No Correlation Between Three Selected Trade-Offs in Birdsong Performance and Male Quality for a Species With Song Repertoires

Gonçalo C. Cardoso[*], **Jonathan W. Atwell**[†], **Yang Hu**[‡], **Ellen D. Ketterson**[†], **Trevor D. Price**[§]

[*]CIBIO—Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Vairão, Portugal

[†]Department of Biology, Indiana University Bloomington, IN, USA

[‡]Department of Zoology, University of Melbourne Melbourne, Vic., Australia

[§]Department of Ecology and Evolution, University of Chicago Chicago, IL, USA

Abstract

Demanding performance of vocal signals, such as birdsong, may be evaluated by trade-offs among acoustic traits. If individuals differ in their ability to sustain physiologically demanding singing, then aspects of song performance resulting from such trade-offs could signal individual quality. Song performance can also differ among song types, and it is not known whether this influences the assessment of individual quality. We asked whether three trade-off-based measures of song performance indicate male age or aspects of condition (body condition, hematocrit and ectoparasite load) in the dark-eyed junco (*Junco hyemalis*), a species with small repertoires. Across a sample of over 100 males, no measure of song performance was related to male age or condition, nor did song performance improve with age for those males recorded in consecutive years. In all cases, the variation in song performance explained by these predictors was small (<4%). Instead, the more song types we recorded from a male, the more likely we were to record high-performance songs, and this sampling effect was stronger than putative correlations with male quality. These results complement a previous study on this population showing that most variation in performance is found among song types rather than among males. Collectively, the lack of association between trade-off-based aspects of song performance and male age or condition, plus variation among song types that interferes with rapid assessment of a male's best performance, indicate that these aspects of song performance do not allow a good assessment of male quality in juncos, and perhaps more generally in species with song repertoires.

Introduction

Signals used in competitive interactions or for attracting mates are expected to encode information on aspects of the quality of the signaller (i.e. on traits related to fitness or competitive ability) or on motivation, and many such signals seem to be elaborate or costly

Correspondence Gonçalo C. Cardoso, CIBIO—Centro de Investigação em Biodiversidade e Recursos Genéticos, Universidade do Porto, Campus Agrário de Vairão, 4485-661 Vairão, Portugal. gcardoso@cibio.up.pt.

and thus encode reliable information on individual quality (reviewed in Johnstone 1996; Searcy & Nowicki 2005). Birdsong is an example of a signal that may encode information on the quality of the signaller in a variety of ways (Gil & Gahr 2002). For example, time spent singing is not spent foraging, and thus the amount of singing at strategic times of day when energetic reserves are more critical, such as dawn, may indicate energetic condition (Hutchinson et al. 1993). Similarly, large repertoires may indicate good neuronal development during a critical period early in life when song learning takes place (Nowicki et al. 1998; Nowicki1 et al. 2000; Spencer et al. 2004).

Structural aspects of song may also indicate aspects of quality of the signaller. As birdsong is often a complex signal that varies in many acoustic traits, and those traits trade-off with each other because of physiological constraints (reviewed in Podos et al. 2004a, 2009), song performance may be reflected in a demanding combination of traits rather than the exaggeration of a single trait. Evidence is appearing that this type of song performance can be related to aspects of individual quality such as age or body condition, and also be preferred by females. For example, longer syllables require longer intervals for respiratory recovery than shorter syllables (e.g. Hartley & Suthers 1989), and male dusky warblers (*Phylloscopus fuscatus*) that are able to sustain songs with long, loud syllables and short intervals in between are preferred as extrapair mates (Forstmeier et al. 2002). Also, many species do not sing trills that simultaneously have a large frequency bandwidth and a fast trill rate, two traits that trade-off against each other (Podos 1997). Older and heavier male swamp sparrows (*Melospiza georgiana*) have songs closer to this limit than younger and lighter males (Ballentine 2009), and female swamp sparrows prefer these higher performance songs in the context of mate choice (Ballentine et al. 2004). These findings indicate that aspects of song performance based on trade-offs between song traits can signal male quality.

One issue that has not been addressed is how variation within song repertoires (i.e. among the song types that a bird sings) can influence the efficacy of songs in signalling quality. Because repertoires may be comprised of song types that differ in performance, Logue & Forstmeier (2008) suggested that the assessment of male singing ability could be complicated by this variation. For example, in dark-eyed juncos (*Junco hyemalis*), several aspects of song performance usually vary much more among the song types that are sung by the same male than among different males (Cardoso et al. 2009). This suggests that in species with song repertoires, quality may be more difficult to assess.

In this study, we test whether trade-off-based aspects of song performance are related to age or measures of physiological condition in dark-eyed juncos, and ask how variation within repertoires affects the assessment of male song performance. We have previously documented trade-offs between song traits in junco song (Cardoso et al. 2007) and, using measures of song performance based on those trade-offs as well as more general measures (Podos 2001; Leadbeater et al. 2005; Holveck & Riebel 2007), found that song performance varies widely among different junco song types (Cardoso et al. 2009). Despite large variation in performance among song types, we detected small but significant individual differences in mean song performance among junco males (Cardoso et al. 2009). Here we use these insights and song data to test whether song performance can indicate aspects of

male quality in juncos, a question that applies generally to understand species with song repertoires. The aspects of performance that we analyse have been shown in other species to either be preferred by females (Forstmeier et al. 2002; Ballentine et al. 2004; Holveck & Riebel 2007), be effective in aggressive contexts (Illes et al. 2006; Cramer & Price 2007; de Kort et al. 2009a) or relate to male quality (Ballentine 2009). If song performance reflects male quality in juncos, we predict that older males or males in better condition will have higher performance songs, in terms of either average or maximum performance. Alternatively, if song performance is more strongly a function of differences among song types than individual quality, we predict that the observed maximum performance will be contingent on the quantity of song types heard from each male. Distinguishing among these scenarios is important to frame the potential communicative role of these song traits.

