Animal Behaviour xxx (2016) e1–e4

Contents lists available at ScienceDirect

Animal Behaviour

journal homepage: www.elsevier.com/locate/anbehav

Forum

Advancing the inference of performance in birdsong

Gonçalo C. Cardoso [a, b, *]

[a] CIBIO — Research Centre in Biodiversity and Genetic Resources, University of Porto, Campus Agrário de Vairão, Vairão, Portugal
[b] Behavioural Ecology Group, Department of Biology, University of Copenhagen, Copenhagen, Denmark

ARTICLE INFO

Article history:
Received 18 July 2016
Initial acceptance 21 September 2016
Final acceptance 4 October 2016
Available online xxx
MS. number: AF-16-00638R

Keywords:
bioacoustics
birdsong
communication
methods
performance

It is appealing to integrate different acoustic traits to infer differences in performance demands among birdsongs, and to use this as a tool for investigating which roles song performance plays in communication. But inferring performance from acoustic measurements introduces a degree of interpretation that can cause disagreement. Here I give an overview of approaches to assess song performance, associated methodological issues, and ways of addressing them. I note advantages and limitations of performance metrics derived from physiological principles or from acoustic trade-offs, discuss issues with the scaling of performance metrics, and with choosing and adapting metrics to different study species and research goals. Throughout I emphasize that these metrics provide tentative assessments of performance, and that empirical results should be interpreted by comparison to alternative hypotheses.

Birdsong is one of the most diverse sexual signals in nature. Many song traits have been correlated with aspects of individual quality, mating success or motivation, and several of those song traits have associated costs or pose performance challenges (Gil & Gahr, 2002). The diversity of song traits, and of ways of combining them, can make it difficult to decide how to quantify birdsong in a way meaningful for communication. In this regard, assessing song performance is appealing because it offers a way of integrating information from many acoustic traits to derive hypotheses on how birds may best evaluate the quality of songs, and doing so in a way that can be customized to species differing in song and in the singing constraints they are subject to (Podos, Huber, & Taft, 2004).

However, assessing song performance often introduces a layer of interpretation and data transformation in between the acoustic measurements used in research and the conclusions taken, which can give rise to disagreement. In this issue, Kroodsma (2017) criticizes the state of the art, raising issues with the empirical evidence and with the methodological rationale of work using the metric ‘vocal deviation’. Here I will focus on methodological issues, both those noted by Kroodsma and additional ones, and give an overview of problems and solutions when assessing performance either using ‘vocal deviation’ (a metric based on the compromise between frequency modulation and rate of syllable repetitions; Podos, 1997, 2001) or other metrics.

SONG PERFORMANCE IS INFERRED AND ASSESSED, NOT MEASURED

Song performance refers to the degree of challenge to the motor system, the respiratory system or other physiological processes involved in singing. Measuring performance would require very precise knowledge on the physiology of singing, such as to translate acoustic differences into quantitative physiological demands. Current knowledge on song production mechanisms is insufficient to do this for complex birdsongs, but suffices to make informed inferences on the directions in which changing song traits may be more or less demanding. For example, longer continuous singing may pose ventilation challenges (Suthers & Zollinger, 2008), louder songs require building larger airsac pressures (Goller & Cooper, 2008), wider frequency modulation should require more movement of the vocal tract (Goller & Cooper, 2008; Suthers & Zollinger, 2008), and two-voiced sounds or precise repetition of syllables are hurdles of neuromotor coordination (Sakata & Vehrencamp, 2012; Suthers & Zollinger, 2008).

But quantifying such acoustic traits, or using a metric that combines more than one trait, does not measure performance. It tentatively assesses (i.e. places a quantitative value on)

* Correspondence: G. C. Cardoso, CIBIO — Research Centre in Biodiversity and Genetic Resources, University of Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal.
E-mail address: gcardoso@cibio.up.pt.

http://dx.doi.org/10.1016/j.anbehav.2016.11.034
0003-3472/© 2016 Published by Elsevier Ltd on behalf of The Association for the Study of Animal Behaviour.

performance, based on the inference that acoustic changes in a certain direction should be more physiologically demanding, at least along part of the quantitative scale. Assessing differs from measuring, in that it places quantitative values indirectly, based on indexes, classifications, judgement, etc. Awareness of the tentative nature of performance metrics is important because it prompts researchers to use them critically, for example by testing the adequacy of a metric to their species, adapting and refining them if possible, and considering alternative interpretations to empirical findings. Most of the proposed solutions to the problems below follow from viewing performance metrics as tentative tools that must be validated by whether they improve empirical insight on communication.

