Behav Ecol Sociobiol (2011) 65:717–726
DOI 10.1007/s00265-010-1073-2

ORIGINAL PAPER

Discrimination of vocal performance by male swamp sparrows

**Adrienne L. DuBois · Stephen Nowicki ·
William A. Searcy**

Received: 16 April 2010 / Revised: 25 August 2010 / Accepted: 4 October 2010 / Published online: 16 October 2010

Abstract In aggressive communication, the interests of signalers and receivers are directly opposed, presenting a challenge to the maintenance of reliable signaling. Index signals, whose production is constrained by physical ability, offer one solution to the reliable signaling problem. Vocal performance, the ability to produce physically challenging songs, is likely such a signal in swamp sparrows. Maximum vocal performance varies between males and is correlated with aspects of quality. However, vocal performance can be modulated in aggressive contexts by increasing the frequency bandwidth and trill rate of songs. This study examines receiver response to (1) differences in performance of the same song types by different signalers and (2) individual modulation of performance between contexts. Results demonstrate that male receivers show differential response to between-male differences in song type performance, but do not show differential response to the smaller scale modulations of performance produced by individuals singing the same song type at different times. This pattern suggests that vocal performance cannot be effectively cheated and may therefore serve as a good example of an index signal.

Communicated by I. Hartley

A. L. DuBois (✉) · W. A. Searcy
Department of Biology, University of Miami,
Coral Gables, FL 33146, USA
e-mail: aldubois@bio.miami.edu

W. A. Searcy
e-mail: wsearcy@miami.edu

S. Nowicki
Department of Biology, Duke University,
Durham, NC 27708, USA
e-mail: snowicki@duke.edu

Keywords Birdsong · Vocal performance · Aggression · Reliable signaling · Communication

Introduction

The evolutionary stability of signal reliability is problematic whenever the interests of signalers and receivers are in conflict. In particular, the maintenance of signal reliability is puzzling in aggressive contexts given that the interests of signalers and receivers are in direct opposition. Signals are considered reliable if they convey dependable information about the signaler or its environment that benefits the receiver of the signal (Searcy and Nowicki 2005). In aggressive contexts, receivers typically benefit from information on either the signaler's fighting ability or its aggressive intentions. Signalers would benefit if they could exaggerate signals to intimidate receivers, but if exaggeration becomes sufficiently widespread, then selection will act on receivers to cease responding to the signals, causing the signaling system to break down (Dawkins and Krebs 1978). Mechanisms that maintain signal honesty are therefore crucial for the evolutionary stability of aggressive signaling systems (Grafen 1990). Various mechanisms have been proposed to maintain reliability in aggressive contexts (Vehrencamp 2000), including the possibility of 'physical constraints' (Hurd and Enquist 2005). Here, we investigated receiver response to a vocal signal—vocal performance—that may be subject to physical constraints in swamp sparrows (*Melospiza georgiana*). We measured the response of male receivers to vocal performance as a test of whether vocal performance functions in aggressive signaling and to explore the mechanisms that maintain the reliability of this signal.

Signals are physically constrained if the mechanism by which the signal is produced determines a fixed relationship between the signal and some characteristic of the signaler. Such signals are termed index signals (Maynard Smith and Harper 2003). Because the intensity of index signals is causally related to the quality being signaled, they are considered "unfakeable." Since index signals are tied to physical attributes of signalers, they are likely candidates for conveying information about fighting ability or resource-holding potential (Hurd and Enquist 2005). Well-established examples of index signals are few, however. One oft-cited case involves the dominant acoustic frequency of calls in some frogs and toads, which is constrained by size (Davies and Halliday 1978, Ryan 1985, Wagner 1989a,b). Larger males possess more massive vocal cords, which allow them to produce calls with lower dominant frequencies (Martin 1972). Larger males also tend to win wrestling contests over resources (Davies and Halliday 1978). Because of this relationship between male size, vocal morphology, and the ability to win contests, males of many anuran species may assess opponents using the dominant frequency of calls as an unbluffable signal of size and fighting ability (Wagner 1989a). Another likely index signal is the formant dispersion of roars in red deer (*Cervus elaphus*, Reby and McComb 2003; Reby et al. 2005). The length of the vocal tract determines its resonant properties, including the spacing of vocal tract resonances or "formants," with individuals having longer vocal tracts producing more closely spaced formants (Fitch 1997; Reby and McComb 2003). Because vocal tract length tends to correlate with overall body size due to anatomical constraints, the spacing of formants is a reliable index of body size.

It turns out, however, that signalers can and do exaggerate these ostensibly reliable signals in aggressive situations (Searcy and Nowicki 2005). Some frogs are able to lower the dominant frequency of their calls in response to aggressive playback (Wagner 1989a; Bee and Perrill 1996; Bee et al. 1999). Red deer are also able to decrease the formant dispersion of their roars by lowering their larynx and extending their vocal tract (Fitch and Reby 2001). This exaggeration would undermine signal reliability if it occurs at a level salient to receivers because both red deer and frogs exaggerate their signals to different degrees in different contexts (Wagner 1989a; Bee and Perrill 1996; Reby et al. 2005).