Methods

Study Populations, Age and Measurements

We studied dark-eyed juncos during the breeding seasons of 2006 and 2007, in two populations in South California: on the campus of the University of California at San Diego and in a mountain population 75 km distant. Details of the populations are in Yeh & Price (2004). Captures were made using mist nets and walk-in traps, and birds were marked with unique combinations of coloured leg rings. For the 128 males with song recordings and age assignments, each individual was captured on average 3.4 times (± 1.71 standard deviation, SD, range 1–8; average interval between recaptures in the same seasons: 34 d ± 28 SD), including 33 males captured in both breeding seasons. Before releasing each bird, we determined age, took several measurements, including body mass, size and ectoparasite loads, and took a blood sample to measure hematocrit level.

Birds were assigned to one of two age categories – second-year (i.e. the year after fledging) or after second-year – based on variation in plumage coloration associated with the retention of juvenile pre-basic I plumage feathers in second-year birds, the shape of the tips of rectrices, and wing and tail length (for details on age criteria see Pyle 1997; Nolan et al. 2002). We also used banding history as supplementary information (i.e. birds banded as adults in previous years are unambiguously after second year). Birds for which we could not assign an age class confidently were not included in the analysis.

Three skeletal measurements (tarsus length, head plus bill length and head breadth) were taken with callipers to the nearest 0.1 mm by JWA or GCC. These measurements were usually taken only once, even if the bird was recaptured. In cases we measured the same bird repeatedly, we used the last measurement. We did not use wing and tail length as measures of size because, unlike the skeletal measurements, these differ between the populations studied and are correlated with age (Rasner et al. 2004). Body mass was measured to the nearest 0.1 g using a 20 g Pesola spring balance.

The load of ectoparasitic feather mites on the rectrices was quantified using a method adapted from Behnke et al. (1995, 1999). Unlike Behnke et al., we quantified ectoparasite loads on tail rather than wing feathers, because in juncos mites were much more abundant on tails than wings (personal observations). Each of the twelve tail feathers was individually

scored as follows: 0 = no parasites; 1 = few parasites, only near rachis, for less than half of its extension; 2 = parasites near the rachis for more than half of its extension; 3 = parasites along most of rachis and extending along barbs. The total parasite load was the sum of the scores from all twelve feathers. For birds lacking some tail feathers, the feathers symmetrical to the lost ones were counted twice. We use ectoparasite loads as an indication of susceptibility to infestation or ability to engage in behavioural or physiological defences.

We collected a small blood sample ($\approx$ 50 μl) from the wing vein. Hematocrit was calculated as the proportion of red blood cells relative to total volume after centrifugation in a micro-capillary tube at 10 000 rpm for 5 min. Hematocrit values are used extensively in avian studies as correlates of health and condition, with higher hematocrit levels generally corresponding to better condition (Brown 1996; Fair et al. 2007).

Mass, ectoparasites and hematocrit can change with date, and we measured them on each recapture of the same birds. Controlling for seasonal and other types of variation in those traits is important for obtaining indexes of condition (for a discussion of this on hematocrit, see Fair et al. 2007). Another reason to control for season effects is that capture and measurements were made at different times from when songs were recorded, and our indexes of condition should reflect long-term differences between individuals rather than short-term variation. We used all mass, ectoparasite and hematocrit data, including measurements on repeated captures of the same bird, in the regressions described below and then the residuals were averaged across the multiple captures of each individual.

To calculate a measure of body condition based on morphology, we computed body size as the first Principal Component (PC1) of a Principal Component Analysis on the three skeletal measurements for each bird. This PC explained 68% of the variation and had high loadings (>0.8) for all traits. We calculated body condition as the residual mass on the size PC, capture date, the square of capture date, year and population. Capture date (and its square), year and population were used to control for possible changes in condition with season or differences between years or populations. Similarly, we obtained the residual hematocrit on capture date, the square of capture date, year and population, and also the residual ectoparasite load on the same variables. Hereafter we refer to these residuals simply as body condition, hematocrit and ectoparasite load, respectively. These three aspects of condition were not correlated across birds (all $|t| < 0.11$, all p > 0.23, $N = 128$ birds).

