Song types, song performance, and the use of repertoires in dark-eyed juncos (*Junco hyemalis*)

Gonçalo C. Cardoso,[a] Jonathan W. Atwell,[b] Ellen D. Ketterson,[b] and Trevor D. Price[c]
[a]Department of Zoology, University of Melbourne, Victoria 3010, Australia, [b]Department of Biology, Indiana University, 1001 East 3rd Street, Bloomington, IN 47405, USA, and [c]Department of Ecology and Evolution, University of Chicago, 1101 East 57th Street, Chicago, IL 60637, USA

Song performance encompasses the idea of how physiologically demanding different songs are to sing, and this is thought to reflect the singing ability of individual birds. In the dark-eyed junco (*Junco hyemalis*), each male has a repertoire of song types, some of which are shared with other males in the population. We used 4 measures of performance, based on trade-offs between song traits, to test if song performance is consistent among the song types making up the repertoire of individual males. We also tested if song types differ consistently in performance regardless of which males sing them. We found low but significant correlations of performance measures among the song types of individual males. This contrasts with highly consistent differences in performance among song types, regardless of which males sing them. We conclude that performance of single song types, as evaluated by trade-off–based measures of performance, gives little information about male singing ability. As song types differ in performance, we asked if males use the song types in their repertoires differently. We found that juncos use higher performance song types during bouts of more motivated singing, as evaluated by the length of songs, suggesting that song types may be preferentially used in different contexts depending on their performance. *Key words:* bird song, dark-eyed junco, *Junco hyemalis*, song performance, song repertoires. *[Behav Ecol 20:901–907 (2009)]*

Males of many bird species have repertoires of song types. Most explanations for the existence of repertoires treat song types as equivalent units (Catchpole and Slater 1995; Beecher and Brenowitz 2005). For example, females often respond more strongly to repertoires than they do to single song types (reviewed in Searcy and Nowicki 2000) and this is believed to indicate a preference for large repertoires per se, rather than for some song characteristic that may more likely appear in a large repertoire. A preference for repertoires per se was confirmed in 1 species, the zebra finch (*Taeniopygia guttata*), by testing preferences for more or less variable repertoires built with the same set of syllables (Collins 1999). Other explanations for the existence of repertoires are that different song types have different functions (e.g., Trillo and Vehrencamp 2005), that repertoires permit males to signal levels of aggressive intent by song type matching (Beecher and Brenowitz 2005), or that repertoires allow inferior males to avoid direct comparison with the songs of superior males by not matching (Logue and Forstmeier 2008).

Even when song types are apparently used interchangeably, different song types or even different renditions of the same type vary in ways that make some more efficient than others in stimulating females or in male–male interactions (e.g., Vallet et al. 1998; Ballentine et al. 2004; Illes et al. 2006; Holveck and Riebel 2007). Songs that are more efficient in stimulating females have been suggested to be more physiologically demanding to sing (reviewed in Podos et al. 2004; Searcy and Nowicki 2005) and, therefore, it is possible that receivers assess the physiological requirements of different songs in deciding how to respond to them. The term song performance has been used to refer to a variety of aspects of singing behavior at different levels of organization, from the proportion of time singing (e.g., Lambrechts and Dhondt 1987) to temporal or frequency consistency among identical song types (Byers 2007). In this paper, we focus on measures of song performance that quantify inferred physiological requirements of different songs, and that allow comparison of different song types. In the literature, these measures of performance have often been measured based on trade-offs between song traits that suggest physiological costs or limitations (reviewed in Podos et al. 2004; Cardoso et al. 2007). For example, longer syllables require longer intervals between them, putatively for respiratory recovery, so that songs with shorter intervals and longer syllables are hypothesized to be more demanding with respect to respiration (Hartley and Suthers 1989; Forstmeier et al. 2002). Also, in the dark-eyed junco (*Junco hyemalis*), aspects of syllable complexity trade off with sound amplitude, so that, at a given amplitude, more complex syllables should be more demanding to sing (Cardoso et al. 2007).

It is possible that song performance reflects the singing ability of different males, as is generally assumed (reviewed in Podos et al. 2004; Searcy and Nowicki 2005; Logue and Forstmeier 2008). For example, some males could be better than others at singing rapid trills or complex syllables. In this case, song performance would be a property of the singer and it should be consistent across the song types in the repertoire of each male. Alternatively, if song types intrinsically differ in performance independently of how different birds may sing them, then the song performance that a male exhibits depends on the song types it chooses to sing. This can have consequences for the function and use of repertoires. For example, males may choose their higher performance song type in aggressive interactions (Logue and Forstmeier 2008).

