



RESEARCH ARTICLE

## Are commonly used metrics of bird song complexity concordant?

Lauryn Benedict,\* and Nadjé A. Najar

School of Biological Sciences, University of Northern Colorado, Greeley, Colorado, USA

\* Corresponding author: [lauryn.benedict@unco.edu](mailto:lauryn.benedict@unco.edu)

Submitted March 26, 2018; Accepted September 18, 2018; Published February 5, 2019

### ABSTRACT

Researchers studying the evolution of animal communication often ask what generates and maintains signal complexity, but they define and measure complexity in different ways. If different metrics are not concordant, then studies using them are probably not comparable. In this study, we asked whether 7 metrics of bird song complexity vary in tandem among individuals of a single species, the Rock Wren (*Salpinctes obsoletus*). The included metrics were chosen because they are regularly used by researchers in the field to test hypotheses within the literature. Results indicated that none of the metrics positively correlated with any others, suggesting that bird songs are not under general selective pressures favoring increased complexity across multiple features. If metrics of signal complexity are not correlated, then care should be taken when designing analyses and comparing results. Researchers should always clearly define the “complexity” under investigation and verify that it has significance to the study species. Contradictory outcomes among existing research into signal “complexity” may result from the fact that this single word is used to represent multiple independent traits.

**Keywords:** bird song, Rock Wren, *Salpinctes obsoletus*, signal complexity, signal repertoires, song complexity

### ¿Son concordantes las métricas comúnmente usadas para medir la complejidad del canto de las aves?

### RESUMEN

Los investigadores que estudian la evolución de la comunicación de los animales usualmente se preguntan qué genera y mantiene la complejidad de las señales, pero definen y miden esta complejidad de modos diferentes. Si las diferentes métricas usadas no son concordantes, entonces los estudios que las usan probablemente no son comparables. En este estudio, nos preguntamos si siete métricas de la complejidad del canto de las aves varían en conjunto entre individuos de una única especie, *Salpinctes obsoletus*. Las métricas incluidas fueron elegidas debido a que son comúnmente usadas en el campo por los investigadores para evaluar hipótesis comunes de la literatura. Los resultados indicaron que ninguna de las métricas se correlacionó positivamente con cualquiera de las otras, sugiriendo que los cantos de las aves no están bajo presiones selectivas generales que favorezcan un aumento de la complejidad a través de múltiples funciones. Si las métricas de la complejidad de las señales no están correlacionadas, entonces se debe tener cuidado cuando se diseñan los análisis y se comparan los resultados. Los investigadores deberían siempre definir claramente la “complejidad” bajo estudio y verificar su importancia para la especie estudiada. Los resultados contradictorios entre las investigaciones existentes sobre la “complejidad” de las señales pueden resultar del hecho de que esta única palabra es usada para representar múltiples rasgos independientes.

**Palabras clave:** canto de las aves, complejidad del canto, complejidad de la señal, repertorios de la señal, *Salpinctes obsoletus*

### INTRODUCTION

Complex traits are of interest to biologists because they beg the question of how simple substrates aggregate to produce increased functionality (Darwin 1859, Herron and Michod 2008, Bradbury and Vehrencamp 2014). The evolution of communication signal complexity, in particular, has been investigated widely across animal taxa, and often invites parallels with studies of human language complexity (Smith and Harper 2003). Animal communication researchers have defined many different axes of signal complexity

including variation in signal form, components, organization and modalities (Hebets and Papaj 2005, Bradbury and Vehrencamp 2011, Kershenbaum 2014, Taylor and Cody 2015). Although combinations of these measures are frequently discussed together as “complexity”, they are likely to be independent and research should account for any distinctiveness among signal components (Hebets et al. 2016, Taylor and Cody 2015).

The tendency to equate multiple measures of signal form is common in studies of bird song where researchers often generalize complexity as a suite of traits or apply the same

hypotheses to different metrics of complexity. Multiple studies have found that different song features have different functions (Botero and Vehrencamp 2007, Wilkins et al. 2015) and yet traits relating to song or syllable repertoire sizes, song structure and delivery patterns are regularly discussed together as components of monolithic complexity, and parallel predictions are made for all metrics (e.g., Palmero et al. 2014, Kaluthota et al. 2016). This can cause confusion when comparing studies. For example, the hypothesis that song complexity increases with latitude has been both supported and refuted by studies that measured complexity in different ways (Soma and Garamszegi 2011, Singh and Price 2015, Najar and Benedict 2018). Similarly, researchers testing the “Cognitive Capacity Hypothesis” have found that song complexity shows no consistent correlation with cognitive task performance across 3 avian species, but studies of those 3 species used different metrics of “complexity” (Templeton et al. 2014). It is possible, therefore, that the lack of concordance stems not from an underlying pattern, but from the fact that the studies measured different signal features. This issue exists not just in cross-species comparisons within the literature, but also within species. For example, research has variously concluded that developmental stress does and does not affect bird song complexity (Spencer and MacDougall-Shackleton 2011), but seemingly contradictory results within Zebra Finches (*Taeniopygia guttata*) come from datasets that used different metrics of complexity (for an example, see Zann and Cash 2008 and Brumm et al. 2009).

The contrasting results of existing research may sometimes be due to an assumption that all “complexity” variables show similar (or at least correlated) patterns of evolution. Individual authors rarely make this explicit claim, but the field as a whole regularly measures different complexity variables interchangeably (Najar and Benedict 2018). To test whether this assumption has validity within a single species we can assess whether individual birds sing songs that are complex in multiple dimensions. There is limited evidence to this effect (Buskirk 1997, Soma et al. 2006), but there is also evidence that elements of song complexity evolve independently and that different measures of complexity are not equivalent in non-avian systems (Botero and Vehrencamp 2007, Kershenbaum et al. 2017). Some studies have explicitly tested for correlations between multiple bird song “complexity” variables (Cardoso and Mota 2007, Woodgate et al. 2012, dos Santos et al. 2018), and they generally find the strongest positive correlations between variables that are linked a priori, such as song length and the number of elements or the number of inflections within a song (Cardoso and Mota 2007, Woodgate et al. 2012, Kaluthota et al. 2016). It is less well known whether correlations exist between variables that are structurally independent.

Here, we assess correlations between 7 potentially independent metrics of bird song complexity. We chose metrics that are commonly used in the literature to describe variation in song repertoire sizes, delivery patterns and acoustic form (see the Appendix for a list of recent studies using these variables). The choice of complexity measures for any study depends on the signal form and delivery patterns of the study species (Kershenbaum et al. 2017). We chose 7 “complexity” variables that are measurable for our study species, the Rock Wren (*Salpinctes obsoletus*), and distinct enough from each other not to be a priori correlated. Our goal was not to identify the “best” measure of complexity for this (or other) species, but to ask whether existing, published metrics of complexity represent a single axis of elaboration for individual birds. We did not assess all measures of “complexity” that have been proposed by researchers in the field, but instead focused on a subset of variables that have been applied to questions relating to song evolution (Appendix). We avoided metrics, such as those relating to performance or entropy, which are composites of features that are sometimes measured independently (Kershenbaum 2014, Podos et al. 2016). The concepts of performance and complexity are intertwined (e.g., Kaluthota et al. 2016) and multiple of our metrics could represent either, but here we focus on them as components of complexity.