PROBLEMS AND RECOMMENDATIONS

Single Acoustic Traits versus Composite Metrics

Aspects of song performance can be reflected by simple acoustic traits (e.g. song length, sound amplitude, etc.), but many such traits trade-off with each other within songs (e.g. Cardoso & Mota, 2009; Nemeth et al., 2013; Podos, 1997). This means that a high-performance song can either exaggerate a single acoustic trait or use a demanding combination of traits. We therefore face the choice of assessing performance with measurements of a single acoustic trait versus metrics that combine several acoustic traits. The former has the advantages of simplicity and ease of interpretation, but the potential disadvantage of capturing only a small part of performance differences.

Composite metrics have the advantage of improved comprehensiveness, but may have a less straightforward interpretation. This problem is substantially ameliorated using post hoc tests asking whether the results obtained for composite metrics of performance are indeed best interpreted as due to a synergistic effect of different acoustic traits (i.e. results are clearer compared to analyses of individual acoustic traits alone), or can be equally understood with one of those acoustic traits or with simpler metrics. Such tests can ground the interpretation of results, but have rarely been used (Cardoso & Atwell, 2016; Cardoso, Atwell, Ketterson, & Price, 2009; Geberzahn & Aubin, 2014a,b; Podos et al., 2016).

For example, skylarks, *Alauda arvensis*, reacting to a song playback use higher-performance singing, as assessed by 'sound density' (a metric of respiratory performance based on singing long syllables with short intervals; Geberzahn & Aubin, 2014a) and by 'vocal gap deviation' (a metric of motor performance based on the speed of frequency changes during intervals within song; Geberzahn & Aubin, 2014b). These composite metrics appear better tuned to capture performance differences relevant for communication, compared to the individual acoustic traits used by them, because song differences when reacting to playbacks are not detected when analysing separately those individual acoustic traits (Geberzahn & Aubin, 2014a,b).

Finding the Direction of Performance Metrics: Physiological Principles

Metrics of performance derived a priori from physiological principles have the advantage of generality, because many sound production mechanisms are shared among bird species. Examples of this approach are 'song consistency', which quantifies the ability to render accurate repetitions of a song or syllable (Sakata & Vehrencamp, 2012), 'sound density' and similar metrics (e.g. 'percentage peak performance'; Forstmeier, Kempanaers, Meyer, & Leisler, 2002) that assess the relation between the length of sound and intervals within song, or 'frequency excursion', which assesses the rate of frequency modulation (Podos et al., 2016). Many measurements of single acoustic traits with implications for performance also fall in this category.

But, even if some physiological principles are universal, bird species differ so much in song that the challenges they experience should also differ (Podos et al., 2004). Depending on which song traits are typical of a species, some physiological constraints will be limiting, while others are less relevant because they are not approached. For example, high sound amplitude requires building higher airsac pressures (Goller & Cooper, 2008) that are likely more challenging to achieve instantaneously than gradually, and should thus be limited by the brevity of sounds. Accordingly, short syllables or syllables split into multiple elements reach lower amplitudes in the songs of several *Serinus* finches and related species (Cardoso & Mota, 2009). But the trade-off between sound amplitude and number of within-syllable elements weakens and disappears for species with simpler syllables (Cardoso & Mota, 2009), meaning that comparing song performance based on this compromise between sound amplitude and syllable complexity may only be relevant for some species.

Therefore, using metrics of performance based on general physiological principles is no guarantee that those metrics are relevant for a particular study species. This problem is reduced when choosing metrics guided by evidence that a particular aspect of vocal performance should be limiting for a focal species. For example, high sound frequency and fast note rate in great tits, *Parus major*, are associated with occasional disruptions in singing, suggesting that those song traits place performance challenges (Lambrechts, 1997), or swamp sparrows, *Melospiza georgiana*, tutored with trills with increased syllable repetition rates produce variants suggestive that, all else being equal, trill rate limits their performance (Podos, 1996). Less direct guidance for likely relevant aspects of performance may be that a species has an exaggerated song trait. For example, we have looked for communication functions of sound frequency and syllable rate in serins (*Serinus serinus*; Cardoso, Mota, & Depraz, 2007; Funghi, Cardoso, & Mota, 2015) guided by this species having the highest sound frequency and shortest syllable intervals among related finches (Cardoso, Hu, & Mota, 2012; Cardoso & Mota, 2007). Using a metric without evidence that it addresses a limiting aspect of performance for the study species can be attempted, and may be validated by finding that animals use or respond to variation in that aspect of song performance. But negative results are ambiguous to interpret in this circumstance, as they can be due to inadequacy of the metric chosen.