Another potential candidate for an index signal is "vocal performance" of birdsong, defined as the ability to sing songs that are physically or physiologically challenging to produce owing to constraints on vocal production mechanisms (Podos 2001; Ballentine et al. 2004; Cardoso et al. 2009; Forstmeier et al. 2002). Vocal performance has been measured in several ways, including vocal deviation (Podos 2001), percentage peak performance (Forstmeier et al. 2002), acoustic density (Cardoso et al. 2009, Holveck and Riebel 2007, Leadbeater et al. 2005), residual intervals, and predicted amplitude (Cardoso et al. 2007; Cardoso et al. 2009). We employ the measure of 'vocal deviation' first described by Podos (2001), a measure of the ability to produce repeated notes, or "trills," both rapidly and with a broad frequency bandwidth. This measure is based on a tradeoff that most likely results from morphological constraints on song production. Sound produced by the syrinx is modified by the vocal tract, which acts as a resonance filter, to allow production of the pure tonal sounds that characterize birdsong (Nowicki 1987). In order to produce pure tones, a bird must modify the configuration of its vocal tract to correspond to the sound frequency produced at the syrinx, specifically by altering bill gape (Nowicki and Marler 1988; Podos and Nowicki 2004a). For example, Westneat et al. (1993) demonstrated that sparrows change bill gape to track the frequency of the sounds that they emit, opening the bill wide for high-frequency sounds and reducing gape to produce low-frequency sounds. Hoese et al. (2000) confirmed that bill gape affects sound production through experiments in which bill movements were constrained. Thus, birds are physically limited by the difficulty of simultaneously making large changes to the angle of their bill (or other elements of the vocal tract) and making these changes rapidly (Podos and Nowicki 2004b), and this limitation produces a tradeoff between frequency bandwidth and trill rate.

Podos's (1997) comparative analysis of songs of the Emberizid family of sparrows provided the first evidence for this tradeoff. Podos found that as trill rates of songs increased, the maximal values of frequency bandwidth decreased. When songs are graphed by trill rate against frequency bandwidth, the result is a triangular distribution of songs in acoustic space. An upper bound regression can be used to estimate the performance limit implied by this triangular distribution, and the "vocal deviation" of a song from the regression line can in turn be used to estimate a song's level of performance (Podos 2001). A vocal performance tradeoff measured this way has since been demonstrated for an independent sample of swamp sparrow songs (Ballentine et al. 2004) and for the songs of species outside the Emberizidae, including domesticated canaries (*Serinus canaria*, Draganoiu et al. 2002), yellow warblers (*Dendroica petechia*, Beebee 2004), banded wrens (*Thryothorus pleurostictus*, Illes et al. 2006), and red-winged blackbirds (*Agaelius phoeniceus*, Cramer and Price 2007).

In swamp sparrows, maximum vocal performance correlates positively with male age and size (i.e., age and size are negatively correlated with vocal deviation). In a given breeding season, males with higher performance (lower deviation) are older and larger, and males tend to improve their maximum vocal performance between their first and second breeding years (Ballentine 2009). Vocal performance affects female response to song, with females responding preferentially to higher performance songs (Ballentine et al. 2004). Thus, females who choose a high-performance singer will likely be pairing with a higher quality male, to the extent that age and size correlate with a male's ability to provide direct and/or indirect benefits to their mates. Signals of size and age are likely also to be important in male–male signaling because older, larger males tend to win encounters with smaller, younger males in many songbird species (Koivula et al. 1993; Richner 1989; Sandell and Smith 1991; Searcy 1979; Yamaguchi and Kawano 2001). Therefore, vocal performance may be used as an assessment signal in male–male aggressive communication.

One criterion to assess whether a signal is used in male–male aggressive communication is to test males for an aggressive response to the signal (Searcy and Beecher 2009). Illes et al. (2006) found that vocal deviation affected aggressive response in male banded wrens. Males presented with both a high- and a low-performance stimulus tended to approach the high-performance stimulus first, though they spent less time overall in close proximity to the high-performance stimulus than to the low-performance stimulus. In a similar test using red-winged blackbirds, Cramer and Price (2007) also found that males discriminated vocal performance differences, responding to lower performance songs with higher song and flight rates, and by spending more time close to the low-performance stimulus. One objective of the present study was to test the hypothesis that male swamp sparrows assess individual differences in vocal performance. That hypothesis predicts that males behaviorally discriminate between two individuals based on typical magnitudes of among-individual differences observed in the population.