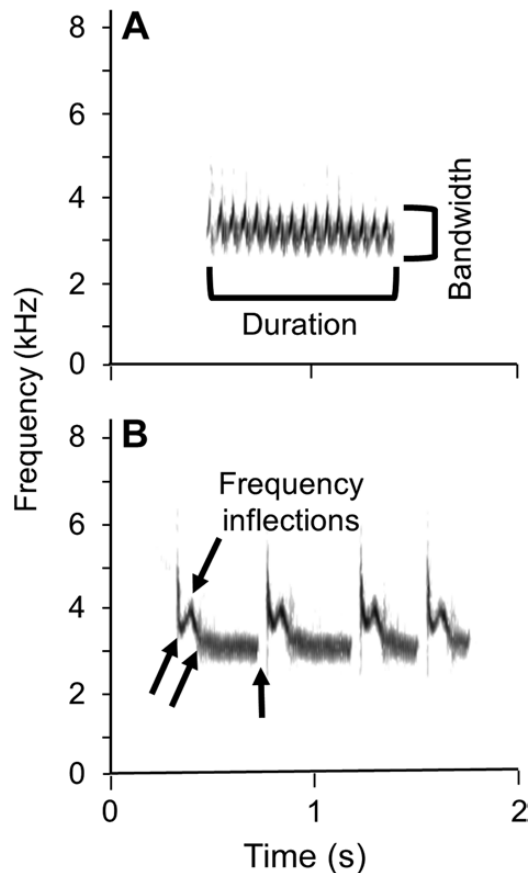
If natural and/or sexual selection promote overall increases in complexity, then individuals should exist along a gradient, with individuals singing songs that are more vs. less complex across multiple metrics. Therefore, positively correlated patterns of complexity among multiple measures across individuals would support the hypothesis that song complexity is a general phenomenon consisting of multiple linked traits. Alternatively, if individual aspects of song complexity evolve independently, complexity metrics are not predicted to correlate across individuals. It is also possible that tradeoffs exist between complex song features, in which case some complexity metrics should be negatively correlated across individuals.

## METHODS

We calculated 7 song complexity measures for 12 wild Rock Wrens (*S. obsoletus*) from one population in northern Colorado, and tested for correlations between these measures using a series of linear regressions. Our sampling focused on obtaining an extensive set of song recordings from each bird to accurately reflect average individual behavior. We recorded birds singing natural broadcast song across multiple days ( $2.75 \pm 1.42$  visits per bird, 4–12 hr of total recording time per bird) during May and June of 2013 and 2014, and examined a mean of 1,588 songs per bird (range: 924–2077) using Raven Sound Analysis Software.

We did not use playback to stimulate singing, and we avoided recording birds that were involved in countersinging or other social interactions (for more details on recording and measurement methods see [Benedict and Warning 2017](#) or [Hedley et al. 2018](#)).

Rock Wrens have large, diverse song repertoires providing an excellent model system for quantification of multiple complexity metrics ([Kroodsma 1975](#)). They sing brief songs and present them in sequences that include some repetition of song types and some introduction of new song types in unpredictable sequences ([Hedley et al. 2018](#)). Rock Wren song types vary widely in acoustic structure ([Figure 1](#)) ([Benedict and Warning 2017](#)). We chose to measure the following 7 variables following existing literature on bird song complexity (Appendix), making sure they were not a priori correlated due to either biological constraints or calculation method. Some of these measures may not appear to be obviously representative of song complexity, but they have been used by other researchers in this context (e.g., bandwidth, 8 studies cited in the [Appendix](#)).



**FIGURE 1.** Spectrograms of 2 Rock Wren song types and some measured features. (A) Song duration and bandwidth. (B) Frequency inflections for a single syllable.

1. Repertoire = the number of total song types after 900 songs recorded. Simple enumeration model.
2. Delivery (songs per min) = the number of songs recorded over total recording time. We only recorded during bouts of continuous broadcast singing. Songs were separated by at least 1 s, but longer gaps of >1 min between song bouts are not included.
3. Switching (proportion) = switching rate calculated as the proportion of songs which represent a type switch from the preceding song.
4. Variety = short-term song diversity, representing the number of song types used in a series of 10 songs.
5. Duration (s) = average song duration ([Figure 1](#)).
6. Bandwidth (Hz) = average song frequency bandwidth ([Figure 1](#)).
7. Inflections per s = the average number of frequency inflections per second across all songs, defining a frequency inflection as either a change from increasing to decreasing frequency, a change from decreasing to increasing frequency, or a break in a continuous trace on the spectrogram where the sound following the break is at a different frequency than the sound preceding the break ([Figure 1](#)).

Repertoire size was calculated from 900 songs. Because this is a cumulative value, we standardized the sample size using our smallest available sample. Song type accumulation curves begin to level out at 500 songs recorded for Rock Wrens ([Kroodsma 1975](#), [Benedict and Warning 2017](#)), such that song repertoire sizes estimated via simple enumeration at 900 songs are expected to be slightly lower than the true repertoire size, but are a representative comparison across individuals ([Botero et al. 2008](#), [Luttrell et al. 2016](#)). Full repertoires are revealed over the course of many hr, therefore we also included some short-term complexity measures: Delivery, Switching, and Variety represent the complexity of singing style that would be evident to a listener over a period of a single min. Delivery was calculated from all 19,058 songs in our dataset. Switching and Variety were calculated from a subsample of 100 songs per bird: 10 sets of 10 songs in series with the identity of those sets randomly chosen out of all recorded song bouts. In each series we counted the number of song types used or the number of switches in 10 opportunities. Reported individual values are means of the 10 sets per bird. We chose to use the means of 10 sets to ensure that these metrics were not influenced by targeted sampling of one singing bout. Other studies have measured variety as the number of syllables or elements types per song, but because Rock Wren songs consist of a single syllable type per song type, short-term variety must be measured via song type use.

Duration, Bandwidth and Inflections per s describe the form of songs and represent the mean value of that trait across all recorded songs per individual. To calculate

these measures we measured the duration, bandwidth or frequency inflections per s in one exemplar of each song type per bird (Figure 1). Because Rock Wrens have large repertoires, these measurements came from a total of 922 song types across all birds. For each song type we chose the highest quality recording for measurement. We multiplied measured values for each song type by the number of times the appropriate song type was used by that bird, summed all multiplied values and divided by the total number of songs sampled from that individual in order to calculate individual average song duration, bandwidth or frequency inflections per s. Across all individuals these calculations included all 19,058 songs in our dataset. Measuring just one exemplar of each song type per bird is representative of the features of all songs of that type because song types are highly stereotyped (Kroodsma 1975, Benedict and Warning 2017).

Traits were correlated against all others using pairwise linear regression between all seven variables, yielding 21 potential correlations with 12 data points each. Because individual complexity maximums may be as informative as averages for some variables, we also identified the maximum Duration, Bandwidth and Inflections per s values produced by each individual, and ran a second set of analyses with those values instead of the average trait values.