Recordings and Song Performance

During the breeding seasons of 2006 and 2007, we recorded long-range song of male dark-eyed juncos. These are usually high amplitude songs broadcast by perched birds (Titus 1998; Nolan et al. 2002). Most long-range song types consist of a single trill (the same syllable repeated several times), but sometimes multi-syllabic songs existed (i.e. two or more trills of different syllables sung without interruption; Newman et al. 2008). Here we treated each of these trills (i.e. each different trilled syllable type) as separate song types. Details of recording procedures are in Cardoso et al. (2007). We obtained recordings of 128 males from which we also have an age assignment and morphological measurements (87 males in the campus population and 41 males in the mountain population). For 122 males, we had morphological measurements from the same year the males were recorded. Six males were

measured in 2006 but recorded in 2007; for these six males, we used the 2006 measurements as the best approximation to their morphometrics. We did not attempt to record complete song repertoires, and some males were recorded more often than others, depending mostly on the centrality of their territories relative to our study area and, consequently, how frequently we visited them. Therefore, our data contain variation on how many song types were sampled from each male. On average we recorded each of these 128 males on 6.22 occasions ($\pm$ 4.61 SD and range 1–22), and obtained 3.4 syllable types per male ($\pm$ 1.71 SD and range 1–8). This compares with an estimated average repertoire size of 4.2 different syllables per male in these populations (Newman et al. 2008).

In each recording, we selected a sequence of five songs of each song type for measurements. Juncos sing with eventual variety, repeating the same song type for long periods before switching (typically about 25 repetitions of the same song type, corresponding to about 3 min of singing, Titus 1998), such that most song bouts and recordings had a single song type. We measured all syllables except the first and last 10% of syllables from each song, rounded up to the nearest integer (because those syllables are sometimes softer and difficult to measure), and averaged the measurements within song type in each recording. In some short or noisy recordings, fewer than five songs were available for measurements. We measured ten syllable traits: frequency bandwidth (highest minus lowest sound frequency of the syllable), peak frequency (frequency with the highest amplitude), number of frequency inflections (times that a rising frequency modulation is followed by a descending one or *vice versa*), number of elements (discontinuous and temporally non-overlapping sounds), length of harmonics (octaves or other harmonics of the fundamental frequency), length of two voices (two simultaneous and independently modulated sounds), length of 'rattles' (harsh sounds with sub-elements repeated at >50 Hz), length of gaps (intervals within syllables), length of syllables and length of intervals between syllables. Measurements were made with Avisoft SA-SLAB version 4.34 (Avisoft Bioacoustics, Berlin, Germany) using recordings with a sampling rate of 22.05 kHz, either on power spectra (peak frequency of syllables), or on spectrograms with a fast Fourier transform length of 512 points, corresponding to a frequency bandwidth of 56 Hz and time resolution of 1.45 ms (all other syllable measurements). More detailed descriptions of these measurements are in Cardoso et al. (2007, 2008).

Cardoso et al. (2009) show that all syllables traits are highly repeatable within song types (even when those are sung by different males). Song types within a male's repertoire are therefore the statistical units for the song analyses in this study. We first averaged the measurements from recordings of the same song type, sung by the same male in the same year. Then we calculated the performance of each song type, for each male, in each year. We calculated 3 measures of performance: 'vocal deviation', which uses the measurements of frequency bandwidth and repetition rate of syllables, 'proportion of sound', which uses the measurements of syllable and interval lengths, and 'residual intervals', which uses the full set of syllable traits listed above. We explain these three measures of performance below, and Cardoso et al. (2009) illustrate examples of song types with high and low scores on each of them.

Vocal deviation is a measure based on the negative relation of frequency bandwidth and trill rate of songs (Podos 2001; Ballentine et al. 2004), which appears to result from a mechanical constraint to either modulate frequency widely or repeat syllables rapidly (Podos et al. 2004b; Riede et al. 2006). Such a negative relation exists in juncos (Cardoso et al. 2007), and we calculated vocal deviation as the orthogonal distance to the upper bound regression line relating bandwidth with trill rate across song types (See Cardoso et al. 2007). Large vocal deviations (song types well below the regression line) represent low-performance songs, and small vocal deviations (or negative, if above the upper bound regression line) represent high-performance songs.

Proportion of sound is a measure of performance based on the relation between the length of syllables and intervals. Intervals between syllables are often used to recover the air volume spent during the phonation of syllables, and in some species longer syllables require longer intervals for recovery. For example, domestic canaries take longer mini-breaths in trills with longer syllables because those use a greater amount of expired air during phonation (Hartley & Suthers 1989) and wild serins, whose song comprises mini-breaths (Depraz et al. 2006), also make longer intervals after longer syllables (Mota & Cardoso 2001). We calculated proportion of sound in juncos as the ratio of syllable length to the total length of syllables plus intervals. This is the same as the 'acoustic density' of Leadbeater et al. (2005) and the 'sound density' of Holveck & Riebel (2007), and in practice very similar to the 'percentage peak performance' of Forstmeier et al. (2002), although the latter is also influenced by changes in sound amplitude. High proportion of sound could indicate higher performance songs in terms of ventilation, because for the same syllable lengths, there are shorter intervals available for air volume recovery.

In juncos, the relation between the length of syllables and intervals is influenced by various other traits: only when controlling for the phonological differences among syllables (see list of syllable traits above) is the relation between syllable and interval length uncovered (Cardoso et al. 2007). In particular, syllable traits that imply interruption of airflow (such as rattles or multiple elements) have negative effects on interval duration (i.e. all else being equal require shorter intervals), and these traits need to be controlled for to detect the expected positive relation between the length of syllables and intervals. This indicates that the interval length is likely set by the phonation requirements of multiple syllable traits, not just syllable length, and that a better measure of performance related to ventilation would take those into account (Cardoso et al. 2007, 2009). Thus, we calculated residual intervals, as the residuals from a multiple regression of the interval lengths between syllables on all measured syllable traits. The regression equation is given in Table 2 of Cardoso et al. (2007). Songs with lower residual intervals are ones that, after controlling for syllable traits, have shorter intervals between syllables for recovery and can thus be inferred high-performance songs.