In this paper, we assess how performance varies among males and among song types in the repertoires of dark-eyed juncos. Our main objective is to test if song performance reliably reflects the singing abilities of different males, in which case it should be consistent for the song types that make up the repertoires of each male, or if performance is an intrinsic property of song types. Some existing measures of performance are
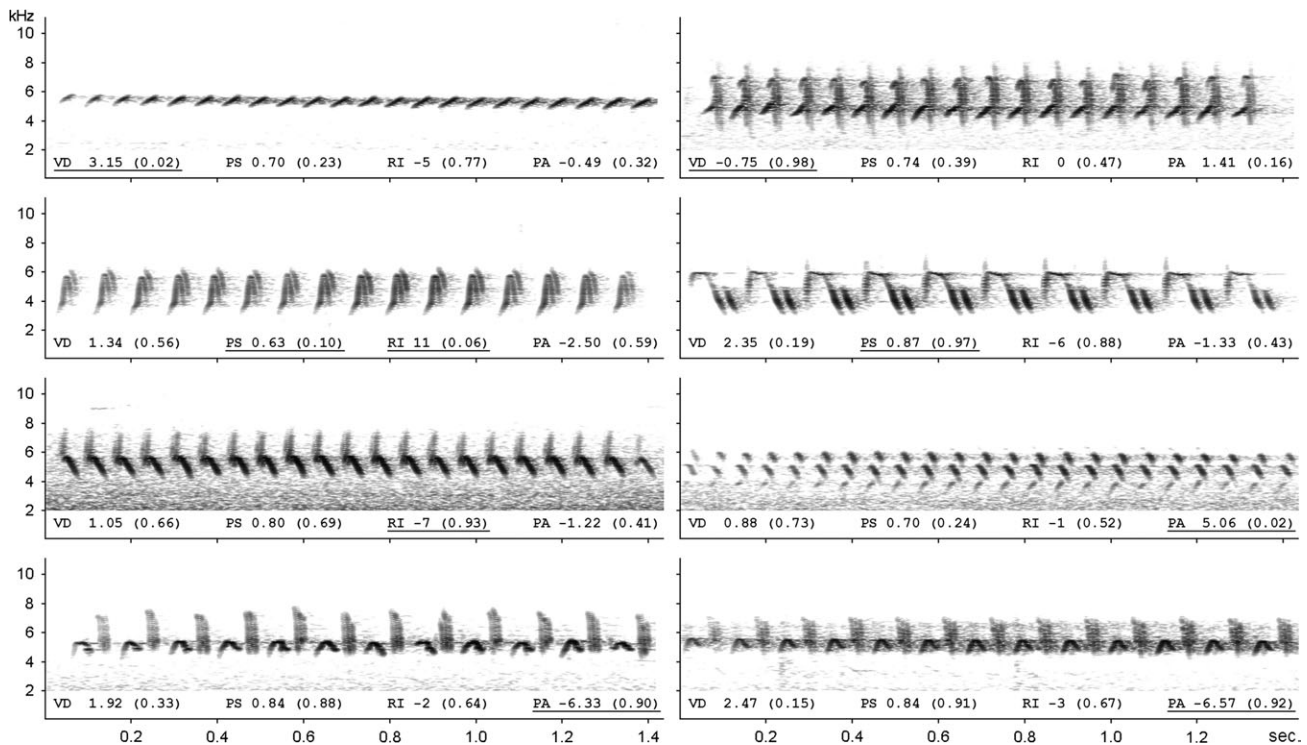
Figure 1

Spectrograms of junco songs illustrating high- and low-performance songs on each of the performance measures. On each spectrogram are the values of each measure (VD–vocal deviation, PS–proportion of sound, RI–residual intervals, in milliseconds, PA–predicted amplitude, in decibels relative to the mean of the analyzed sample) and, in parentheses, the percentile relatively to the entire sample of song types/males. Underlined is 1 example of a high-performance (percentile > 0.90) and a low-performance (percentile < 0.10) song for each measure. Note that, except for proportion of sound, lower values mean higher performance. The bottom row shows the same song type as sung by 2 different males.

based on the interaction of only a few song traits: trill rate and frequency bandwidth (Podos 1997, 2001) or the length of syllables and intervals (Leadbeater et al. 2005; Holveck and Riebel 2007; see also Forstmeier et al. 2002). However, song types differ in many other phonological characteristics. We previously showed that different song traits trade off either with the brevity of intervals between syllables or with sound amplitude in dark-eyed juncos, suggesting that certain combinations of song traits are difficult to sing (Cardoso et al. 2007). Based on these trade-offs, we proposed measures of song performance that take into account the effects of multiple song traits and, thus, may be more appropriate to compare among different song types.