Recently, DuBois et al. (2009) addressed the issue from the signaler's perspective and found that male swamp sparrows actively increase vocal performance in aggressive situations. When males used the same song type in both an aggressive context and a more neutral one, they sang that song type with a higher performance level (i.e., a lower vocal deviation as measured from the upper bound regression reference line) in the aggressive context. This change was accomplished by increasing both the trill rate and frequency bandwidth of the song. Similarly, Cardoso et al. (2009) found that dark-eyed juncos' (*Junco hyemalis*) increased vocal performance during aggressive singing, although this increase was primarily the result of choice of song type. These examples of individual modulation could be interpreted as "cheating" on an index signal.

Changes in vocal performance produced by individual modulation, however, tend to be small relative to the performance differences seen between males, as well as performance differences between some song types within males. It is important to note here that although it is clear that there are important differences between males in the performance of specific song types, it has not been shown that males differ consistently in performance across their entire repertoires. In testing female swamp sparrows for response to songs differing in vocal performance, Ballentine et al. (2004) used pairs of songs of the same song type obtained from different males, and differing from each other by 5–18 units of vocal deviation (as measured from Podos's 1997 Emberizid line, with vocal deviation derived as the orthogonal distance to the reference line). In contrast, individual males in the DuBois et al. (2009) study were capable of modulating their vocal deviation to a maximum of only five deviation units, with most males falling well short of this maximum. Thus, male swamp sparrows might be able to detect differences in vocal performance on the order shown between males without being able to discriminate the smaller differences produced by within-male modulation. Accordingly, in this study, we not only tested for male discrimination of "between-male" differences in performance but also tested the hypothesis that within-individual modulation in vocal performance functions as an acoustic bluff. That hypothesis predicts that males would behaviorally discriminate between low- and high-performance song renditions sung by the same individual.

Methods

Experiments were conducted between 24 April and 9 June 2009 on a population of territorial male swamp sparrows in Conneaut Marsh, Crawford County, PA, USA. Territories remain stable during the height of the breeding season, so territorial males can be identified by location; a subset of males ($N=4$) was color banded as well and none moved during the testing period. Of the total 34 males tested, 11 males were tested in all three experiments, 12 in two experiments, and 11 in one experiment each. No individual was presented with the same stimulus set in more than one experiment.

Experiment 1: response to between-male differences in vocal performance

Experiment 1 investigated whether males respond differentially to between-male differences in vocal performance (inter-male discrimination). We employed a single speaker, paired design modified from Cramer and Price (2007). Subjects were presented with stimulus pairings consisting of high- and low-vocal performance exemplars of the same song type, sung by different males (Fig. 1). The stimuli used in experiment 1 were the same 20 low- and high-performance stimulus pairings employed in a study of female preference conducted by Ballentine et al. (2004). These stimuli included 14 unique song types. When the

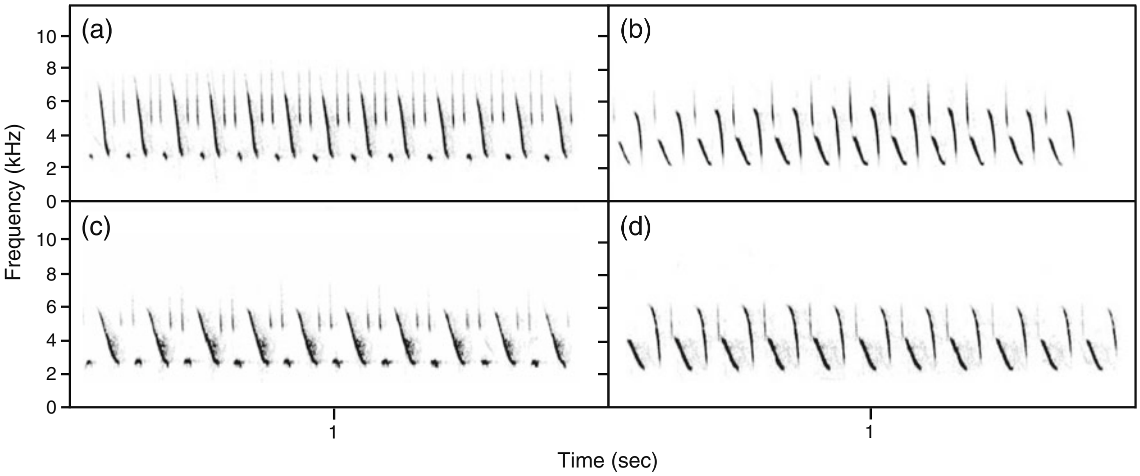

Fig. 1 Two examples of stimulus pairings used in the between-male discrimination trials (experiment 1). Spectrograms **a** and **c** constitute one pairing of the same song type sung by two different males, while **b** and **d** are a different pairing. Songs **a** and **b** represent the higher performance (low deviation) renditions of the song type, while songs **c** and **d** represent the lower performance renditions

same song type was used more than once, the multiple sets of stimuli were taken from different pairs of males. The mean difference in vocal deviation was 9.9 (range, 5.4–18.1).

Each male subject received a high-performance trial and a low-performance trial, separated by 48 h. Stimuli were presented in an alternate random fashion—random for the first male, reverse for the second, random for the third, and so forth.