Comparisons Across Males

From our sample of 128 males, we wished to compare second-year (i.e. first breeding season) and after second-year males. For birds recorded in both 2006 and 2007 and that were in their second year in 2006, we used data only from 2006 (i.e. when they were in their

second year) and discarded the data from 2007 (i.e. when they were older than second year), because overall there are less second-year birds (n = 60) in our sample than after second-year birds (n = 68). For after second-year birds with the same song type recorded in both years, we used the average performance of the song type across both years (weighted by number of recordings each year). The two age classes did not differ in any of the measures of condition (*t*-tests, all $|t_{126}| < 1.6$, all p > 0.13).

For each measure of song performance, we ran a General Linear Model with a normally distributed error term, in which the dependent variable was the average performance across the song types recorded from each male's repertoire; age class and population were fixed factors; and body condition, hematocrit, and parasite load were covariates. The model also included the interaction term between the fixed factors (age and population).

For each measure of song performance, we also ran a similar model using the performance of the best song type recorded from each male, rather than average performance across its song types. This is because male singing ability may be better revealed by the most difficult song it sings (Ballentine 2009), rather than by the repertoire average. These models are identical to the above, but the number of song types recorded from each male was added as an additional covariate, to control for the increased chances of finding a high-performance song type in males with more song types sampled. Inspection of scatter plots of residuals against predicted values indicated that the assumptions of homoscedastic, independently distributed residuals were met for the linear models.

Longitudinal Comparisons

Thirty-three males were recorded in both breeding seasons. For these, we compared the average performance of song types recorded in 2006 and in 2007 from each male, with paired t-tests.

We also looked at the single best performance song type of these males in each year and tested whether it changed longitudinally while controlling for the number of song types recorded each year. To do this, we calculated the difference in best performance between years for each male (calculated so that for all measures a positive difference means an improvement in performance) and regressed it on the relative number of song types recorded each year. The relative number of song types recorded each year was computed as the proportion of song types recorded in 2007 [i.e. types recorded in 2007/(types recorded in 2006 + types recorded in 2007)] −0.5. Subtracting 0.5 makes the origin coincide with same number of song types recorded each year. A relation between best performance and relative number of song types is tested by the regression coefficient between those two variables. A longitudinal change in best performance is indicated by an intercept significantly different from the origin (above the origin if performance increased, or below if it decreased).

For 30 of these 33 males, we could record the same song types both years (average of 1.57 song types recorded both years per male). With data from these 30 males, we tested for within-song type longitudinal changes in performance with repeated measures ANOVA, where the measurement of performance in each year is the dependent variable, and bird

identity is a between-subjects variable to nest changes in individual song types within males. All statistical tests were run in SPSS 13.0 (SPSS Inc., Chicago, IL, USA).

Results

Comparisons Across Males

Table 1 shows the results of models relating each of the measures of performance (average for the song types recorded from each male) with age, population and measures of condition. Neither age (all $F_{1,121} < 3.3$, all p > 0.07) nor measures of condition (all $F_{1,121} < 2.4$, all p > 0.13) were significantly related with song performance measures (Table 1). The proportion of variation explained by each of these effects was always very low (all $\eta^2 < 0.03$, Table 1).

The only significant effect was the population difference in vocal deviation ($F_{1,121} = 10.7$, p = 0.001, significant after standard Bonferroni correction). This was because, as shown elsewhere, juncos in the urban setting have higher minimum song frequency (Slabbekoorn et al. 2007; Newman et al. 2008; Cardoso & Atwell 2011a,b) and consequently a smaller frequency bandwidth and higher vocal deviation (see methods). As this population difference can be explained as an adaptation to urban noise (Slabbekoorn et al. 2007; Cardoso & Atwell 2011a,b), rather than differences in male quality, it will not be discussed further.

Table 2 shows similar analyses, looking at the performance of the best song type recorded from each male, rather than the average of its song types. In all cases, the number of song types recorded from each male was strongly and positively related with the performance of the best song type (all $F_{1,120} > 21$, all p < 0.001, Table 2). In addition to this, males with better body condition tended to have a song type with higher proportion of sound ($F_{1,120} = 4.43$, p = 0.037, Table 2), but this is not robust to correction for multiple comparisons, and the proportion of variation explained by this effect was small ($\eta^2 = 0.036$).

After stepwise removal of non-significant effects in the above models (not shown), results were qualitatively identical.

Longitudinal Comparisons

Average performance of the song types recorded from the same male in consecutive years did not change significantly for any measure of performance: for vocal deviation, the average paired difference between the 1st and 2nd years was 0.17 ($\pm$ 0.11 standard error, SE, N = 33 males, an improvement in performance; $t_{32} = 1.59$, p = 0.12); for proportion of sound 2.0% ($\pm$ 1.5 SE, a decrease in performance; $t_{32} = 1.36$, p = 0.18); and for residual intervals −0.2 ms ($\pm$ 1.0 SE, a decrease in performance; $t_{32} = -0.20$, p = 0.84). We also found no evidence for longitudinal changes within-song type, as the effect of recording year on all measures of performance was weak (repeated measures ANOVAs, vocal deviation: $F_{1,17} = 0.45$, p = 0.51; proportion of sound: $F_{1,17} = 0.98$, p = 0.38; residual intervals: $F_{1,17} = 0.34$, p = 0.57).