Because our dataset included only 12 individuals, non-significant results could reflect a lack of power. Therefore, we assessed the potential for our existing dataset to yield positive results by examining relationships between variables that have been shown to correlate in previous studies (e.g., Duration and Inflections per song rather than Inflections per s). We also ran a set of analyses designed to test whether our observed correlations fall above or below a random null expectation. To produce a null expectation, we created a model that used our data to generate a set of randomized correlations. The model randomly assigned one of the 12 data points within each of our 7 variables to each individual and correlated the randomized variables using the same method that we used for the observed data. This process was iterated 1,000 times, and our observed mean  $r^2$  value was compared with the distribution of 1,000 randomly generated  $r^2$  values. Modeling was performed in R 3.4.0 and linear regressions in JMP 9.0.

## RESULTS

All 7 complexity metrics were variable across individuals (Table 1), but none showed significant positive correlations with any other (i.e. birds with large repertoires did not tend to have high rates of song delivery, switch song types more often, sing longer songs, etc.) (Table 2). Out of 21 relationships between complexity variables we found 9 positive correlations and 12 negative

correlations (Table 2). The only significant regression analysis ( $P = 0.0039$ ) reflected a negative correlation between song length and song delivery rate: birds that sang longer songs delivered them more slowly (Table 2). Because individual complexity maximums may be as informative as averages for some variables, we also identified the maximum song length, bandwidth and inflections per s values for each individual, and ran similar analyses with those values instead of the trait means. That analysis yielded 7 positive correlations, 14 negative correlations and no significant relationships among complexity metrics (Table 3).

Our randomized data model yielded 1,000 correlations with a mean average  $r^2$  of 0.092. Our observed mean  $r^2$  was 0.086 ( $n = 21$ , Table 2), a value that falls below the center of the random distribution (Figure 2). Several lines of evidence suggest that strong correlations could be observed, given our range of data for each variable. The maximum mean  $r^2$  in our randomization model was 0.24, and the maximum possible mean  $r^2$  when all variables were sorted in parallel across individuals was 0.86. Additionally, we assessed correlations between other variables measured from the same songs and found strong relationships. For example, song bandwidth correlated positively with song high frequency ( $r^2 = 0.81$ ,  $P < 0.0001$ ), and although song length did not correlate with the number of frequency inflections per s (Table 2) it did correlate positively with the number of frequency inflections per song ( $r^2 = 0.53$ ,  $P = 0.0073$ ).

## DISCUSSION

Results indicate that components of Rock Wren song “complexity” vary independently. We found no evidence that one individual bird might display a suite of traits that were all measurably more complex than the traits of his neighbors. Birds with large song repertoires did not tend to sing longer songs or songs with more frequency inflections. Some of these results are rather non-intuitive, for example, one might posit that a bird with a large song

**TABLE 1.** Rock Wren song complexity metrics.

Metric	Mean $\pm$ SD	Range
Repertoire (total number of song types)	71 $\pm$ 19.40	44–108
Delivery (songs min <sup>-1</sup> )	9.6 $\pm$ 1.25	7.07–12.03
Switching (proportion of transitions with a song type switch)	0.89 $\pm$ 0.093	0.69–0.99
Variety (number of song types per 10 songs)	4.26 $\pm$ 0.58	3.40–5.40
Length (s)	1.43 $\pm$ 0.16	1.25–1.75
Bandwidth (Hz)	2891 $\pm$ 250.90	2458–3251
Inflections s <sup>-1</sup> (frequency inflections s <sup>-1</sup> )	12.54 $\pm$ 1.19	10.86–14.43



**TABLE 2.** Results of linear regressions describing correlations between complexity measures of the songs of 12 Rock Wrens.<sup>a</sup>

	Repertoire	Delivery	Switching	Variety	Length	Bandwidth	Inflections s <sup>-1</sup>
Repertoire		$-r^2 < 0.001$	$+r^2 = 0.043$	$-r^2 = 0.043$	$-r^2 = 0.014$	$+r^2 = 0.20$	$+r^2 = 0.05$
Delivery	$P = 0.99$		$+r^2 < 0.001$	$-r^2 = 0.034$	$-r^2 = 0.58 *$	$+r^2 = 0.052$	$-r^2 < 0.001$
Switching	$P = 0.52$	$P = 0.93$		$+r^2 = 0.30$	$-r^2 = 0.0048$	$+r^2 < 0.001$	$+r^2 = 0.16$
Variety	$P = 0.52$	$P = 0.56$	$P = 0.066$		$-r^2 = 0.0027$	$-r^2 = 0.11$	$+r^2 = 0.16$
Length	$P = 0.71$	$P = 0.0039 *$	$P = 0.83$	$P = 0.87$		$-r^2 = 0.049$	$-r^2 = 0.020$
Bandwidth	$P = 0.15$	$P = 0.48$	$P = 0.97$	$P = 0.29$	$P = 0.49$		$-r^2 < 0.001$
Inflections s <sup>-1</sup>	$P = 0.49$	$P = 0.97$	$P = 0.20$	$P = 0.20$	$P = 0.66$	$P = 0.95$	

<sup>a</sup> Correlation direction and strength (– and +) are indicated in the upper right. Significance of the linear regression is indicated in the lower left. Asterisk (\*) indicates significance at  $P < 0.05$ .

**TABLE 3.** Results of linear regressions describing correlations between complexity measures of the songs of 12 Rock Wrens using maximum duration, bandwidth and frequency inflections s<sup>-1</sup>.<sup>a</sup>

	Repertoire	Delivery	Switching	Variety	Maximum length	Maximum bandwidth	Maximum inflections s <sup>-1</sup>
Repertoire		$-r^2 < 0.001$	$+r^2 = 0.043$	$-r^2 = 0.043$	$-r^2 = 0.14$	$-r^2 = 0.0006$	$-r^2 = 0.17$
Delivery	$P = 0.99$		$+r^2 < 0.001$	$-r^2 = 0.034$	$-r^2 = 0.0084$	$-r^2 = 0.074$	$-r^2 = 0.063$
Switching	$P = 0.52$	$P = 0.93$		$+r^2 = 0.30$	$+r^2 = 0.065$	$-r^2 = 0.0058$	$-r^2 = 0.054$
Variety	$P = 0.52$	$P = 0.56$	$P = 0.067$		$+r^2 = 0.18$	$+r^2 = 0.0022$	$-r^2 = 0.038$
Maximum length	$P = 0.24$	$P = 0.78$	$P = 0.43$	$P = 0.17$		$-r^2 = 0.034$	$-r^2 = 0.056$
Maximum bandwidth	$P = 0.94$	$P = 0.39$	$P = 0.81$	$P = 0.89$	$P = 0.56$		$+r^2 = 0.012$
Maximum inflections s <sup>-1</sup>	$P = 0.18$	$P = 0.43$	$P = 0.47$	$P = 0.55$	$P = 0.46$	$P = 0.27$	