Playback stimuli were broadcast as WAV files at approximately 89 dB (measured at 1 m with a BK Precision 32A sound level meter) using an iPod mini attached to an Advent Powered Partners AV570 speaker (35 W) mounted on a 1-m post. Each trial consisted of 6 min of pre-playback observation and recording, followed by observation and recording for 3 min during playback and another 3 min after playback. Playback songs were repeated at a rate of six songs/minute. Subject songs, as well as spoken narration of behavior and location by an observer (ALD), were recorded using a Marantz 660 solid-state digital recorder and a RadioShack 33-3001 directional microphone mounted in a Sony Parabolic Reflector-330. Behaviors recorded were as follows: the number of broadcast songs sung by the subject, number of soft songs (low-amplitude songs associated with aggression, Ballentine et al. 2008), number of flights, and number of wing waves (the visual display most closely associated with attack, Ballentine et al. 2008). The observer also estimated the subject's distance from the speaker at all times (henceforth referred to as approach distance). Flagging markers set at 2, 4, and 8 m in either direction from the speaker facilitated estimation of approach distance. Distance estimates were recorded on flow sheets divided into 10-s intervals, from which mean approach distance for each observation period (pre-playback, during playback, and post-playback) was calculated. Males were classified as "within 2 m" (in which case 1 m was used in approach distance calculations for that 10-s period), "within 4 m" (3 m in calculations), "within 8 m" (6 m in calculations), "within 16 m" (12 m in calculations), and "greater than 16 m" (24 m in calculations). To control for variation in a male's behavior on separate playback days, we calculated the rates of behaviors (per min) for each observation period (pre-, during, and post-playback). We subtracted the rates of each behavior during the pre-playback period from the mean rate of that behavior during and post-playback. Thus, all values used for analysis represent the change in behavior once playback began (Cramer and Price 2007).

To compare subject responses to low- and high-performance playback in both experiments, we used a principal components analysis (PCA) with changes in broadcast song rate, soft song rate, flight rate, wing waving rate, and approach distance as variables. We used Wilcoxon signed-rank tests to compare resultant PCA scores, as well as the subject's behavioral rates in response to stimuli. Twenty-six males were presented with both high- and low-performance playback. Behavioral responses of subjects who were tested with playback of the same stimuli were averaged with the response of the other subject to avoid pseudoreplication so that the sample size in the statistical tests is number of stimulus sets (20) rather than number of subjects (26).

Experiment 2: response to within-male differences in vocal performance, 1-speaker trials

Experiment 2 investigated whether males respond differentially to two levels of performance modulation of a single simulated rival (intra-male discrimination). We followed the same experimental protocol as in experiment 1, but replaced the stimuli representing inter-male differences in vocal deviation with digitally manipulated stimuli that simulated intra-male modulation of vocal performance. High and low vocal performance stimuli were artificially created using SIGNAL v. 4.0 software (Engineering Design, Belmont, MA, USA) to mimic a natural level of performance modulation by a single male (mean difference in performance=4.19, range=3.02–6.24).

We chose to use digitally manipulated stimuli in experiment 2 (rather than natural recordings as in experiment 1) because we did not have a sufficient number of natural stimulus pairs recorded during the DuBois et al. (2009) study that met our criteria for this discrimination study. Eighteen males in the DuBois et al. (2009) study increased their vocal performance in aggressive contexts, with the change in vocal deviation ranging from 0.45 to 5.11. We wished to present vocal performance differences from the upper end of this range. By digitally manipulating songs instead of using natural song pairs, we were able to produce pairs of stimuli identical except for the vocal performance differences while achieving performance differences in the desired range.

We created 16 stimulus sets, consisting of a base song (low performance) and a test song (high performance) (Fig. 2). These stimulus sets included 11 unique song types. When the same song type was used more than once, song exemplars were recorded from different males. To create a stimulus set, we began with a natural song exemplar. Exemplars were chosen based on having a broad frequency bandwidth and high recording quality. For each song exemplar, we measured the average syllable length, the average length of time between syllables (average interval), the length of time between notes in the syllable (inter-note intervals), and the maximum frequency of each syllable. To create a base song, we chose a single syllable from the natural song and digitally inserted silent inter-note intervals of approximately equal length to those of the natural song and then concatenated this control manipulated syllable to create a song of the same trill rate and duration as the natural song. To create a high-performance (test) song, we followed the same methods as in the creation of the base song, with two modifications. First, we exchanged the broadest bandwidth note with a corresponding note obtained from a different exemplar of the same syllable type, but one having a broader frequency bandwidth. We then concatenated this manipulated syllable to create a full song, but did so at an increased trill rate from the original. The vocal deviation of each newly created song was measured following the methods of Ballentine et al. (2004), using the upper bound regression line calculated for the Emberizidae by Podos (1997) as a reference. The difference in vocal deviation between the low-performance (base) and high-performance (test) songs mimicked a natural range of vocal deviation modulation found by DuBois et al. (2009) (mean difference=4.19, range=3.02–6.24). As before, the behavioral responses of males who were tested with the same stimulus set were averaged to avoid pseudoreplication so that the sample size in statistical tests is number of stimulus sets (16) instead of number of subjects (26).