Looking at the song type with the best performance each year, there was always a strong effect of the relative number of song types recorded each year: longitudinal changes in

performance of the best song types were positively related to changes in the number of song types recorded from each male (Fig. 1; vocal deviation: standardized regression coefficient $\beta_{st} = 0.64$, $N = 33$, p < 0.001; proportion of sound: $\beta_{st} = 0.51$, p = 0.003; residual intervals: $\beta_{st} = 0.62$, p < 0.001). After accounting for this, there was no evidence for consistent longitudinal changes in any measure of performance, as the intercepts were never significantly different from the origin (Fig. 1; vocal deviation: $t_{32} = 1.52$, p = 0.14; proportion of sound: $t_{32} = 0.01$, p = 0.99; residual intervals: $t_{32} = 1.23$, p = 0.23).

It might be that the most pronounced longitudinal change in best song performance takes place between the second and third years of age (Corbitt & Deviche 2005; Ballentine 2009), rather than when the birds are older. Seven birds that were recorded in both years had an age assignment of second-year in 2006 (plotted as triangles in Fig. 1), which is insufficient to test this hypothesis formally. But inspection of the existing data in Fig. 1 does not suggest that longitudinal changes in these younger birds differ from the larger sample.

Discussion

Long-range song of male dark-eyed juncos is primarily involved in territorial defence and interactions between males (Titus et al. 1997; Titus 1998; McGlothlin et al. 2007), and possibly also in female attraction (Hostetter 1961; Titus et al. 1997). In both these contexts, receivers should pay attention to signals that encode information on the quality of individuals (Grafen 1990; Johnstone 1995; Searcy & Nowicki 2005). We tested whether aspects of song performance based on trade-offs between song traits are a signal of male quality in dark-eyed juncos. These aspects of song performance were not related to body condition or to other measures of long-term physiological condition. Song performance was not related to male age either, and did not improve consistently as males grew older. Instead, for all measures of vocal performance, best performance was strongly related to the number of song types recorded from each male; that is, the more song types are sampled from each male, the more likely the sample was to include a high-performance song type.

The hypothesis that trade-off-based aspects of song performance can indicate male quality is very sensible because individuals likely differ in their ability to sustain physiologically demanding singing, which these measures of performance aim to quantify (Podos et al. 2004a; Cardoso et al. 2007). Also, in some species, there is evidence that high-performance songs are either preferred by females (Forstmeier et al. 2002; Ballentine et al. 2004) or more effective in aggressive interactions (Illes et al. 2006; Cramer & Price 2007; de Kort et al. 2009a), and recently Ballentine (2009) showed that older and heavier swamp sparrow males have better maximum performance. In contrast, we found no indication that in male dark-eyed juncos these aspects of vocal performance are related to age or quality, as evaluated by the measures of long-term condition that we used. This was not because of lack of statistical power, because the proportion of variation explained was consistently small (<4%) and sample sizes were large (over 100 males in the main analyses and 33 in the longitudinal comparisons). Our study is also the most comprehensive to date in terms of examining a greater variety of potential trade-offs between song traits as indicators of male quality, having used different trade-offs amongst a comprehensive set of traits in the frequency and

time domains (Podos 2001; Leadbeater et al. 2005; Cardoso et al. 2007, 2009; Holveck & Riebel 2007).

One possibility is that these traits are not good candidates for indicating quality in juncos. Such aspects of song performance are expected to indicate individual quality when songs are close to a performance constraint, in which case they may reveal high motor ability (Byers et al. 2010). Our global sample of junco songs appears close to performance constraints because the expected trade-offs could be detected (between frequency modulation and repetition rate, and between lengths of syllables and intervals, Cardoso et al. 2007), but it is still possible that many songs fall far from these constraints thus reducing the potential signalling value of the measures of performance.

Another possibility is that song repertoires make it difficult to assess differences in male quality based on song performance (Logue & Forstmeier 2008). Birds can only sample a subset of an individual's repertoire at a time and repertoires add variation to the potential information in song performance. This is particularly so in dark-eyed juncos, where differences in performance among song types within the repertoire of individual males are much larger than the average difference between males (Cardoso et al. 2009). Accordingly, we found that in all cases the number of song types recorded from each male had a strong effect on the best performance encountered. This shows that variation in performance of song types within repertoires can override the assessment of male quality based on singing ability. It also suggests that the quantity of song types sampled from each male, or even differences in repertoire sizes, can be a confounding factor when testing for differences in best song performance among males.

It could still be that male juncos signal quality relative to other males that sing the same song types, if we assume that birds are better at comparing song performance when listening to the same song type (Logue & Forstmeier 2008). However, song type sharing in dark-eyed juncos is low, with the majority of males not sharing song types with neighbours (Newman et al. 2008). Together, the lack of a relation between song performance and aspects of male quality (this study), the large differences in performance among song types within repertoires (Cardoso et al. 2009) and the low levels of song type sharing (Newman et al. 2008) imply that dark-eyed junco songs do not signal male quality via these trade-off-based measures.