Finding the Direction of Performance Metrics: Acoustic Trade-offs

This type of metric is ambitious in that acoustic trade-offs are used not only to select likely limiting aspects of performance, but also to compute the performance metric itself. Examples of these metrics are 'vocal deviation' (Podos, 2001), 'vocal gap deviation' (Geberzahn & Aubin, 2014b), and metrics based on multidimensional trade-offs of acoustic traits with an aspect of song output (e.g. with the brevity of intervals or sound amplitude; Cardoso, Atwell, Ketterson, & Price, 2007; Cardoso et al., 2009). The rationale is to infer performance trade-offs from the co-distribution of acoustic traits across a large sample of songs, and then assess how individual songs are positioned in relation to these trade-offs. This type of metric can be made comprehensive by integrating effects of many acoustic traits in a multidimensional analysis. Comprehensiveness, however, can be advantageous or not depending on the comparisons intended (see below).

Out of many potential performance limits, this approach focuses on the trade-offs that actually affect singing in a focal species or group, providing metrics tailored for their apparent performance challenges. For example, sound amplitude can be limited by sound frequency, either because producing some frequencies requires different air pressures (e.g. Amador, Goller, & Mindlin, 2009; Beckers, Suthers, & ten Cate, 2003) or because the ability to produce loud sound decreases towards the limits of a species' potential frequency range (Lambrechts, 1996). In the latter case, it is plausible that species using different sound frequencies, relative to their potential frequency range, experience different performance challenges. Looking again at *Serinus* and related finches, sound frequencies differ markedly among species and are not related to differences in body size (Cardoso & Mota, 2007), suggesting that these species use different regions of their frequency range. In some species sound amplitude is not linearly related to the sound frequency of syllables, in other species amplitude increases with frequency, and yet in others amplitude decreases with frequency (Cardoso & Mota, 2009; see also Figure A1 of Cardoso & Atwell, 2012). Therefore, knowing which acoustic trade-offs affect a species may not only adjust, but even reverse, the direction of a performance metric in cases where either extreme of an acoustic trait could be physiologically demanding.

One problem associated with this approach is that acoustic trade-offs are not proof of performance challenges. Alternative interpretations may apply even to acoustic trade-offs that are consistent with a physiological compromise. For example, the absence of fast trills with wide frequency modulation, amidst hundreds of trills from several related species (Podos, 1997), is convincing evidence for a motor limitation on the speed of frequency modulation. But this evidence is often weaker for individual species, especially if using small sample sizes (Wilson, Bitton, Podos, & Mennill, 2014), where it may be interpreted as frequency modulation accumulating randomly along the length of syllables rather than limited by the speed of modulation (Cardoso, Atwell et al., 2007). Similarly, other correlations among acoustic traits might be due to functional or developmental reasons, rather than to performance challenges. The possibility of alternative interpretations for trade-offs is not impeditive of deriving tentative metrics of performance, but calls for critical interpretation of these metrics (e.g. checking whether they explain empirical results better than simpler acoustic traits do).

Finding the Scale of Performance Metrics

Two difficult issues related to scale, whether of acoustic measurements or of composite performance metrics, are: which scale best depicts movement (e.g. configuration of the vocal tract) or other changes (e.g. air pressure or expenditure) during singing? And, along which part of that scale do performance challenges increase the most? For example, changing acoustic traits in a direction of higher performance should be more challenging when near a physiological limit than when further away, and scaling to reflect this asymmetry might render metrics more meaningful. Ideally, scaling would be informed by physiological evidence but, even more here than for finding the direction of performance metrics, this is generally insufficient to calibrate the scale of performance metrics in a study species.