Experiment 3: response to within-male differences in vocal performance, 2-speaker trials

We performed a second test for behavioral discrimination of within-male vocal performance differences, using a two-speaker discrimination design. This design may provide a more sensitive test of discrimination (Searcy et al. 1995, 1999), at the cost of limiting the number of clearly aggressive response measures that can be associated with each stimulus to one: approach distance. In this design, we placed two matched speakers (Advent Powered Partners AV570) mounted on 1-m metal posts on the subject's territory, 12 m apart. We placed flagging at 3 and 6 m from each speaker in either direction to facilitate estimating the subject's distance from either speaker at all times. We also incorporated a control song for each stimulus set (Fig. 2). To create a control song, we followed the same protocol as above, but exchanged the note responsible for the maximum frequency of the syllable with the corresponding note in a different syllable with similar or equal maximum frequency to the maximum frequency of the base song. This substitution served as a control for the modifications made to the high-performance song. We also kept the trill rate equal to that of the base song. Base and control song pairs were nearly identical with respect to vocal performance (mean difference=0.17, range=0.00–0.44). During an initial phase, the base song type was played from both speakers for 42 min at a rate of six songs/minute, with the second speaker offset by 5 s to prevent song overlap. After this initial phase, both speakers switched to new stimuli, and playback continued for another 6 min. One speaker switched to the low-performance (control) song, while the other speaker switched to the high-performance (test) song, simulating an increase in vocal performance. The order and location of the high-performance song stimulus were balanced across subjects. The territory owner was thus asked to judge which stimulus was more dissimilar to the base song type by approaching one speaker more closely than the other. This design was used successfully to demonstrate stronger discrimination between song types than between song variants in song sparrows (*Melospiza melodia*, Searcy et al. 1999).

During playback, an observer (WAS) noted the distance of the subject from both speakers at all times. This information was recorded in the field on time flow sheets

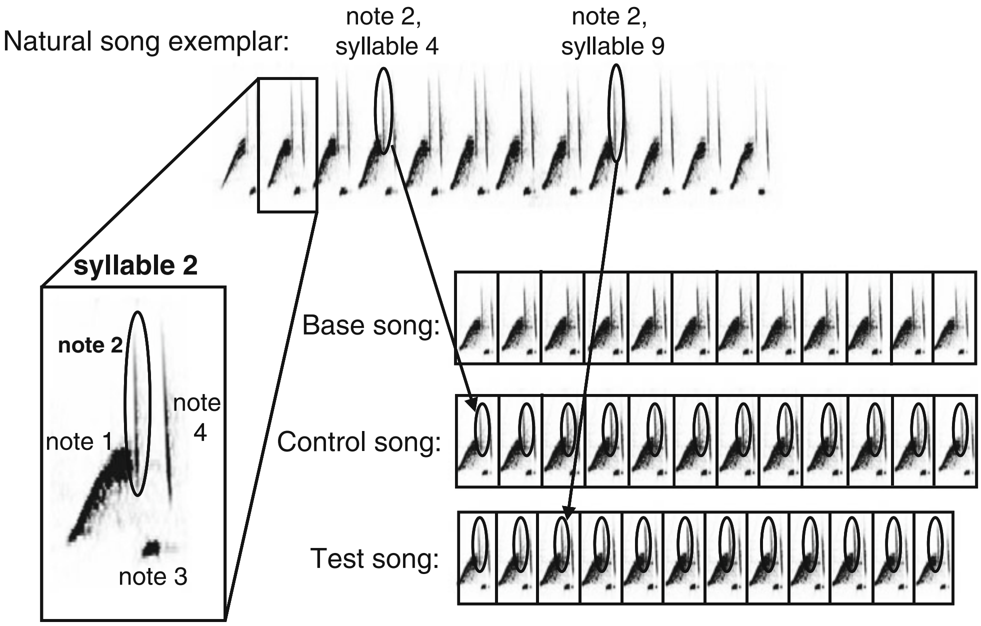

Fig. 2 Creation of intra-male "modulated" stimuli. To form a base song, a single syllable was chosen from a natural song exemplar. Notes from that syllable were saved into individual time buffers and reassembled with internote and intersyllable silent intervals equal to the natural song (resulting in equal trill rates). To form the control song, the note responsible for the maximum frequency (in this example, "note 2") of the song was substituted with a note from a different syllable with equivalent frequency bandwidth. Internote and intersyllable silent intervals were again equal to the natural and base song types. To form the test song, "note 2" was substituted with a note with a broader frequency bandwidth, and intersyllable silent intervals were reduced to increase the trill rate of the song

by a second observer (ALD), as in the previous experiments. We used the subject's mean distance to the speaker in the last 6 min before the switch minus his average distance in the 6 min after the switch as a measure of change in approach distance, and compared this measure toward the low- and high-performance speakers using a Wilcoxon signed-rank test to determine whether subjects discriminated modulation in vocal performance.