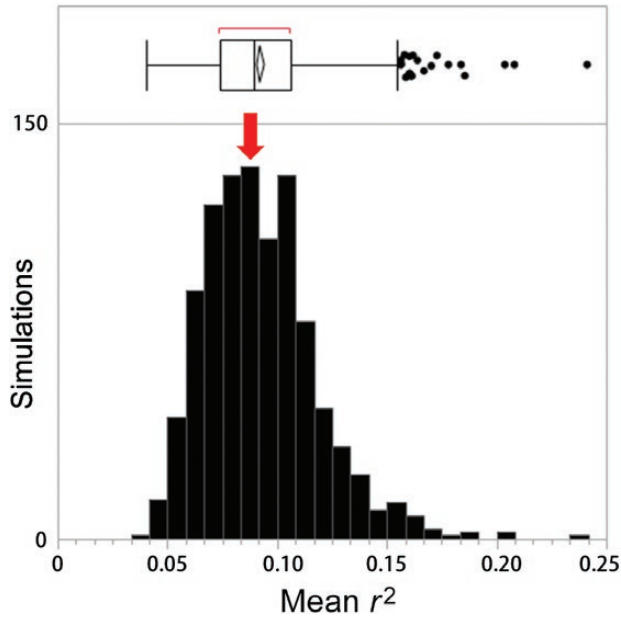
<sup>a</sup> Correlation direction and strength (– and +) are indicated in the upper right. Significance of the linear regression is indicated in the lower left. Asterisk (\*) indicates significance at  $P < 0.05$ .

repertoire should display a higher song switching rate and a higher local song variety, but those predictions were not supported. In fact, the only significant relationship between complexity metrics was a negative one between delivery rate and song length. This adds some support (albeit minimal) to the notion that tradeoffs in complexity may be more common than positive correlations (Gil and Gahr 2002, Cardoso and Hu 2011).

Our study included only 12 birds, but provided a comprehensive measurement of average song behavior in those individuals. As such, the dataset used here differs from those used in some studies of signal complexity, which may rely on only a few sampled vocalizations from each individual or species under investigation (Buskirk 1997, Cardoso and Hu 2011, Kaluthota et al. 2016). By sampling many songs per individual we ensured that our measurements of complexity represent true individual average behavior, and the lack of correlations between those averages is not due to short-term variation in song form and context. Instead, results indicate that the different metrics of song complexity used in this study do not correlate tightly across individuals. It is possible that the lack of significant correlations stems from a relatively low sample size. If the traits measured in this study are weakly correlated, we may not have detected that with our linear regressions. However, 2 sets of results argue against this interpretation. First, the presence of strong correlations between structurally related variables (e.g., high frequency

and bandwidth) confirms that our approach could detect such relationships within the dataset. Second, the average measured  $r^2$  value for any 2 complexity metrics was smaller than the average randomly generated  $r^2$  value for those same metrics in our model, suggesting that Rock Wren song complexity traits are less correlated than we would expect by chance. Weak, non-significant, correlations would be predicted to fall above the average value within the model. Instead, our results suggest that over prolonged singing periods, which display the full repertoire and range of broadcast singing behavior, Rock Wren song “complexity” measures vary independently. Patterns may be different in short-term contexts, such as one-on-one contest or mating situations, but our results are likely to be relevant to any research that examines the complexity of general broadcast signals.

Existing studies have found mixed support for the idea that aspects of bird song repertoires, delivery patterns and form are biologically linked. For example, multiple syntactical or structural song features correlate in some species, but those correlations are often between variables that are non-independent, such as multiple frequency measures or duration and frequency inflections per song (Cardoso and Mota 2007, Soma and Garamszegi 2011, Woodgate et al. 2012, Kaluthota et al. 2016). Other work has shown that different elements of song form and complexity may be independent or serve different functions, making them likely to be under different selective



**FIGURE 2.** Distribution of mean  $r^2$  values from 1,000 simulated correlations between 7 song complexity measures using random permutations of the observed data. The boxplot indicates median, quartiles and 97.5% distributions. The diamond indicates the distribution mean and the red arrow shows where the observed data fall.

pressures (Botero and Vehrencamp 2007, Cardoso and Hu 2011, Wilkins et al. 2015, dos Santos et al. 2018). Our results add to that evidence and extend it to include multiple song features that have not often been previously compared. Many existing studies use either song repertoire size or single-song structural traits to measure song complexity, but for most species we do not know whether those repertoire-based and song-based metrics are correlated. If traits do correlate, they might represent multiple ornaments with redundant or distinct signal content (Moller and Pomiankowski 1993, Gil and Gahr 2002). If traits do not correlate, as is the case for Rock Wren song, this may indicate no need for multiple ornaments, such that elaboration of a single feature provides sufficient information. Among Rock Wrens, repertoire sizes are remarkably large, making this trait an obvious candidate target for selection, and future work will investigate the factors promoting large repertoires. The results presented here suggest that the same factors will not promote the elaboration of other Rock Wren song features, including structural complexity or short-term complexity. A similar pattern was found among House Wrens (*Troglodytes aedon chilensis*), where the majority of measured song complexity traits were not collinear (although, intriguingly, some were) (Dos Santos et al 2018).

Overall, the existing evidence does not support the idea that song “complexity” traits evolve in tandem (Singh and

Price 2015, Collins et al. 2009, Najar and Benedict 2018). Because the term “complexity” is used to refer to many features of bird song, the complete body of literature on this topic includes a diverse array of metrics which are probably not analogous across or even within species. When different features are elaborated in different species, researchers face the challenge of how to draw inferences about similarities between them. Future research should choose metrics of signal complexity carefully, should define those metrics clearly, and should consider when different metrics may or may not be valid and/or comparable (Gil and Gahr 2002, Kershenbaum et al. 2017). When choosing metrics, researchers may benefit from considering the song features of related species and looking for traits that fall outside the typical range of variation within the clade.

The term “complexity” strongly suggests an expected outcome—nearly all studies that measure song complexity using the metrics analyzed in this paper predict it to increase with their treatment (Appendix). This suggests that when we use the term “complexity” to describe the parts of a signal, we assume selection for those traits and may ignore other possible explanations for why song might change, including random or non-functional variation (Podos et al. 2004). Investigators studying signal complexity within and across groups should always verify that the “complexity” in question is meaningful to the animals themselves if it is to be subject to selection.

We expect our conclusions to apply to many signaling systems beyond bird song. Researchers studying other animal acoustic communication including human language complexity have grappled with similar issues (Joseph and Newmeyer 2012, Taylor and Cody 2015). Linguists recognize that language complexity may be measured in different ways relating to phonology, morphology, syntax or other language traits, and have argued that no single metric of overall complexity exists, or indeed that “complexity” as a concept is a meaningful way to assess languages (Deutscher 2009). As demonstrated here, for a single animal communication signal, 7 different complexity variables can show 7 different expression levels across individuals, making it important to differentiate the traits that have communicative significance from those which are non-functional. Researchers studying signal complexity should be explicit in defining metrics, and should be cautious when drawing general conclusions from multiple studies with different methodological approaches.

## ACKNOWLEDGMENTS

Field access was provided by Colorado Parks and Wildlife, the Larimer County Department of Natural Resources, and the City of Fort Collins. N. Warning conducted field work. W. Hoffman assisted with statistical analyses. K. Hoke and 4

anonymous reviewers provided comments on a draft of the manuscript.