Given differences in performance among song types, we also ask if males use the song types in their repertoires differently depending on their performance. To address this, we test if song performance changes with social context or with singing intensity.

MATERIALS AND METHODS

Recordings and song measurements

Here, we analyze the same dataset of male dark-eyed junco recordings as in Cardoso et al. (2007), and details of the study populations, recording procedures, and measurement of songs are described in detail there. Briefly, we recorded male dark-eyed juncos during the breeding season of 2006 in 2 banded populations in southern California: on the campus of the University of California at San Diego (32°52′N, 117°14′W) and at the Laguna Mountain recreation area (32°51′N, 116°26′W). We recorded both spontaneous song (468 recordings) and

song elicited by playbacks (114 recordings) of 97 free ranging males. On average, each male was recorded on 6.0 recording sessions (±4.83 standard deviation [SD]), mostly on different days. We noted if neighboring males were singing, either during or shortly before the recording. Dark-eyed juncos have trilled songs, and syllables are defined as the unit that repeats within the trill (Figure 1; Konishi 1964). Song types are defined by syllable structure. A total of 193 different song types were identified, 62 of which were shared by at least 2 males. We did not attempt to record the complete repertoires of the males. Some males were recorded over many days and from 42 of them we obtained from 4 to 8 different song types, which is on the high end of the range of repertoire sizes for this species (Williams and MacRoberts 1977, 1978; Newman et al. 2008). Other males were recorded few times and, therefore, for some, we only have a single song type (21 males, Figure 2). On average, we recorded 3.23 (±1.78 SD) different song types per male.

We measured a sequence of 5 songs of each type adjacent in each recording. For those songs, we measured all syllables except the first and last 10% of each song (because those are sometimes softer and difficult to measure) and averaged the measurements within song type in each recording. In short recordings or recordings with noisy portions, fewer than 5 songs were sometimes measured; on average, we measured 4.45 songs per recording. We measured 10 syllable traits: frequency bandwidth, peak frequency, number of frequency inflections, number of elements, length of harmonics, length of 2 voices (simultaneous and independent use of both sides of the syrinx), length of "rattles," length of gaps, length of syllables, and length of intervals between syllables. Most song types consist of a single trill, but in 13 recordings (of 582), multisyllabic songs existed that contain 2 or more trills of different
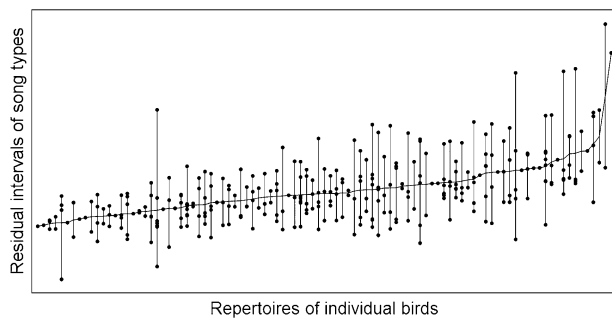
Figure 2
Plot of a measure of song performance, residual intervals (residuals from a regression of intervals between syllables on all other song traits measured), over the song types in the repertoire of male dark-eyed juncos. Points are averages of each song type for each male and vertical lines connect the song types in the repertoire of each male. Birds are ordered by decreasing average performance of their song types (curved line). High-performance song types are the ones with low residual intervals.

syllables (Konishi 1964; Williams and MacRoberts 1977; Newman et al. 2008). We treated each of these trills as separate song types. Definitions of these traits and details on measurements are in Cardoso et al. (2007); a more detailed description of the frequency measurements is in Cardoso et al. (2008). In addition, we measured average song length in each recording.

Song traits in the 2 recorded populations are similar except for the campus birds having, on average, higher minimum frequency (as shown for these populations by Slabbekoorn et al. 2007; Newman et al. 2008; our data). The difference between populations (from our data, on average 0.42 kHz) is less than the variation between song types in the repertoire of individual males (average SD within the repertoire of males is 0.49 kHz), and therefore, we analyze the 2 populations together.