Results

Experiment 1: between-male discrimination

Principal component analysis reduced five univariate measures of response to two principal components (Table 1). Song rate and wing waving rate load highly on PC1, which explains 36.2% of the variance. Soft song rate and approach distance load highly on PC2, (approach distance is negatively correlated, as greater changes in approach are more negative values), which explains 29.4% of the variance. PC1 scores for subjects were significantly greater in response to high-performance song (Wilcoxon signed-rank test, $N=20$, $Z=-2.58$, $p=0.01$, Fig. 3a), as were PC2 scores ($Z=-2.17$, $p=0.03$, Fig. 3a).

Responses were greater toward high-performance song on all five univariate measures, and the differences were significant for three of these: song rates ($N=20$: $Z=-2.58$, $p=0.01$, Fig. 3b), soft song rates ($Z=-2.06$, $p=0.039$, Fig. 3b), and flight rates ($Z=-2.44$, $p=0.015$, Fig. 3b). Differences were not significant for wing waving rates ($Z=-1.22$, $p=0.22$, Fig. 3b) or approach distance ($Z=-1.87$, $p=0.062$, Fig. 3c).

Experiment 2: within-male discrimination, one-speaker

Principal component analysis again revealed two principal components (Table 2). PC1 was most associated with wing waving rate and explained 37.8% of the variance in response measures, while PC2 was associated with approach distance and explained an additional 29.6%. Comparison of PC1

Table 1 Loadings of response measures on the first two principal components in the between-male discrimination experiment (experiment 1)

	Loadings	
Response measures	PC1	PC2
Song rate	0.866	−0.028
Soft song rate	0.400	0.729
Flight rate	0.474	0.313
Wing waving rate	0.820	0.143
Approach distance	0.066	−0.905

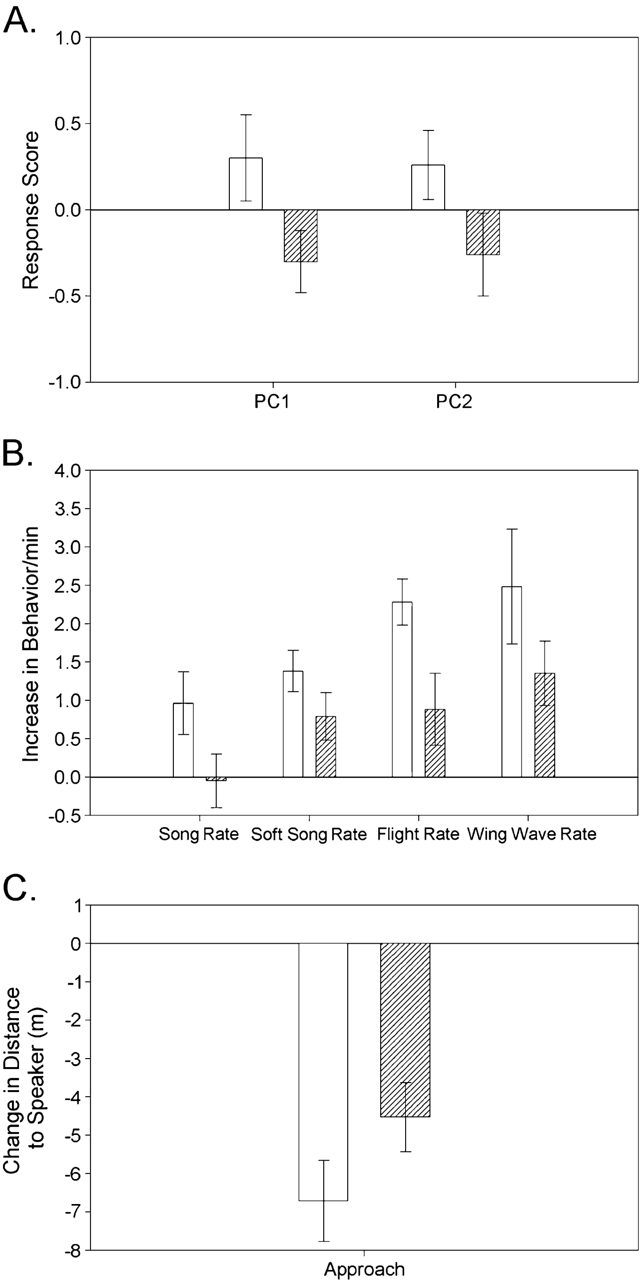

Fig. 3 Response to between-male differences in vocal performance. **a** Mean principal component scores (±SE) in response to high-performance (*white*) and low-performance (*striped*) playback song ($N=20$); **b** mean (±SE) behavioral responses to playback of high- and low-performance song playback. Values are increases in behavioral rates during and after playback, relative to behavioral rates prior to playback; **c** mean (±SE) change in approach distance. Lower values represent a closer approach to the speaker during and after playback, relative to approach distance prior to playback

scores revealed no significant difference between responses to the two stimulus categories ($N=16$; $Z=-0.931$, $p=0.352$, Fig. 4a). Comparison of PC2 scores also revealed no difference ($Z=-1.45$, $p=0.148$, Fig. 4a).