**Funding statement:** Support came from The University of Northern Colorado's Provost Fund. The funder did not have any influence on the content of the submitted or published manuscript. The funder did not require approval of the final manuscript to be published.

**Ethics statement:** Research was conducted under University of Northern Colorado IACUC permit 1506C, Colorado permit TRb2041 and USGS/BBL permit 23741. All work followed the ethical guidelines of those agencies.

**Author contributions:** L.B. collected acoustic data and conducted analyses. L.B. and N.A.N. conceived the idea and wrote the manuscript.

**Date deposits:** Supporting data can be found at <http://datadryad.org/resource/doi:10.5061/dryad.453v0/1>

## LITERATURE CITED

- Alward, B. A., F. A. Madison, W. T. Gravley, and G. F. Ball (2016). Antagonism of syringeal androgen receptors reduces the quality of female-preferred male song in canaries. *Animal Behaviour* 119:201–212.
- Barišić, S., D. Čiković, V. Tutiš, J. Kralj, and H. Hoi. (2018). Song structure and variability in the Black-headed Bunting (*Emberiza melanocephala*). *Journal of Ornithology* 159:389–400.
- Barnett, C. A., and J. V. Briskie (2007). Energetic state and the performance of dawn chorus in Silvereyes (*Zosterops lateralis*). *Behavioral Ecology and Sociobiology* 61:579–587.
- Bartsch, C., M. Weiss, and S. Kipper (2015). Multiple song features are related to paternal effort in Common Nightingales. *BMC Evolutionary Biology* 15:115.
- de Boer, R. A., M. Eens, and W. Müller (2016). 'Out of tune': consequences of inbreeding on bird song. *Proceedings of the Royal Society B* 283:20161142.
- Benedict, L., and N. Warning (2017). Rock Wrens preferentially use song types that improve long distance signal transmission during natural singing bouts. *Journal of Avian Biology* 48:1254–1262.
- Boogert, N. J., R. C. Anderson, S. Peters, W. A. Searcy, and S. Nowicki (2011). Song repertoire size in male Song Sparrows correlates with detour reaching, but not with other cognitive measures. *Animal Behaviour* 81:1209–1216.
- Boogert, N. J., L. Giraldeau, and L. Lefebvre (2008). Song complexity correlates with learning ability in Zebra Finch males. *Animal Behaviour* 76:1735–1741.
- Botero C. A., and S. L. Vehrencamp (2007). Responses of male Tropical Mockingbirds (*Mimus gilvus*) to variation in within-song and between-song versatility. *Auk* 124:185–196.
- Botero, C. A., A. E. Mudge, A. M. Koltz, W. M. Hochachka, and S. L. Vehrencamp (2008). How reliable are the methods for estimating repertoire size? *Ethology* 114:1227–1238.
- Bradbury, J. W., and S. L. Vehrencamp (2011). *Principles of Animal Communication*. Sinauer, Sunderland, MA, USA.
- Bradbury, J. W., and S. L. Vehrencamp (2014). Complexity and behavioral ecology. *Behavioral Ecology* 25:435–442.
- Briefer, E., Osiejuk, T. S., Rybak, F., and T. Aubin (2009). Are bird song complexity and song sharing shaped by habitat structure? An information theory and statistical approach. *Journal of Theoretical Biology* 262:151–164.
- Brumm, H., and P. J. B. Slater (2006). Ambient noise, motor fatigue, and serial redundancy in chaffinch song. *Behavioral Ecology and Sociobiology* 60:475–481.
- Brumm, H., S. A. Zollinger, and P. J. Slater (2009). Developmental stress affects song learning but not song complexity and vocal amplitude in Zebra Finches. *Behavioral Ecology and Sociobiology* 63:1387–1395.
- Brunner, P., and G. Pasinelli (2010). Variation in singing style use in the Reed Bunting *Emberiza schoeniclus*: Influencing factors and possible functions. *Journal of Avian Biology* 4:388–379.
- Buskirk, J. V. (1997). Independent evolution of song structure and note structure in American Wood Warblers. *Proceedings of the Royal Society B* 264:755–761.
- Cardoso, G. C., and Y. Hu (2011). Birdsong performance and the evolution of simple (rather than elaborate) sexual signals. *American Naturalist* 178:679–686.
- Cardoso, G. C., and P. G. Mota (2007). Song diversification and complexity in canaries and seedeaters (*Serinus* spp.). *Biological Journal of the Linnean Society* 92:183–194.
- Caro, S. P., K. B. Sewall, K. G. Salvante, and K. W. Sockman (2010). Female Lincoln's Sparrows modulate their behavior in response to variation in male song quality. *Behavioral Ecology* 21:562–569.
- Catchpole, C. K., and P. J. Slater (2003). *Bird Song: Biological Themes and Variations*. Cambridge University Press, Cambridge, UK.
- Collins, S. A., S. R. de Kort, J. Pérez-Tris, and J. L. Tellería (2009). Migration strategy and divergent sexual selection on bird song. *Proceedings of the Royal Society B* 276:585–590.
- Crouch, N. M. A., and R. J. Mason-Gamer (2018). Structural equation modeling as a tool to investigate correlates of extra-pair paternity in birds. *PLoS ONE* 13:e0193365.
- Darwin, C. (1859). *The Origin of Species by Means of Natural Selection: or, the Preservation of Favored Races in the Struggle for Life*. Murray, London, UK.
- Darolová, A., J. Krištofík, H. Hoi, and M. Wink (2012). Song complexity in male Marsh Warblers: Does it reflect male quality? *Journal of Ornithology* 153:431–439.
- Deutscher, G. (2009). "Overall complexity": A wild goose chase? In *Language Complexity as an Evolving Variable* (G. Sampson, D. Gil, and P. Trudgill, Editors). Oxford University Press, Oxford, UK, pp. 243–251.
- DuBois, A. L., S. Nowicki, S. Peters, K. D. Revera-Cáceres, and W. A. Searcy (2018). Song is not a reliable signal of general cognitive ability in a songbird. *Animal Behaviour* 137:205–213.
- Farrell, T. M., K. Weaver, Y. An, and S. A. MacDougall-Shackleton. (2011). Song bout length is indicative of spatial learning in European Starlings. *Behavioral Ecology* 23:101–111.
- Footo, J. R., K. L. D. Marini, and H. AL-Ani (2018). Understanding the function of nocturnal song in ovenbirds: Males do not countersing at night. *Journal of Avian Biology* 2018:e01729.
- Forstmeier, W., D. Hasselquist, S. Bensch, and B. Leisler (2006). Does song reflect age and viability? A comparison between two populations of the Great Reed Warbler *Acrocephalus arundinaceus*. *Behavioral Ecology and Sociobiology* 59:634–643.
- Garamszegi, L. Z., M. Zagalska-Neubauer, D. Canal, G. Blázi, M. Laczi, G. Nagy, E. Szöllösi, É. Vaskuti, J. Török, and S. Zsebök