Approximations to biologically meaningful scales can nevertheless be made, informed by principles of sound production. For example, logarithmic transformation of sound frequency should best predict conformational changes of the vocal tract associated with resonant frequencies (Cardoso, 2013). It may also be advantageous to log-transform durations when quantifying the brevity of events (e.g. of syllables or of intervals), because small reductions in the duration of a brief event likely pose a greater motor challenge than the same small reduction in the duration of a long event. Likewise, when quantifying rates (e.g. rate of repetition, rate of frequency or of amplitude modulation), it may be advisable to use an inverse log transformation (since rate = 1/duration). Another type of approximation is to explore nonlinear relations between a metric of performance and how birds respond to or use song, to learn whether a transformation of scale would render the metric more meaningful for communication.

Comprehensiveness versus Generality

Comparing performance among renditions of the same song type is easier than comparing among song types, because the latter differ in many acoustic traits relevant to performance. Comparisons among song types may thus require integrating the effects of many acoustic traits, or using several different metrics. For example, comparisons among dark-eyed juncos, *Junco hyemalis*, singing different song types used several metrics of performance, including some that integrate effects of many acoustic traits (Cardoso et al., 2009; Cardoso, Atwell, Hu, Ketterson, & Price, 2012), and comparisons among swamp sparrows singing different song types used a metric that assesses frequency modulation quite exhaustively throughout the entire length of songs (‘frequency excursion’; Podos et al., 2016). Otherwise, it would not have been convincing to show absence of individual differences in performance. Conversely, studies using the metric ‘vocal deviation’ often avoid comparing among song types (reviewed in Kroodsma, 2017), and work using ‘percentage peak performance’ typically controls statistically for differences among song types (e.g. Forstmeier et al., 2002). This acknowledges implicitly (and sometimes explicitly; Cardoso, 2014; DuBois, Nowicki, & Searcy, 2009; Podos, Lahti, & Moseley, 2009, Podos et al., 2016) that comparisons of performance are only sensible when there is little variation in acoustic traits not accounted for by the metric used. Also for this reason, playback experiments of song performance may benefit from artificially modifying stimuli (despite associated challenges; Kroodsma, 2017) to avoid unwanted variation in relevant acoustic traits.

A counterintuitive exception to the above reasoning (comprehensiveness is needed to compare among song types) is the metric ‘vocal gap deviation’, which assesses motor performance based on the brevity of within-song intervals in relation to the changes in sound frequency during those intervals (Geberzahn & Aubin, 2014b). The gambit here is to forgo evaluating the voiced portions of song to try obtaining a metric that may still be representative of differences in motor performance during singing, and that can be compared among song types (or even among species) unencumbered by their many differences in sound modulation (Cardoso, 2014; Geberzahn & Aubin, 2014b). Other exceptions are aspects of performance likely not affected by fine differences among songs (e.g. song rate or time spent singing; Brumm, Lachlan, Riebel, & Slater, 2009), and that may therefore be compared irrespective of song type.

A disadvantage of performance metrics aiming at comprehensiveness is that they tend to be less applicable for comparisons among species. This is because some acoustic trade-offs that constrain singing should differ across species (Podos et al., 2004), and metrics taking into account the effects of many acoustic traits need to be adjusted accordingly (see above). Among-species comparisons of song performance, which are yet rare in the literature, will likely need to compromise on comprehensiveness to be applicable across different species. Similarly to other uses of performance metrics, their validity for among-species comparisons hinges on whether these metrics uncover evolutionary patterns that are not detected using simpler acoustic traits.

Please cite this article in press as: Cardoso, G. C., Advancing the inference of performance in birdsong, Animal Behaviour (2016), http://dx.doi.org/10.1016/j.anbehav.2016.11.034

CONCLUSION

In this overview of issues to be aware of when planning to infer song performance from acoustic recordings, I echoed some of Kroodsma's (2017) concerns (e.g. scaling metrics, so that acoustic differences scale more evenly with performance demands) and disagreed with others (e.g. interpreting acoustic changes in a certain direction as increased performance). Since performance metrics are tentative assessments rather than true measurements, I suggested that they be used flexibly, to fit the study species and level of analysis (e.g. within song type, among song types or among species), and results interpreted by comparison to simpler alternatives.