Measures of performance

We used 4 measures of performance, hereafter referred to as 1) "vocal deviation," 2) "proportion of sound," 3) "residual intervals," and 4) "predicted amplitude."

Vocal deviation is the measure proposed by Podos (2001), based on the negative relation of frequency bandwidth and trill rate of songs. Such a negative relation exists in juncos and, following Ballentine et al. (2004), we calculated the upper bound regression relating bandwidth with trill rate across the syllable types with the largest bandwidth in each 1-Hz rate class (Figure 3 in Cardoso et al. 2007). Then, for each recording, we measured the orthogonal distance to the upper bound regression line. Large vocal deviations (points well below the regression line) represent low-performance songs and vice versa.

The second measure of performance, proportion of sound, quantifies the relative lengths of syllables and intervals between syllables, and is calculated as the ratio: (average length of syllables)/(average length of syllables plus intervals). This is the same as the "acoustic density" and "sound density" of Leadbeater et al. (2005) and Holveck and Riebel (2007), and similar to the "percentage peak performance" of Forstmeier et al. (2002), though the latter is also influenced by changes in sound amplitude. While singing, songbirds can recover the air volume expired during phonation taking short inspirations at the intervals between syllables (mini-breaths, Hartley and Suthers 1989). Some syllable types (31 of 193) had short gaps within syllables, but those are much shorter than the intervals between syllables, and too short for mini-breaths (average = 7.0 ms ±5.0 SD, $n = 31$; see Hartley and Suthers 1989), and therefore were not considered here. Their

possible effect on the length of intervals, together with that of various other syllable traits, is accounted for with the next measure of performance.

Residual intervals is a measure designed to quantify how demanding songs are in respiratory terms, taking into account variation in syllable traits. Various syllable traits influence the duration of intervals between syllables (Cardoso et al. 2007), and therefore taking those into account may provide an accurate measure of the respiratory requirements of songs. Residual intervals are calculated from a multiple regression of the intervals between syllables on all the syllable traits. The syllable traits are listed in the previous section, and together they describe variation in syllable phonology comprehensively. This regression was computed based on data from 188 different syllable types, each data point being a different syllable type, and the regression equation is given in Table 2 of Cardoso et al. (2007). Songs with low residuals are ones that, after controlling for syllable traits, have short intervals between syllables and are therefore best considered as high-performance songs, and vice versa for songs with high residuals. Proportion of sound and residual intervals are correlated ($r = -0.63$, $N = 193$ syllable types, $P < 0.01$), which is expected because these are 2 alternative ways to quantify a respiratory trade-off.

Our fourth measure of performance (predicted amplitude) is based on the trade-off between sound amplitude and aspects of syllable complexity in dark-eyed juncos (Cardoso et al. 2007). This is based on comparisons of amplitudes within recordings that contained multiple song types, because recording conditions vary among different recordings. During these recordings ($N = 34$), gain levels were not adjusted, and the distance to the singing bird remained unchanged. These trade-offs with amplitude provide a tool to quantify syllable phonology in a way that reflects its difficulty of singing: We used the multiple regression of sound amplitude on syllable traits in Cardoso et al. (2007) to predict how demanding different songs are to sing at a given amplitude (i.e., how low is their predicted relative amplitude).

Figure 1 shows examples of songs with high and low scores on each of the above measures of performance. All 4 measures of performance are composites based on trade-offs between song traits (Podos 1997; Forstmeier et al. 2002; Cardoso et al. 2007). We repeated our analyses using each of the individual song traits instead of the measures of performance, to determine if the patterns we find with performance can instead be attributed to variation in simpler song traits.

Consistency of performance within males across song types and within song types across males

Using average values for the recordings of each song type within each male, we calculated the intraclass correlations (Sokal and Rohlf 1981) of song performance and of song traits within males (i.e., the correlation across the song types in the repertoire of each male) and within song types (i.e., the correlation across different males' renditions of the same song type). Positive correlations of performance among the song types sung by each male would show that song performance is a property of the individual males. Positive correlations of performance within song types indicate that song performance is a property of the song types, independently of which males sing them.