- (2018). MHC-mediated sexual selection on birdsong: Generic polymorphism, particular alleles and acoustic signals. *Molecular Ecology* 27:2620–2633.
- Gil, D., and M. Gahr (2002). The honesty of bird song: Multiple constraints for multiple traits. *Trends in Ecology & Evolution* 17:133–141.
- Gil, D., P. J. B. Slater, and J. A. Graves (2007). Extra-pair paternity and song characteristics in the Willow Warbler *Phylloscopus trochilus*. *Journal of Avian Biology* 38:291–297.
- Gomes, A. C. R., C. Funghi, M. Soma, M. D. Sorenson, and G. C. Cardoso (2017). Multimodal signaling in Estrildid Finches: Song, dance and colour are associated with different ecological and life-history traits. *Journal of Evolutionary Biology* 30:1336–1346.
- Goretskaia, M. I., I. R. Beme, D. V. Popova, N. Amos, K. L. Buchanan, P. Sunnucks, and A. Pavlova (2018). Song parameters of the Fuscous Honeyeater *Lichenostomus fuscus* correlate with habitat characteristics in fragmented landscapes. *Journal of Avian Biology* 2018:e01493.
- Grunst, M. L., A. S. Grunst, V. A. Formica, R. A. Gonser, and E. M. Tuttle (2018) Multiple signaling functions of song in a polymorphic species with alternative reproductive strategies. *Ecology and Evolution* 8:1369–1383.
- Grunst, M. L., A. S. Grunst, and J. T. Rotenberry (2014). Phenotypic plasticity in nest departure calls: Weighing costs and benefits. *Animal Behaviour* 90:31–39.
- Hamao, S. (2013). Acoustic structure of songs in island populations of the Japanese Bush Warbler, *Cettia diphone*, in relation to sexual selection. *Journal of Ethology* 31:9–15.
- Hart, P. J., E. Sebastián-González, A. Tanimoto, A. Thompson, T. Speetjens, M. Hopkins, and M. Atencio-Picado (2018). Birdsong characteristics are related to fragment size in a Neotropical forest. *Animal Behaviour* 137:45–52.
- Hebets, E. A., A. B. Barron, C. N. Balakrishnan, M. E. Hauber, P. H. Mason, and K. L. Hoke (2016). A systems approach to animal communication. *Proceedings of the Royal Society B* 283:20152889.
- Hebets, E. A., and D. R. Papaj (2005). Complex signal function: Developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology* 57:197–214.
- Hedley, R. W., D. M. Logue, L. Benedict, and D. J. Mennill (2018). Assessing the similarity of song-type transitions among birds: Evidence for interspecies variation. *Animal Behaviour* 140:161–170.
- Herron, M. D., and R. E. Michod (2008). Evolution of complexity in the volvocine algae: Transitions in individuality through Darwin's eye. *Evolution* 62:436–451.
- Hesler, N., R. Mundry, and T. Dabelsteen (2011). Does song repertoire size in Common Blackbirds play a role in an intra-sexual context? *Journal of Ornithology* 152:591–601.
- Hill, S. D., D. H. Brunton, M. G. Anderson, and W. Ji (2018a). Fighting talk: Complex song elicits more aggressive responses in a vocally complex songbird. *Ibis* 160:257–268.
- Hill, S. D., M. D. M. Pawley, M. G. Anderson, and W. Ji (2018b). Higher song complexity and intruder pressure at dawn in a vocally complex songbird. *Emu – Austral Ornithology* 118:147–157.
- Hill, S. D., W. Ji, K. A. Parker, C. Amiot, and S. J. Wells (2013). A comparison of vocalizations between mainland tui (*Prosthemadera novaeseelandiae novaeseelandiae*) and Chatham Island tui (*P. n. chathamensis*). *New Zealand Journal of Ecology* 37:214–223.
- Hill, S. D., M. D. M. Pawley, and W. Ji (2017). Local habitat complexity correlates with song complexity in a vocally elaborate honeyeater. *Austral Ecology* 42:590–596.
- Holveck, M., A. C. V. de Castro, R. F. Lachlan, C. ten Cate, and K. Riebel (2008). Accuracy of song syntax learning and singing consistency signal early condition in Zebra Finches. *Behavioral Ecology* 19:1267–1281.
- Joseph, J. E., and F. J. Newmeyer (2012). All languages are equally complex. *Historiographia linguistica* 39:341–368.
- Kershenbaum, A. (2014). Entropy rate as a measure of animal vocal complexity. *Bioacoustics* 23:195–208.
- Kershenbaum, A., É. C. Déaux, B. Habib, B. Mitchell, V. Palacios, H. Root-Gutteridge, and S. Waller (2017). Measuring acoustic complexity in continuously varying signals: How complex is a wolf howl? *Bioacoustics* 27:215–229.
- Kaluthota, C., B. E. Brinkman, E. B. dos Santos, and D. Rendall (2016). Transcontinental latitudinal variation in song performance and complexity in House Wrens (*Troglodytes aedon*). *Proceedings of the Royal Society B* 283:20152765.
- Křištofik, J., A. Darolová, J. Majtan, M. Okuliarová, M. Zeman, and H. Hoi (2014). Do females invest more into eggs when males sing more attractively? Postmating sexual selection strategies in a monogamous Reed Passerine. *Ecology and Evolution* 4:1328–1339.
- Kroodsma, D. E. (1975). Song patterning in the Rock Wren. *The Condor* 77:294–303.
- Laiolo, P., J. R. Obeso, and Y. Roggia (2011). Mimicry as a novel pathway linking biodiversity functions and individual behavioural performances. *Proceedings of the Royal Society B* 278:1072–1081.
- Lattin, C., and G. Ritchison (2009). Intra- and intersexual functions of singing by male Blue Grosbeaks: The role of within-song variation. *The Wilson Journal of Ornithology* 121:714–721.
- Luttrell, S. A., M. E. Gallagher, and B. Lohr (2016). Multiple estimation methods suggest similar repertoire sizes for Gulf Coast and eastern Marsh Wrens with no correlation between repertoire size and migratory distance. *Behaviour* 153:287–311.
- MacDougall-Shackleton, S. A., L. Dindia, A. E. M. Newman, D. A. Potvin, K. A. Stewart, and E. A. MacDougall-Shackleton (2009). Stress, song and survival in sparrows. *Biology Letters* 5:746–748.
- MacDougall-Shackleton, E. A., K. A. Stewart, D. A. Potvin, and E. Tennenhouse (2009). The rich get richer: Song complexity predicts song element sharing and song output in Song Sparrows *Melospiza melodia*. *Animal Behaviour* 78:141–146.
- Markman, S., S. Leitner, C. Catchpole, S. Barnsley, C. T. Müller, D. Pascoe, and K. L. Buchanan (2008). Pollutants increase song complexity and the volume of the brain area HVC in a songbird. *PLoS ONE* 3:e1674.
- McMullen, H., R. Schmidt, and H. P. Kunc (2014). Anthropogenic noise affects vocal interactions. *Behavioural Processes*, 103:125–128.
- Møller, A. P., and A. Pomiankowski (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology* 32:167–176.
- Morinay, J., G. C. Cardoso, C. Doutrelant, and R. Covas (2013). The evolution of birdsong on islands. *Ecology and Evolution* 3:5127–5140.



