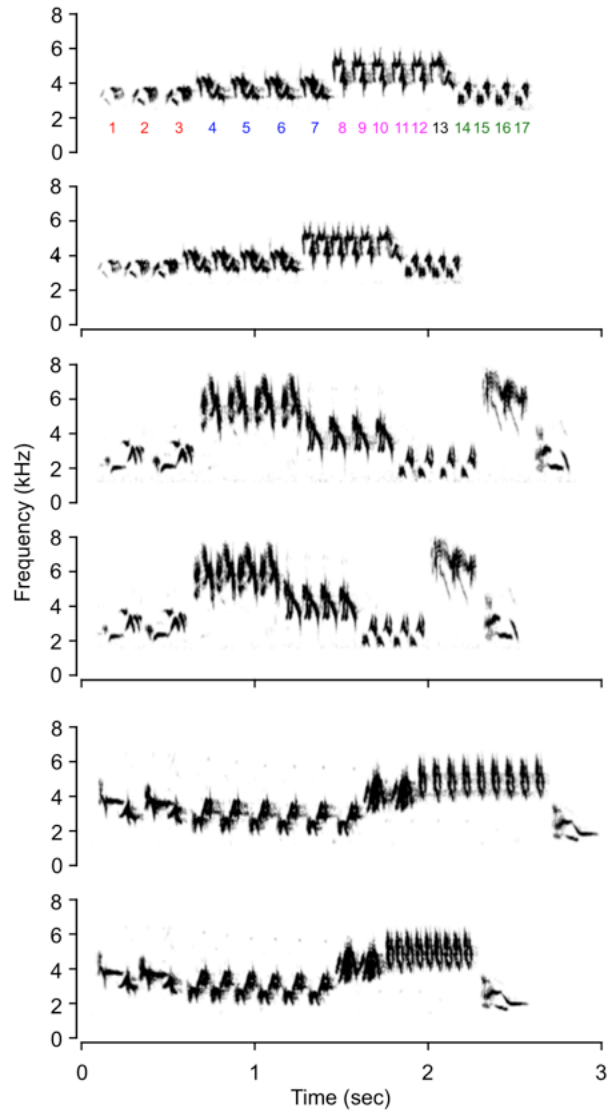


Figure 1

789

790



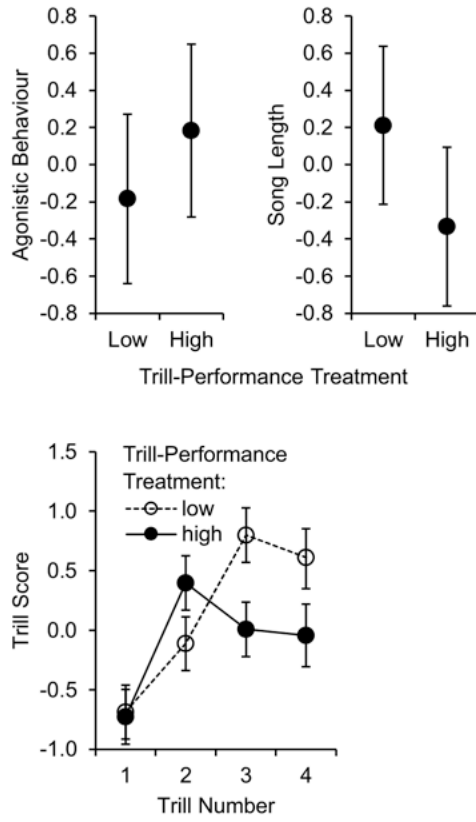
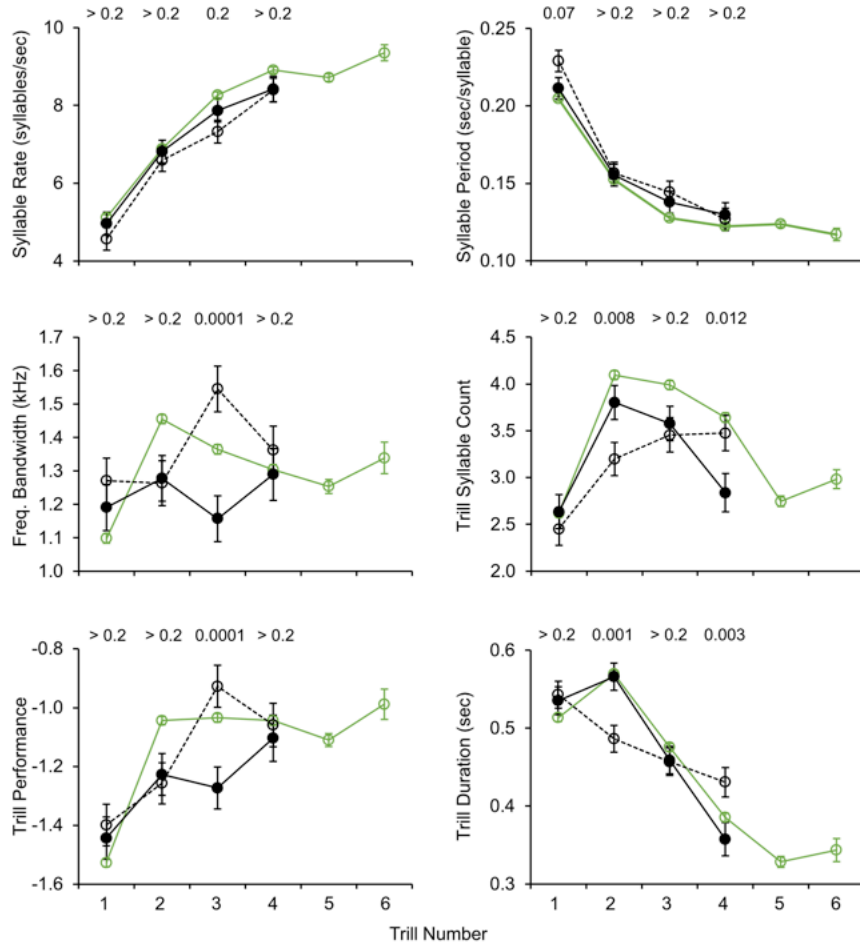


Figure 2

791

792



793

Figure 3