Performance and motivation

We compared performance measures and individual song traits between recordings elicited by playbacks, recordings without playbacks but that were made when neighboring birds sang, and recordings that were made when no other birds sang. This is a coarse categorization of the intensity of social

Table 1

Intraclass correlation coefficients (r) within males and within song types

	Intraclass correlations within males			Intraclass correlations within song types		
	r	$F_{96,216}$	P	r	$F_{192,120}$	P
Measures of performance						
Vocal deviation	0.119	1.433	0.016[a]	0.756	6.022	<0.001[a]
Proportion of sound	0.119	1.434	0.016[a]	0.645	3.935	<0.001[a]
Residual intervals	0.141	1.529	0.005[a]	0.554	3.006	<0.001[a]
Predicted amplitude	0.081	1.282	0.071	0.919	19.298	<0.001[a]
Song traits						
Frequency bandwidth	0.196	1.782	0.0002[b]	0.705	4.870	<0.001[b]
Peak frequency	0.025	1.081	0.318	0.564	3.096	<0.001[b]
Number of frequency inflections	0.018	1.062	0.356	1[c]	—	—
Number of elements	−0.001	0.996	0.500	1[c]	—	—
Length of harmonics	−0.045	0.861	0.798	0.980	81.825	<0.001[b]
Length of 2 voices	0.001	1.002	0.486	0.901	15.637	<0.001[b]
Length of "rattles"	−0.021	0.935	0.642	0.945	29.804	<0.001[b]
Length of gaps	−0.036	0.893	0.735	0.763	6.221	<0.001[b]
Length of syllables	−0.022	0.931	0.652	0.865	11.378	<0.001[b]
Length of intervals	0.131	1.484	0.009	0.588	3.309	<0.001[b]

[a] Significant P values using sequential Bonferroni correction for testing 4 measures of performance.

[b] Significant P values using sequential Bonferroni correction for testing 10 different song traits.

[c] Perfect intraclass correlation because there is no variation within song types for these traits.

interactions, dictated by the type of information that we could collect reliably during field recordings. We treated recordings as the replicate and included male identity as a random factor in analyses of variance (ANOVAs) in order to control for differences between males. Recordings that used playbacks but where neighbors also sang were included in the category "playbacks," because we consider this the stronger stimulus.

We found that, the length of songs is a plastic trait that varies with social context in a way that seems to reflect overall motivation or singing intensity (see results). Therefore, we used the average song length in each recording as an index of motivation, as it may reflect the motivational state of singing males more accurately than our simple categories of social context. We excluded recordings in which birds sang more than 1 song type or syllable type (34 of 582 recordings). We correlated residual song length with each measure of performance and song trait across recordings controlling for difference between males and between song types. In a first analysis, the identity of the recorded male was controlled for by using residual song length from a random-effects ANOVA of song length on males. In this first analysis, associations of song performance or song traits with song length could be due to males choosing different song types during more intense singing, or to changes within the same song types. In a second analysis, both male identity and song type were controlled for. Here, we used residual song length from a random-effects ANOVA where both male and song type are random factors. In this second analysis, associations of song performance or traits with song length are due only to changes within song types. In this last analysis, we used only 4 of the simple song traits (length of syllables, length of intervals, peak frequency, and frequency bandwidth) because there is no variation within song type and male for the other traits.

Throughout the Results section, we present simple significance values and refer to Bonferroni corrections when appropriate.

RESULTS

### Consistency of performance within males and within song types

The 4 measures of performance had low intraclass correlations within males (all r between 0.08 and 0.14), meaning that the song types in the repertoire of each male vary in performance, but these intraclass correlations were always significant or close to significance ($0.005 < P < 0.071$, Table 1). In contrast, intraclass correlations within song types as sung by different males were high for all 4 measures of performance (all $r >$ 0.55, all $P < 0.01$, Table 1), meaning that the performance of the same song type is consistent across males. In Figure 2, we plot the residual intervals of all sampled song types, which was the measure of performance with the highest intraclass correlation within males. The figure shows the generally small differences between males when compared with the large variation among song types.

As with the measures of performance, the simple syllable traits had weak correlations within males (all $r < 0.20$, Table 1) and were highly consistent within song types (all $r > 0.56$, all $P < 0.01$, Table 1). Only the length of intervals between syllables ($r = 0.13$, $F_{96,216} = 1.48$, $P = 0.01$) and frequency bandwidth ($r = 0.20$, $F_{96,216} = 1.78$, $P < 0.01$) showed a tendency to be consistent within males. The result on frequency bandwidth was due to both minimum and maximum frequencies being consistent within males ($r = 0.19$, $F_{96,216} = 1.78$, $P < 0.01$ and $r = 0.16$, $F_{96,216} = 1.63$, $P < 0.01$, respectively).