- Motes-Rodrigo, A., A. Labra, and H. M. Lampe. (2016). Breeding experience and not age modulates the song development of Pied Flycatchers (*Ficedula hypoleuca*). *Ethology* 123:197–204.
- Najar, N., and L. Benedict (2018). The relationship between latitude, migration and the evolution of bird song complexity. *Ibis* 161:1–12.
- Opaev, A. (2016). Relationships between repertoire size and organization of song bouts in the Grey-crowned Warbler (*Seicercus tephrocephalus*). *Journal of Ornithology* 157:949–960.
- Osiejuk, T. S. (2011). The song of the Cinnamon-breasted Bunting, *Emberiza tahapisi*, in the Bamenda Highlands (NW Cameroon). *Journal of Ornithology* 152:651–659.
- Palmero, A. M., J. Espelósín, P. Laiolo, and J. C. Illera (2014). Information theory reveals that individual birds do not alter song complexity when varying song length. *Animal Behaviour* 87:153–163.
- Palmero, A. M., J. C. Illera, and P. Laiolo (2012). Song characterization in the Spectacled Warbler (*Sylvia conspicillata*): A circum-Mediterranean species with a complex song structure. *Bioacoustics* 21:175–191.
- Pang-Ching, J. M., K. L. Paxton, E. H. Paxton, A. A. Pack, and P. J. Hart (2018). The effect of isolation, fragmentation, and population bottlenecks on song structure of a Hawaiian Honeycreeper. *Ecology and Evolution* 8:2076–2087.
- Pavlova, D. Z., R. Pinxten, and M. Eens (2010). Age-related changes of song traits in female European Starlings (*Sturnus vulgaris*). *Animal Biology* 60:43–59.
- Petrusková, T., T. S. Osiejuk, P. Linhart, and A. Petrusek (2008). Structure and complexity of perched and flight songs of the Tree Pipit (*Anthus trivialis*). *Annales Zoologici Fennici* 45:135–148.
- Pilowsky, J. A. and D. R. Rubenstein (2013). Social context and the lack of sexual dimorphism in song in an avian cooperative breeder. *Animal Behaviour* 85:709–714.
- Podos, J., S. K. Huber, and B. Taft (2004). Bird song: The interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics* 35:55–87.
- Podos, J., D. L. Moseley, S. E. Goodwin, J. McClure, B. N. Taft, A. V. Strauss, C. Rega-Brodsky, and D. C. Lahti. (2016). A fine-scale, broadly applicable index of vocal performance: frequency excursion. *Animal Behaviour* 116:203–212.
- Poot, H., A. ter Maat, L. Trost, I. Schwabl, R. F. Jansen, and M. Gahr (2012). Behavioural and physiological effects of population density on domesticated Zebra Finches (*Taeniopygia guttata*) held in aviaries. *Physiology and Behavior* 105:821–828.
- Potviní, D. A., and E. A. MacDougall-Shackleton (2010). Paternal song complexity predicts offspring sex ratios close to fledging, but not hatching, in Song Sparrows. *The Wilson Journal of Ornithology* 122:146–152.
- Read, A. F., and D. M. Weary (1992). The evolution of bird song: Comparative analyses. *Philosophical Transactions of the Royal Society B* 338:165–187.
- Roper, M. M., A. M. T. Harmer, and D. H. Brunton (2018). Developmental changes in song production in free-living male and female New Zealand bellbirds. *Animal Behaviour* 140:57–71.
- Sandoval, L., G. Barrantes, and D. R. Wilson (2018). Conceptual and statistical problems with the use of the Shannon-Weiner entropy index in bioacoustics analyses. *Bioacoustics* (doi: 10.1080/09524622.2018.1443286).
- dos Santos, E. B., P. E. Llambías, and D. Rendall (2018). Male song diversity and its relation to breeding success in Southern House Wrens (*Troglodytes aedon chilensis*). *Journal of Avian Biology* 49:e01606.
- Schwabl, K. L., A. A. Furlonger, J. M. Lapierre, E. A. MacDougall-Shackleton, and S. A. MacDougall-Shackleton (2012). Regulation of the HPA axis is related to song complexity and measures of phenotypic quality in Song Sparrows. *Hormones and Behavior* 61:652–659.
- Schwabl, K. L., S. D. Moore, E. A. MacDougall-Shackleton, and S. A. MacDougall-Shackleton (2013). Early-life stress affects song complexity, song learning and volume of the brain nucleus RA in adult male Song Sparrows. *Animal Behaviour* 86:25–35.
- Schwabl, H., J. Dowling, D. T. Baldassarre, M. Gahr, W. R. Lindsay, and M. S. Webster (2015). Variation in song system anatomy and androgen levels does not correspond to song characteristics in a tropical songbird. *Animal Behaviour* 104:39–50.
- Seneviratne, S. S., I. L. Jones, and E. H. Miller (2009). Vocal repertoires of auklets (Alcidae: Aethiini): Structural organization and categorization. *The Wilson Journal of Ornithology* 121:568–584.
- Sewall, K. B., E. C. Dankoski, and K. W. Sockman (2010). Song environment affects singing effort and vasotocin immunoreactivity in the forebrain of male Lincoln's Sparrows. *Hormones and Behavior* 58:544–553.
- Singh, P., and T. D. Price (2015). Causes of the latitudinal gradient in birdsong complexity assessed from geographical variation within two Himalayan warbler species. *Ibis* 157:511–527.
- Smith, J. M., and D. Harper (2003). *Animal Signals*. Oxford University Press, Oxford, UK.
- Soma, M., M. Hiraiwa-Hasegawa, and K. Okanoya (2009a). Early ontogenetic effects on song quality in the Bengalese Finch (*Lonchura striata* var. *domestica*): Laying order, sibling competition, and song syntax. *Behavioral Ecology and Sociobiology* 63:363–370.
- Soma, M., M. Hiraiwa-Hasegawa, and K. Okanoya (2009b). Song-learning strategies in the Bengalese Finch: Do chicks choose tutors based on song complexity? *Animal Behaviour* 78:1107–1113.
- Soma, M., and L. Z. Garamszegi. (2011). Rethinking birdsong evolution: Meta-analysis of the relationship between song complexity and reproductive success. *Behavioral Ecology* 22:363–371.
- Soma, M., M. Takahashi, T. Hasegawa, and K. Okanoya (2006). Trade-offs and correlations among multiple song features in the Bengalese Finch. *Ornithological Science* 5:77–84.
- Soma, M., M. Takahashi, M. Ikebuchi, H. Yamada, M. Suzuki, T. Hasegawa, and K. Okanoya (2006). Early rearing conditions affect the development of body size and song in Bengalese Finches. *Ethology* 112:1071–1078.
- Spencer, K. A., and S. A. MacDougall-Shackleton (2011). Indicators of development as sexually selected traits: The developmental stress hypothesis in context. *Behavioral Ecology* 22:1–9.
- Supriya, K., T. D. Price, and M. Rowe (2018). Positive correlations between pre- and post-copulatory sexual traits in warblers. *Journal of Avian Biology* 2018:e01694.