Table 2 Loadings of response measures on the first two principal components in the within-male, one-speaker discrimination experiment (experiment 2)

	Loadings	
Response measures	PC1	PC2
Song rate	0.523	0.577
Soft song rate	0.798	0.063
Flight rate	0.380	0.699
Wing waving rate	0.901	0.092
Approach distance	0.155	−0.803

Male responses to low- and high-performance manipulations of song types did not differ with respect to song rate ($N=16$, $Z=-0.691$, $p=0.489$, Fig. 4b), soft song rate ($Z=-0.525$, $p=0.60$, Fig. 4b), wing waving rate ($Z=-1.22$, $p=0.221$, Fig. 4b), or approach distance ($Z=-0.982$, $p=0.326$, Fig. 4c). Flight rate was significantly higher in response to the higher performance song ($Z=-1.989$, $p=0.047$, Fig. 4b).

Experiment 3: within-male discrimination, two-speaker

The initial playback phase of 48 min was divided evenly into 14 3-min time blocks, followed by two 3-min time blocks post-switch. Approach distance toward the two speakers differed only during one of the initial 14 time blocks (block 11; $Z=-2.05$, $p=0.041$) but was not significantly different during time blocks 1–10 or 12–14 (Fig. 5). Thus, subjects ($N=16$) did not show a bias toward either speaker. After the switch to new stimuli, approach was slightly closer to the low-performance songs, though the differences were not significant (time block 15, $Z=-0.672$, $p=0.501$; time block 16, $Z=-1.704$, $p=0.088$, Fig. 5). There was no difference in approach between the last 6 min pre-switch and first 6 min post-switch toward the speaker switching to a higher performance song versus the speaker switching to a control performance level song ($Z=1.60$, $p=.109$, $df=15$).

Discussion

Male swamp sparrows discriminated between song exemplars of a single song type, which differed in vocal performance, when the differences in performance were on the level exhibited among different individuals. This result supports our first hypothesis that males assess individual differences in vocal performance. The difference in PC1 and PC2 scores in the between-male experiment is particularly telling because PC1 was strongly associated with wing waving rate, while PC2 combined soft song and

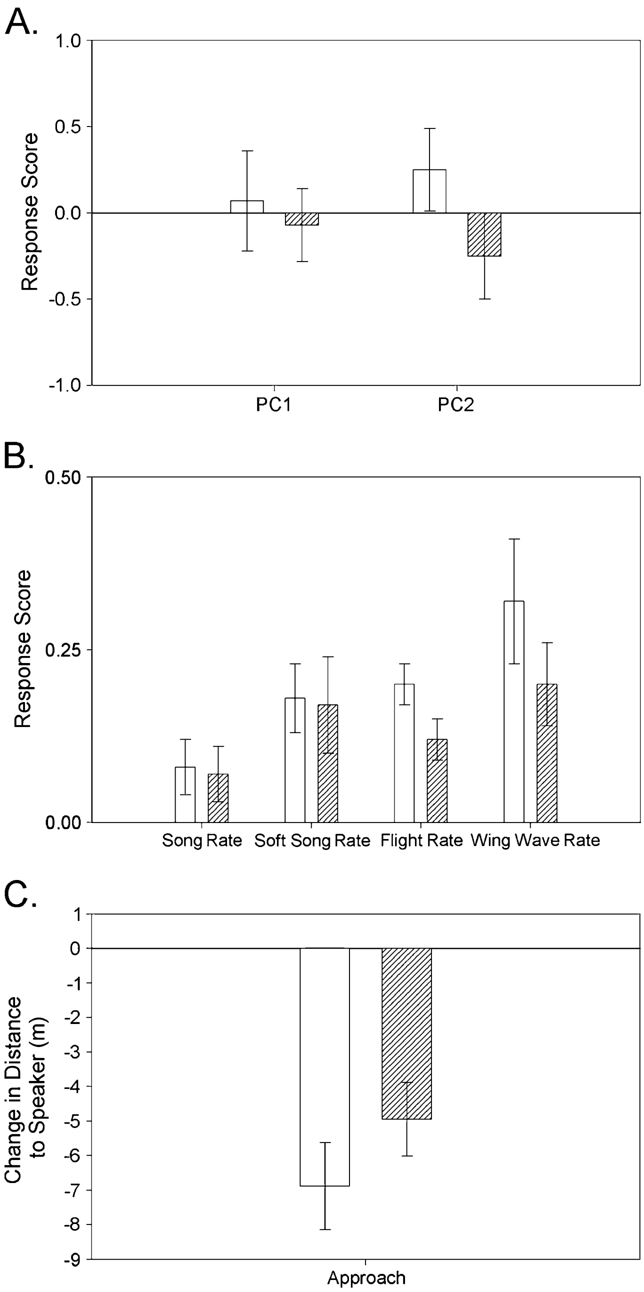

Fig. 4 Response to within-male differences in vocal performance. **a** Mean principal component scores (±SE) in response to higher performance (*white*) and lower performance (*striped*) song stimuli (N=16); **b** mean (±SE) behavioral responses to playback of higher performance and lower performance stimuli. Values are increases in behavioral rates during and after playback, relative to behavioral rates prior to playback; **c** mean (±SE) change in approach distance. Lower values represent a closer approach to the speaker during and after playback, relative to approach distance prior to playback