Performance and motivation

Song length was greater in recordings elicited by playbacks (raw average across all recordings was 1.45 s ± 0.02 standard error, $N = 114$ recordings) or when other birds sing (1.46 ± 0.02, $N = 134$), than in recordings where no other birds sing (1.36 ± 0.01, $N = 334$; controlling for differences between males, $F_{2,483} = 15.94$, $P < 0.01$, significant differences in post hoc Tukey test were playbacks > none, and other birds > none, both $P < 0.01$). Additionally, after controlling for differences in song length among song types, by using residuals of an ANOVA of song length on song types, differences between recording contexts remained equally significant ($F_{2,483} = 14.56$, $P < 0.01$). Therefore, song length is a plastic trait within males and within song types, and seems to reflect motivation. No other song traits or measures of performance varied significantly with social context (all $F_{2,519} < 2.20$, all $P > 0.11$).
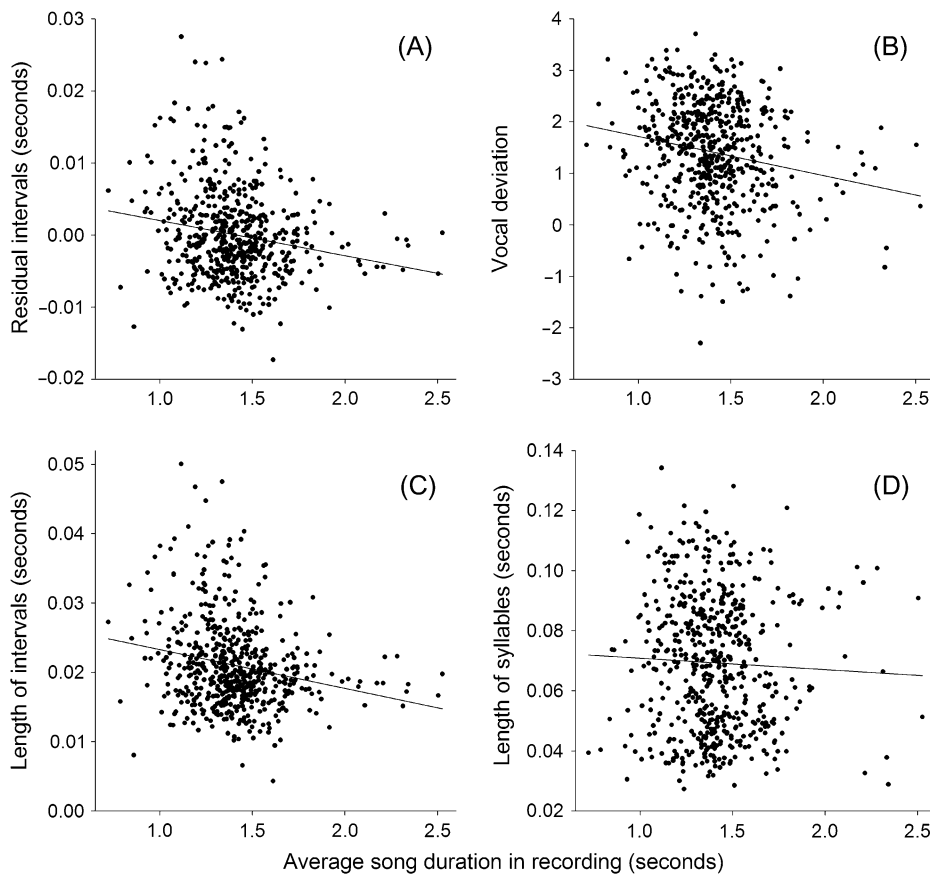
Figure 3
Scatter plots of the significant relations of measures of performance (A, B) and song traits (C, D) with the average song length in each recording. Points are different recordings. These plots are only for illustration purposes, as tests of these relations were done controlling for differences among males (see text).

We used average song length as an index of singing motivation and regressed it on each song trait and measure of performance among recordings. Controlling for differences between males, 2 measures of performance were related to song length: Residual intervals and vocal deviation both decreased with song length ($r = -0.15$ and -0.14, respectively, both $N = 548$, both $P < 0.01$; Figure 3A,B).