Overall, I see the field maturing and benefiting from unanticipated findings. It is true that song repertoires and cultural conformity can make it difficult to reveal differences in performance ability among individuals (Kroodsma, 2017; but see also Ballentine, 2009; Kagawa & Soma, 2013; Sockman, 2009), which has been a central hypothesis in this field. But, having learned this, research is now appearing on how song repertoires may compensate for limitations that cultural conformity places on song performance (Cardoso & Atwell, 2016; Podos et al., 2016), or how birds may adaptively choose not to sing high-performance songs (Logue & Forstmeier, 2008; Poesel & Nelson, 2015). I expect that critical use of performance metrics and reacting to unexpected findings will continue to move the field in novel directions.

Acknowledgments

I thank the editor and two anonymous referees for their corrections and discussion. G.C.C. was supported by a fellowship from the Fundação para a Ciência e a Tecnologia (SFRH/BPD/110165/2015).

References

Amador, A., Goller, F., & Mindlin, G. B. (2009). Frequency modulation during song in a suboscine does not require vocal muscles. *Journal of Neurophysiology, 99*, 2383–2389.

Ballentine, B. (2009). The ability to perform physically challenging songs predicts age and size in male swamp sparrows, *Melospiza georgiana*. *Animal Behaviour, 77*, 973–978.

Beckers, G. J. L., Suthers, R. A., & ten Cate, C. (2003). Mechanisms of frequency and amplitude modulation in ring dove song. *Journal of Experimental Biology, 206*, 1833–1843.

Brumm, H., Lachlan, R. F., Riebel, K., & Slater, P. J. B. (2009). On the function of song type repertoires: Testing the 'antiexhaustion hypothesis' in chaffinches. *Animal Behaviour, 77*, 37–42.

Cardoso, G. C. (2013). Using frequency ratios to study vocal communication. *Animal Behaviour, 85*, 1529–1532.

Cardoso, G. C. (2014). Studying the silent side of birdsong. *BMC Biology, 12*, 62. http://dx.doi.org/10.1186/s12915-014-0062-8.

Cardoso, G. C., & Atwell, J. W. (2012). On amplitude and frequency in birdsong: A reply to Zollinger et al. *Animal Behaviour, 84*(4), e10–e15.

Cardoso, G. C., & Atwell, J. W. (2016). Shared songs are of lower performance in the dark-eyed junco. *Royal Society Open Science, 3*, 160341.

Cardoso, G. C., Atwell, J. W., Hu, Y., Ketterson, E. D., & Price, T. D. (2012). No correlation between three selected trade-offs in birdsong performance and male quality for a species with song repertoires. *Ethology, 118*, 584–593.

Cardoso, G. C., Atwell, J. W., Ketterson, E. D., & Price, T. D. (2007). Inferring performance in the songs of dark-eyed juncos (*Junco hyemalis*). *Behavioral Ecology, 18*, 1051–1057.

Cardoso, G. C., Atwell, J. W., Ketterson, E. D., & Price, T. D. (2009). Song types, song performance, and the use of repertoires in dark-eyed juncos (*Junco hyemalis*). *Behavioral Ecology, 20*, 901–907.

Cardoso, G. C., Hu, Y., & Mota, P. G. (2012). Birdsong, sexual selection, and the flawed taxonomy of canaries, goldfinches and allies. *Animal Behaviour, 84*, 111–119.

Cardoso, G. C., & Mota, P. G. (2007). Song diversification and complexity in canaries and seedeaters (*Serinus* spp.). *Biological Journal of the Linnean Society, 92*, 183–194.

Cardoso, G. C., & Mota, P. G. (2009). Loudness of syllables is related to syntax and phonology in the songs of canaries and seedeaters. *Behaviour, 146*, 1649–1663.

Cardoso, G. C., Mota, P. G., & Depraz, V. (2007). Female and male serins (*Serinus serinus*) respond differently to derived song traits. *Behavioral Ecology and Sociobiology, 61*, 1425–1436.

DuBois, A. L., Nowicki, S., & Searcy, W. A. (2009). Swamp sparrows modulate vocal performance in an aggressive context. *Biology Letters, 5*, 163–165.

Forstmeier, W., Kempanaers, B., Meyer, A., & Leisler, B. (2002). A novel song parameter correlates with extra-pair paternity and reflects male longevity. *Proceedings of the Royal Society B: Biological Sciences, 269*, 1479–1485.

Funghi, C., Cardoso, G. C., & Mota, P. G. (2015). Increased syllable rate during aggressive singing in a bird with complex and fast song. *Journal of Avian Biology, 46*, 283–288.