- Taylor, C. E., and M. L. Cody (2015). Bird song: A model complex adaptive system. *Artificial Life and Robotics* 20:285–290.
- Templeton, C. N., K. N. Laland, and N. J. Boogert (2014). Does song complexity correlate with problem-solving performance in flocks of Zebra Finches? *Animal Behaviour* 92:63–71.
- Tobias, J. A., and N. Seddon (2009). Sexual selection and ecological generalism are correlated in antbirds. *Journal of Evolutionary Biology* 22:623–636.
- Weir, J. T., and D. Wheatcroft (2011). A latitudinal gradient in rates of evolution of avian syllable diversity and song length. *Proceedings of the Royal Society B* 278:1713–1720.
- Weiss, M., H. Hultsch, I. Adam, C. Scharff, and S. Kipper (2014). The use of network analysis to study complex animal communication systems: A study on nightingale song. *Proceedings of the Royal Society B* 281:20140460.
- Wilkins, M. R., D. Shizuka, M. B. Joseph, J. K. Hubbard, and R. L. Safran. (2015). Multimodal signalling in the North American Barn Swallow: A phenotype network approach. *Proceedings of the Royal Society B* 282:20151574.
- Woodgate, J., Leitner, S., Catchpole, C. K., Berg, M. L., Bennett, A. T. D., and K. L. Buchanan (2011). Developmental stressors that impair song learning in males do not appear to affect female preferences for song complexity in the Zebra Finch. *Behavioral Ecology* 22:566–573.
- Woodgate, J. L., M. M. Mariette, A. T. Bennett, S. C. Griffith, and K. L. Buchanan (2012). Male song structure predicts reproductive success in a wild Zebra Finch population. *Animal Behaviour* 83:773–781.
- Zann, R., and E. Cash (2008). Developmental stress impairs song complexity but not learning accuracy in non-domesticated Zebra Finches (*Taeniopygia guttata*). *Behavioral Ecology and Sociobiology* 62:391–400.
- Zsebők, S., G. Herczeg, G. Blázi, M. Laczi, G. Nagy, J. Török, and L. Z. Garamszegi (2018). Minimum spanning tree as a new, robust repertoire size comparison method: Simulation and test on birdsong. *Behavioral Ecology and Sociobiology* 72:48.
- Petrusková et al. (2008).
- Potviní and MacDougall-Shackleton (2010).
- dos Santos et al. (2018).
- Schmidt et al. (2012).
- Schmidt et al. (2013).
- Soma et al. (2009a).
- Soma et al. (2009b).
- Weiss et al. (2014).
- Zsebők et al. (2018).
2. Delivery
- Opaev (2016).
- dos Santos et al. (2018).
- Pilowsky and Rubenstein (2013).
3. Switching
- Briefer et al. (2009).
- Forstmeier et al. (2006).
- Hill et al. (2018a).
- Hill et al. (2018b).
- Krištofík et al. (2014).
- Opaev (2016).
- Petrusková et al. (2008).
- Weiss et al. (2014).
- Zann and Cash (2008).
4. Variety (song or syllable/element, depending on the study species)
- Alward et al. (2016).
- Barišić et al. (2018).
- de Boer et al. (2016).
- Boogert et al. (2008).
- Briefer et al. (2009).
- Brumm et al. (2009).
- Cardoso and Mota (2007).
- Caro et al. (2010).
- Collins et al. (2009).
- Crouch and Mason-Gamer (2018).
- Darolová et al. (2012).
- Garamszegi et al. (2018).
- Gomes et al. (2017).
- Grunst et al. (2014).
- Hart et al. (2018).
- Hill et al. (2018a).
- Hill et al. (2018b).
- Holveck et al. (2008).
- Laiolo et al. (2011).
- McMullen et al. (2014).
- Moller and Pomiankowski (1993).
- Motes-Rodrigo et al. (2016).
- Opaev (2016).
- Palmero et al. (2012).
- Pavlova et al. (2010).
- Pilowsky and Rubenstein (2013).
- Poot et al. (2012).

## APPENDIX

Studies ( $n=76$ ) citing the included metrics as bird song “complexity” since 2006. Some references cite multiple metrics and so are repeated.

### 1. Repertoire

- Bartsch et al. (2015).
- de Boer et al. (2016).
- Boogert et al. (2008).
- Boogert et al. (2011).
- Cardoso and Mota (2007).
- DuBois et al. (2018).
- Gil et al. (2007).
- Grunst et al. (2018).
- Krištofík et al. (2014).
- MacDougall-Shackleton et al. (2009).
- Markman et al. (2008).
- Motes-Rodrigo et al. (2016).
- Opaev (2016).

Sandoval et al. (2018).  
 dos Santos et al. (2018).  
 Schwabl et al. (2015).  
 Sewall et al. (2010).  
 Supriya et al. (2018).  
 Tobias and Seddon (2009).  
 Weir and Wheatcroft (2011).  
 Zann and Cash (2008).

#### 5. Duration

Boogert et al. (2008).  
 Brumm and Slater (2006).  
 Brumm et al. (2009).  
 Cardoso and Mota (2007).  
 Crouch and Mason-Gamer (2018).  
 Farrell et al. (2011).  
 Gil et al. (2007).  
 Gomes et al. (2017).  
 Hamao (2013).  
 Hill et al. (2013).  
 Hill et al. (2017).  
 Hill et al. (2018a).  
 Hill et al. (2018b).  
 Kaluthota et al. (2016).  
 Laiolo et al. (2011).  
 Pang-Ching et al. (2018).  
 Pilowsky and Rubenstein (2013).  
 Roper et al. (2018).  
 Soma et al. (2006).  
 Supriya et al. (2018).  
 Weir and Wheatcroft (2011).  
 Woodgate et al. (2011).  
 Zann and Cash (2008).

#### 6. Bandwidth

Cardoso and Mota (2007).  
 Crouch and Mason-Gamer (2018).  
 Gomes et al. (2017).  
 Hamao (2013).  
 Hill et al. (2017).  
 Kaluthota et al. (2016).  
 Laiolo et al. (2011).  
 Morinay et al. (2013).  
 Zann and Cash (2008).

#### 7. Inflections s<sup>-1</sup>

Cardoso and Mota (2007).  
 Crouch and Mason-Gamer (2018).  
 Hamao (2013).  
 Pang-Ching et al. (2018).

### UNDEFINED COMPLEXITY MEASURES

Some papers measure song metrics and use the term “complexity” but do not commit to calling any one metric “complexity.” This practice further illustrates the varied and imprecise use of that term within the bird song literature. The following list is not exhaustive but rather illustrative.

Barnett and Briskie (2007).  
 Brunner and Pasinelli (2010).  
 Foote et al. (2018).  
 Goretskaia et al. (2018).  
 Hesler et al. (2011).  
 Lattin and Ritchison (2009).  
 MacDougall-Shackleton et al. (2009).  
 Osiejuk (2011).  
 Seneviratne et al. (2009).