Our results do not imply that these measurements of song performance are uninformative in juncos. As the performance a male exhibits depends on the song type it chooses from its repertoire, aspects of song performance may be employed in a dynamic way. For example, variation in performance among song types could be used to match the performance of other males during countersinging (Logue & Forstmeier 2008), or the different song types could function as a graded signal of aggression or motivation. According to the later possibility, there is some evidence that dark-eyed juncos use higher performance song types (types with fast trill rates and short intervals, and thus with low 'vocal deviation' and low 'residual intervals') during more motivated singing (Cardoso et al. 2009).

Other aspects of song performance, which we did not consider here, have been shown to signal aspects of male quality in other species. One possibility is repertoire size, which increases with age in some species (e.g. Gil et al. 2001; Kiefer et al. 2006). However, repertoire size is not related to age in other species (e.g. Nordby et al. 2002), including juncos (Newman et al. 2008), making it an unlikely candidate. Other examples are song rate and time spent singing, which signal energetic reserves and other aspects of condition in a number of species (reviewed in Podos et al. 2004a). Recently it has been found that singing with high consistency (i.e. identical syllables in a trill or identical songs in a song bout) can be related to male age or aspects of quality (Byers 2007; Botero et al. 2009; de Kort et al. 2009b) or that learning accuracy can also reflect past condition (Holveck et al. 2008; Brumm et al. 2009). Those other aspects of performance may not differ between song types, and therefore, it may be expected that they can signal male quality unaffected by variation among song types. In contrast, the song trade-offs we studied depend on the phonology of individual song types, and our results show that in dark-eyed juncos, this variation among song types obscures a possible signalling function of male quality. We conclude that possible signalling roles of these aspects of performance should be understood in a more dynamic way, taking into account variation within song repertoires and repertoire usage (Cardoso et al. 2009).

Acknowledgments

Literature Cited

Ballentine B. 2009; The ability to perform physically challenging songs predicts age and size in male swamp sparrows, *Melospiza georgiana*. Anim Behav. 77:973–978.

Ballentine B, Hyman J, Nowicki S. 2004; Vocal performance influences female response to male bird song: an experimental test. Behav Ecol. 15:163–168.

Behnke JM, McGregor PK, Shepherd M, Wiles R, Barnard C, Gilbert FS, Hurst JL. 1995; Identity, prevalence and intensity of infestation with wing feather mites on birds (Passeriformes) from Setúbal Peninsula of Portugal. Exp Appl Acarol. 19:443–458.

Behnke J, McGregor P, Cameron J, Hartley I, Shepherd M, Gilbert F, Barnard C, Hurst J, Gray S, Wiles R. 1999; Semi-quantitative assessment of wing feather mite (Acarina) infestations on passerine birds from Portugal. Evaluation of the criteria for accurate quantification of mite burdens. J Zool. 248:337–347.

Botero CA, Rossman RJ, Caro LM, Stenzler LM, Lovette IJ, de Kort SR, Vehrencamp SL. 2009; Syllable type consistency is related to age, social status and reproductive success in the tropical mockingbird. Anim Behav. 77:701–706. [PubMed: 20161558]

Brown ME. 1996; Assessing body condition in birds. Curr Orn. 13:67–135.

Brumm H, Zollinger SA, Slater PJB. 2009; Developmental stress affects song learning but not song complexity and vocal amplitude in zebra finches. Behav Ecol Sociobiol. 63:1387–1395. [PubMed: 19554102]

Byers BE. 2007; Extrapair paternity in chestnut-sided warblers is correlated with consistent vocal performance. Behav Ecol. 18:130–136.

Byers J, Hebets E, Podos J. 2010; Female choice based upon male motor performance. Anim Behav. 79:771–778.

Cardoso GC, Atwell JW. 2011a; Directional cultural change by modification and replacement of memes. Evolution. 65:295–300. [PubMed: 20722726]

Cardoso GC, Atwell JW. 2011b; On the relation between loudness and the increased song frequency of urban birds. Anim Behav. 82:831–836. [PubMed: 32952202]

Cardoso GC, Atwell JW, Ketterson ED, Price TD. 2007; Inferring performance in the songs of dark-eyed juncos (*Junco hyemalis*). Behav Ecol. 18:1051–1057.

Cardoso GC, Mamede AT, Atwell JW, Mota PG, Ketterson ED, Price TD. 2008; Song frequency does not reflect differences in body size among males in two oscine species. Ethology. 114:1084–1093.

Cardoso GC, Atwell JW, Ketterson ED, Price TD. 2009; Song types, song performance, and the use of repertoires in dark-eyed juncos (*Junco hyemalis*). Behav Ecol. 20:901–907.

Corbitt C, Deviche P. 2005; Age-related difference in size of brain regions for song learning in adult male dark-eyed juncos (*Junco hyemalis*). Brain Behav Evol. 65:268–277. [PubMed: 15761218]

Cramer ERA, Price JJ. 2007; Red-winged blackbirds *Ageliaus phoeniceus* respond differently to song types with different performance levels. J Avian Biol. 38:122–127.

Depraz V, Suthers R, Mota PG. 2006; Motor correlates of complexity, speed and frequency band-width in the song of the Serin: role of the duplex vocal organ and respiratory movement. J Ornithol. 147(Suppl.1):155–156.

Fair J, Whitaker S, Pearson B. 2007; Sources of variation in haematocrit in birds. Ibis. 149:535–552.