Geberzahn, N., & Aubin, T. (2014a). How a songbird with a continuous singing style modulates its song when territorially challenged. *Behavioral Ecology and Sociobiology, 68*, 1–12.

Geberzahn, N., & Aubin, T. (2014b). Assessing vocal performance in complex birdsong: A novel approach. *BMC Biology, 12*, 58. http://dx.doi.org/10.1186/s12915-014-0058-4.

Gil, D., & Gahr, M. (2002). The honesty of bird song: Multiple constraints for multiple traits. *Trends in Ecology & Evolution, 17*, 133–141.

Goller, F., & Cooper, B. G. (2008). Peripheral mechanisms of sensorimotor integration during singing. In H. P. Zeigler, & P. Marler (Eds.), *Neuroscience of birdsong* (pp. 99–114). Cambridge, U.K.: Cambridge University Press.

Kagawa, H., & Soma, M. (2013). Song performance and elaboration as potential indicators of male quality in Java sparrows. *Behavioural Processes, 99*, 138–144.

Kroodsma, D. (2017). Birdsong performance studies: A contrary view. *Animal Behaviour, 125*, e1–e16.

Lambrechts, M. M. (1996). Organization of birdsong and constraints on performance. In D. E. Kroodsma, & E. H. Miller (Eds.), *Ecology and evolution of acoustic communication in birds* (pp. 305–320). Ithaca, NY: Cornell University Press.

Lambrechts, M. M. (1997). Song frequency plasticity and composition of phrase versions in great tits *Parus major*. *Ardea, 85*, 99–109.

Logue, D. M., & Forstmeier, W. (2008). Constrained performance in a communication network: Implications for the function of song-type matching and for the evolution of multiple ornaments. *American Naturalist, 172*, 34–41.

Nemeth, E., Pieretti, N., Zollinger, S. A., Geberzahn, N., Partecke, J., Miranda, A. C., et al. (2013). Bird song and anthropogenic noise: Vocal constraints may explain why birds sing higher-frequency songs in cities. *Proceedings of the Royal Society B: Biological Sciences, 280*, 20122798.

Podos, J. (1996). Motor constraints on vocal development in a songbird. *Animal Behaviour, 51*, 1061–1070.

Podos, J. (1997). A performance constraint on the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). *Evolution, 51*, 537–551.

Podos, J. (2001). Correlated evolution of morphology and vocal signal structure in Darwin's finches. *Nature, 409*, 185–188.

Podos, J., Huber, S. K., & Taft, B. (2004). Bird song: The interface of evolution and mechanism. *Annual Review of Ecology, Evolution, and Systematics, 35*, 55–87.

Podos, J., Lahti, D. C., & Moseley, D. L. (2009). Vocal performance and sensorimotor learning in songbirds. *Advances in the Study of Behavior, 40*, 159–195.

Podos, J., Moseley, D. L., Goodwin, S. E., McClure, J., Taft, B. N., Strauss, A. V. H., et al. (2016). A fine-scale, broadly applicable index of vocal performance: Frequency excursion. *Animal Behaviour, 116*, 203–212.

Poesel, A., & Nelson, D. A. (2015). Changes in performance of shared and unshared songs within and between years in the white-crowned sparrow. *Ethology, 121*, 850–860.

Sakata, J. T., & Vehrencamp, S. L. (2012). Integrating perspectives on vocal performance and consistency. *Journal of Experimental Biology, 215*, 201–209.

Sockman, K. W. (2009). Annual variation in vocal performance and its relationship with bill morphology in Lincoln's sparrows, *Melospiza lincolnii*. *Animal Behaviour, 77*, 663–671.

Suthers, R. A., & Zollinger, S. A. (2008). From brain to song: The vocal organ and vocal tract. In H. P. Zeigler, & P. Marler (Eds.), *Neuroscience of birdsong* (pp. 78–98). Cambridge, U.K.: Cambridge University Press.

Wilson, D. R., Bitton, P. P., Podos, J., & Mennill, D. J. (2014). Uneven sampling and the analysis of vocal performance constraints. *American Naturalist, 183*, 214–228.

Please cite this article in press as: Cardoso, G. C., Advancing the inference of performance in birdsong, Animal Behaviour (2016), http://dx.doi.org/10.1016/j.anbehav.2016.11.034