Among the simple syllable traits, the length of intervals and of syllables were both negatively related to song length ($r = -0.17$ and -0.12, respectively, both $N = 548$, both $P < 0.01$; Figure 3C,D) and this explains the results obtained with the measures of performance. First, interval length was more strongly associated with song length than residual intervals were and, therefore, the result with residual intervals can be explained by simple differences in interval length. Second, trill rate (which is the inverse of the length of syllables plus intervals) was more strongly associated with song length ($r = 0.16$, $N = 548$, $P < 0.01$) than vocal deviation was and, therefore, the result with vocal deviation (which is calculated from a regression on trill rate) can be explained by simple differences in trill rate. All above correlations are significant after Bonferroni corrections for multiple comparisons across 4 measures of performance or 10 syllable traits. None of the other syllable traits was correlated with song length (all $|r| < 0.09$, all $P > 0.05$).

These associations of song traits with average song length could be due to choice of different song types during more motivated singing or to changes within the same type. When we control for differences between song types, there was no significant correlation of any measure of performance or syllable trait with song length (all $|r| < 0.05$, all $P > 0.30$). Therefore, long songs, which putatively reflect more motivated singing, are sung with shorter intervals and shorter syllables mostly because males choose song types with short intervals and syllables rather than modify these traits within song types.

DISCUSSION

We found that song performance is highly consistent within song types, independently of which males sing them, but that consistency is low across song types within the repertoire of individual males ($r \approx 0.1$). Therefore, the song performance that a male shows, as evaluated by these trade-off–based measures of performance, depends mostly on which of the song types in his repertoire the male sings. Despite this, the intraclass correlations within the repertoire of individual males, though low, are significant or nearly significant for all the measures of performance. Therefore, we do not exclude the possibility that song performance conveys information about the quality of males, as suggested by studies with dusky warblers (*Phylloscopus fuscatus*, Forstmeier et al. 2002) and chestnut-sided warblers (*Dendroica pensylvanica*, Byers 2007), where males with higher song performance obtained more extrapair paternity. In the dusky warbler study (Forstmeier et al. 2002), song performance also had low repeatability within males, but it was still a good predictor of extrapair fertilizations. The main conclusion from our results is that, because most of the variation in performance depends on the song type, a receiver that compares a few song types from different males is likely to obtain little information about performance differences between males.

Many of the syllable traits are used in the assignment of songs to song types, so it is not surprising to find that quantitative measurements of those traits are correlated within the same song type from a male to another, and that measures of performance are therefore similar for the same song type.

However, different males could conceivably learn the same song type and then sing it according to their singing abilities by adjusting key traits, such as the intervals between syllables. Birds learning song types of performance above their physiological capabilities must modify them in some way (Podos 1996), but we do not know if birds learning song types with performance below their capabilities will improve them. The high consistency of performance measures in song types sung by different males suggests that males do not generally improve the performance of learned song types.

Because song types appear to differ in performance, we might expect that birds choose to learn only song types of high performance, within their capabilities, or that they would improve the poorer song types. We suggest some possibilities for why this was not the case in dark-eyed juncos. It is possible that repertoires of dissimilar song types are maintained in order to engage in song matching (Beecher and Brenowitz 2005; or avoid song matching, Logue and Forstmeier 2008), or because females prefer diversity (reviewed in Catchpole and Slater 1995) or copying accuracy (Nowicki et al. 2002; Coleman et al. 2007). It could also be useful to learn song types of lower performance if singing ability by adulthood cannot be well predicted as juveniles and birds cannot modify learned song types according to their singing abilities.

If song performance has a signaling function, males may use their higher performance song types during more intense social interactions, and in this way the presence of both low- and high-performance song types in repertoires may be favored. For example, using vocal deviation, Beebee (2004) found that the song types that yellow warblers (*Dendroica petechia*) use mostly in male–male interactions are high-performance ones. Dark-eyed juncos sang longer songs in response to playbacks or when neighboring males are singing. Using average song length as an index of motivation, we found that longer songs are indeed sung with higher performance, in terms of respiratory trade-offs. This is due to choice of high-performance song types rather than to changes within song type. Note that from a purely physiological standpoint, we would expect the opposite trend: High-performance songs, which have short intervals between syllables, should tend to be shorter because of air volume depletion or hypoxia (Hartley and Suthers 1989). Thus, the positive relationship between performance and song length strongly suggests a motivational phenomenon.