Forstmeier W, Kempanaers B, Meyer A, Leisler B. 2002; A novel song parameter correlates with extra-pair paternity and reflects male longevity. Proc Roy Soc B. 269:1479–1485.

Gil D, Gahr M. 2002; The honesty of bird song: multiple constraints for multiple traits. Trends Ecol Evol. 17:133–141.

Gil D, Cobb JLS, Slater PJB. 2001; Song characteristics are age dependent in the willow warbler, *Phylloscopus trochilus*. Anim Behav. 62:689–694.

Grafen A. 1990; Biological signals as handicaps. J Theor Biol. 144:517–546. [PubMed: 2402153]

Hartley RS, Suthers RA. 1989; Airflow and pressure during canary song: direct evidence for mini-breaths. J Comp Physiol A. 165:15–26.

Holveck M-J, Riebel K. 2007; Preferred songs predict preferred males: consistency and repeatability of zebra finch females across three testing contexts. Anim Behav. 74:297–309.

Holveck M-J, Vieira de Castro AC, Lachlan RF, ten Cate C, Riebel K. 2008; Accuracy of song syntax learning and singing consistency signal early condition in zebra finches. Behav Ecol. 19:1267–1281.

Hostetter DR. 1961; Life history of the Carolina junco *Junco hyemalis carolinensis Brewster*. Raven. 32:97–145.

Hutchinson JMC, McNamara JM, Cuthill IC. 1993; Song, sexual selection, starvation and strategic handicaps. Anim Behav. 45:1153–1177.

Illes AE, Hall ML, Vehrencamp SL. 2006; Vocal performance influences male receiver response in the banded wren. Proc Roy Soc B. 273:1907–1912.

Johnstone RA. 1996; Sexual selection, honest advertisement and the handicap principle: reviewing the evidence. Biol Rev. 70:1–65.

Kiefer S, Spiess A, Kipper S, Mundry R, Sommer C, Hultsch H, Todt D. 2006; First-year common nightingales (*Luscinia megarhynchos*) have smaller song-type repertoire sizes than older males. Ethology. 112:1217–1224.

de Kort SL, Eldemire ERB, Cramer ERA, Vehrencamp SL. 2009a; The deterrent effect of bird song in territory defense. Behav Ecol. 20:200–206. [PubMed: 19337589]

de Kort SL, Eldemire ERB, Valderrama S, Botero CA, Vehrencamp SL. 2009b; Trill consistency is an age-related assessment signal in banded wrens. Proc Roy Soc B. 276:2315–2321.

Leadbeater E, Goller F, Riebel K. 2005; Unusual phonation, covarying song characteristics and song preferences in female zebra finches. Anim Behav. 70:909–919.

Logue DM, Forstmeier W. 2008; Constrained performance in a communication network: implications for the function of song-type matching and for the evolution of multiple ornaments. Am Nat. 172:34–41. [PubMed: 18518774]

McGlothlin JW, Jawor JM, Ketterson ED. 2007; Natural variation in a testosterone-mediate trade-off between mating effort and parental effort. Am Nat. 170:864–875. [PubMed: 18171169]

Mota PG, Cardoso GC. 2001; Song organisation and patterns of variation in the serin (*Serinus serinus*). Acta Ethol. 3:141–150.

Newman MM, Yeh PJ, Price TD. 2008; Song variation in a recently founded population of the dark-eyed junco (*Junco hyemalis*). Ethology. 114:164–173.

Nolan, V, Ketterson, ED, Cristol, DA, Rogers, CM, Clotfelter, ED, Titus, RC, Schoech, SJ, Snajdr, E. Dark-eyed junco (*Junco hyemalis*), No. 716. In: Poole, A, Gill, F, editors. The Birds of North America. The Academy of Natural Sciences and The American Ornithologists Union; Philadelphia and Washington D.C.: 2002.

Nordby JC, Campbell SE, Beecher MD. 2002; Adult song sparrows do not alter their song repertoires. Ethology. 108:39–50.

Nowicki S, Searcy WA, Podos J. 1998; Song learning, early nutrition and sexual selection in songbirds. Am Zool. 38:179–190.

Nowicki1 S, Hasselquist D, Bensch S, Peters S. 2000; Nestling growth and song repertoire size in great reed warblers: evidence for song learning as an indicator mechanism in mate choice. Proc Roy Soc B. 267:2419–2424.

Podos J. 1997; A performance constraint in the evolution of trilled vocalizations in a songbird family (Passeriformes: Emberizidae). Evolution. 51:537–551. [PubMed: 28565357]

Podos J. 2001; Correlated evolution of morphology and vocal signal structure in Darwin's finches. Nature. 409:185–188. [PubMed: 11196640]

Podos J, Huber SK, Taft B. 2004a; Bird song: the interface of evolution and mechanism. Ann Rev Ecol Evol Syst. 35:55–87.

Podos J, Southall JA, Rossi-Santos MR. 2004b; Vocal mechanics in Darwin's finches: correlation of beak gape and song frequency. J Exp Biol. 207:607–619. [PubMed: 14718504]

Podos J, Lahti DC, Moseley DL. 2009; Vocal performance and sensorimotor learning in songbirds. Adv Stud Behav. 40:159–195.

Pyle, P. Identification Guide to North American Birds: Part I Columbidae to Ploceidae. Slate Creek Press; Bolinas: 1997.