Measures of performance are useful insofar as they show biologically meaningful results that are not observed with the simple song traits used in their computation. For most of the analyses presented here, this was not the case. Results obtained with the measures of performance that quantify respiratory trade-offs (proportion of sound and residual intervals) were similar to the ones obtained with simple interval lengths, which is a key trait in their computation. Therefore, for this species and for the analyses in this paper, these 2 measures of performance did not improve on the insight from studying simple interval lengths. Likewise, the results that we obtained using vocal deviation can be explained either by rate or bandwidth alone, and never by the interaction of the 2: The consistency of vocal deviation within repertoires was due to bandwidth alone, and vocal deviation was related to song length due to rate alone, as males choose faster song types during more intense song bouts. Therefore, vocal deviation did not improve on results obtained with the traits used in its computation. On the contrary, it confounded distinct effects of these 2 traits. Finally, predicted amplitude (the measure of performance based on trade-offs with amplitude) was the only measure showing a trend (it tends to be consistent within the repertoires of each male) that is not found with any of the traits important in its computation (Cardoso et al. 2007). Together with the finding that loud-

ness is the aspect of song output with which song traits trade off more strongly (Cardoso et al. 2007), this suggests that amplitude may be an important aspect of how birds evaluate song performance.

The finding that song performance, assessed by trade-off–based measures of performance, is a property of the different song types rather than the singers suggests that it is not a reliable signal of male quality, particularly when receivers evaluate only 1 or a few song types. This has implications for the function of repertoires. For example, high-performance song types seem to be used by dark-eyed juncos during more motivated singing. Other possibilities, not tested here, are that the choice of song types depends simultaneously on the performance and similarity of the opponent's song (Logue and Forstmeier 2008), or that high-performance songs are used when directing singing to visible conspecifics (Kreutzer et al. 1999) or to fertile females (Ballentine et al. 2003). These possibilities require experimental approaches to clarify how differences of performance among song types shape repertoire usage.

REFERENCES

Ballentine B, Badyaev A, Hill GE. 2003. Changes in song complexity correspond to periods of female fertility in blue grosbeaks (*Guiraca caerulea*). Ethology. 109:55–66.

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

Beebee MD. 2004. Variation in vocal performance in the songs of a Wood-Warbler: evidence for the function of distinct singing modes. Ethology. 110:531–542.

Beecher MD, Brenowitz EA. 2005. Functional aspects of song learning in songbirds. Trends Ecol Evol. 20:143–149.

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

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.

Catchpole CK, Slater PJB. 1995. Bird song. Biological themes and variations. Cambridge: Cambridge University Press.

Coleman SW, Patricelli GL, Coyle B, Siani J, Borgia G. 2007. Female preferences drive the evolution of mimetic accuracy in male sexual displays. Biol Lett. 3:463–466.

Collins SA. 1999. Is female preference for male repertoires due to sensory bias? Proc R Soc Lond B Biol Sci. 266:2309–2314.

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

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

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

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

Konishi M. 1964. Song variation in a population of Oregon juncos. Condor. 66:423–436.

Kreutzer M, Beme I, Vallet E, Kiosseva L. 1999. Social stimulation modulates the use of the 'A' phrase in male canary songs. Behaviour. 136:1325–1334.

Lambrechts M, Dhondt AA. 1987. Differences in singing performance between male great tits. Ardea. 75:43–52.

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.

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.

Nowicki S, Searcy WA, Peters S. 2002. Quality of song learning affects female response to male bird song. Proc R Soc Lond B Biol Sci. 269:1949–1954.

Podos J. 1996. Motor constraints on vocal development in a songbird. Anim Behav. 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 SK, Taft B. 2004. Bird song: the interface of evolution and mechanism. Ann Rev Ecol Evol Syst. 35:55–87.

Searcy WA, Nowicki S. 2000. Male-male competition and female choice in the evolution of vocal signalling. In: Espmark Y, Amundsen T, Rosenqvist G, editors. Animal signals: signalling and signal design in animal communication. Trondheim (Norway): Tapir Academic Press. p. 301–315.

Searcy WA, Nowicki S. 2005. The evolution of animal communication. Princeton (NJ): Princeton University Press.

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

Sokal RR, Rohlf FJ. 1981. Biometry. San Francisco (CA): WH Freeman and Co.

Trillo PA, Vehrencamp SL. 2005. Song types and their structural features are associated with specific contexts in the banded wren. Anim Behav. 70:921–935.

Vallet E, Beme I, Kreutzer M. 1998. Two-note syllables in canary songs elicit high levels of sexual display. Anim Behav. 55:291–297.

Williams L, MacRoberts MH. 1977. Individual variation in songs of dark-eyed juncos. Condor. 79:106–112.

Williams L, MacRoberts MH. 1978. Song variation in dark-eyed juncos in Nova Scotia. Condor. 80:237–240.