Rasner CA, Yeh P, Eggert LS, Hunt KE, Woodruff DS, Price TD. 2004; Genetic and morphological evolution following a founder event in the dark-eyed junco, *Junco hyemalis thurberi*. Mol Ecol. 13:671–681. [PubMed: 14871370]

Riede T, Suthers RA, Fletcher NH, Blevins WE. 2006; Songbirds tune their vocal tract to the fundamental frequency of their song. Proc Nat Acad Sci U S A. 103:543–5548.

Searcy, WA, Nowicki, S. The Evolution of Animal Communication. Princeton University Press; Princeton: 2005.

Slabbekoorn H, Yeh P, Hunt K. 2007; Sound transmission and song divergence: a comparison of urban and forest acoustics. Condor. 109:67–78.

Spencer KA, Buchanan KL, Goldsmith AR, Catchpole CK. 2004; European starling (*Sturnus vulgaris*). Proc Roy Soc B. 271:S121–S123.

Titus RC. 1998; Short-range and long-range songs: use of two acoustically distinct song classes by dark-eye juncos. Auk. 115:386–393.

Titus RC, Chandler CR, Ketterson ED, Nolan V. 1997; Song rates of dark-eyed juncos do not increase when females are fertile. Behav Ecol Sociobiol. 41:165–169.

Yeh PJ, Price TD. 2004; Adaptive phenotypic plasticity and the successful colonization of a novel environment. Am Nat. 164:532–542.
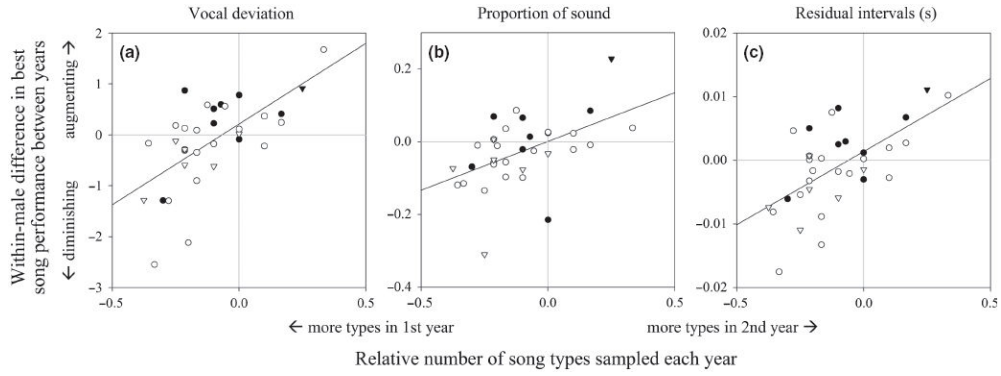
Fig. 1.

Relations between changes in performance of the best song type each year (vertical axis, zero means no change in performance) and relative number of song types recorded from each male in the 2 yrs (horizontal axis, zero means same number of song types recorded in both years). Filled and open symbols are mountain and campus males, respectively. Triangles are birds known to be changing from their second to third years of age. In all graphs positive differences in performance mean longitudinal improvements. Consistent longitudinal change in performance would be indicated by an intercept significantly different from the origin (not significant in any case, see text).

Table 1

Linear models on the average performance of all song types of each male. Indicated are the test statistic and its significance (F and p) and, in the second line, the proportion of variation explained by each factor (partial η^2). Significant effects are marked in bold

	Vocal deviation	Proportion of sound	Residual intervals
Age	$F_{1,121} = 2.41$ (p = 0.12)	0.59 (0.44) <0.01	0.09 (0.76) <0.01
	$\eta^2 = 0.02$		
Population	**10.7 (0.001)** 0.08	2.43 (0.12) 0.02	1.91 (0.17) 0.02
Age × Population interaction	2.38 (0.13) 0.02	0.45 (0.50) <0.01	0.81 (0.37) 0.01
Body condition	1.90 (0.17) 0.02	1.98 (0.16) 0.02	2.32 (0.13) 0.02
Hematocrit	1.94 (0.34) 0.01	1.34 (0.25) 0.01	0.34 (0.56) <0.01
Ectoparasite load	0.31 (0.58) <0.01	0.07 (0.79) <0.01	0.01 (0.91) <0.01

Table 2

Linear models on the performance of the best song type in each male's repertoire. Statistics as in Table 1

	Vocal deviation	Proportion of sound	Residual intervals
Age	$F_{1,120} = 1.96$ (p = 0.16)	0.29 (0.59) <0.01	0.01 (0.90) <0.01
	$\eta^2 = 0.02$		
Population	2.98 (0.09) 0.02	0.58 (0.45) <0.01	0.69 (0.41) 0.01
Age × Population interaction	1.08 (0.30) 0.01	0.26 (0.61) <0.01	0.17 (0.68) <0.01
Body condition	1.46 (0.23) 0.01	**4.43 (0.04)** 0.04	1.76 (0.19) 0.01
Hematocrit	0.84 (0.36) 0.01	0.663 (0.42) 0.01	<0.01 (0.99) <0.01
Ectoparasite load	0.01 (0.89) <0.01	0.23 (0.64) <0.01	0.01 (0.87) <0.01
Number of song types recorded	**21.1 (<0.001)** 0.15	**32.1 (<0.001)** 0.21	**41.3 (<0.001)** 0.27