approach, and these three measures have been shown to predict attack in swamp sparrows (Ballentine et al. 2008). We are thus safe in concluding that aggressive response in this first experiment was stronger toward higher performance songs. Differential response of males to songs differing in vocal performance, in swamp sparrows as well as in other species (Cramer and Price 2007, Illes et al. 2006), suggests that receivers recognize vocal performance as a signal in male–male communication.

Distinguishing potential rivals based on vocal performance may allow males to assess an individual's quality as a competitor, in the same way that females are able to assess a male's quality as a potential mate using vocal performance (Ballentine et al. 2004, Ballentine 2009). While females presumably attend to vocal performance as an indicator of potential direct and indirect benefits, males may be attending to vocal performance as an indicator of potential fighting ability. Maximum vocal performance is correlated with age and size in swamp sparrows, and these attributes have been associated with fighting ability and the ability to win contests against intraspecific competitors in other songbird species (Koivula et al. 1993, Richner 1989, Sandell and Smith 1991, Searcy 1979, Yamaguchi and Kawano 2001). Thus, it is possible that males can garner information about an opponent based on his vocal performance, whether or not any information is conveyed through modulation of this characteristic.

Although we found convincing evidence of discrimination in between-male performance, subjects did not respond differentially to vocal performance differences at the within-male level. Thus, the results failed to support our second hypothesis that within-individual modulation in vocal performance functions as an acoustic bluff. The only behavioral variable to differ significantly between low- and high-performance trials in experiment 2 was flight rate.

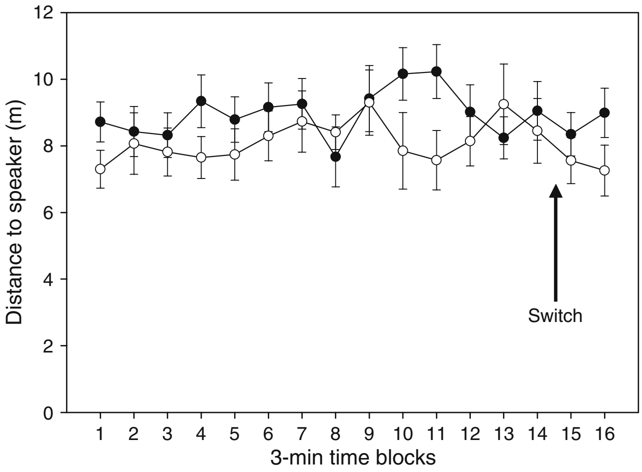

Fig. 5 Mean approach distance (±SE) to speakers during the two-speaker intra-male discrimination test. One speaker (*white circle*) switched from a base song type during the habituation phase (time blocks 1–14) to a low-performance (control) song playback during the recovery phase (time blocks 15–16). The other speaker (*black circle*) switched to a high-performance (test) song playback. *Arrow* indicates point at which playback switched from the base song to either the control or test song (N=16)

Increased flight rate suggests that males were perhaps slightly more agitated, but flight rate is not likely to be an important aggressive behavior, and none of the other behavioral variables more closely associated with aggression (soft song, wing waving, and approach) differed between trials. The two-speaker assay (experiment 3) also failed to demonstrate discrimination. This method has been used successfully to detect discrimination of other fine scale differences in song variation in a closely related species, the song sparrow (Searcy et al. 1999), but it failed to detect any discrimination of within-male differences in vocal performance in swamp sparrows.

The exhibition of differential responses to inter-male differences and an absence of response differences for intra-male modulation found in this study are consistent with the suggestion that vocal performance serves as an index signal. Several species that use index signals are able to modulate these signals (Searcy and Nowicki 2005), suggesting that index signals are only loosely constrained within limits, and that the signal may be modulated adaptively within those limits. For example, Wagner (1989a) found that male Blanchard's cricket frogs (*Acris crepitans blanchardi*) lower the dominant frequency of their calls during aggressive interactions. Importantly, the level of modulation of dominant frequency in cricket frogs has been shown to be functionally significant in terms of its effect on receiver response. Male receivers tend to retreat from playback of calls that decrease in dominant frequency more so than from playback calls of constant or increasing frequency (Wagner 1992). This alteration implies that dominant frequency is a signal that can be effectively cheated. In the same vein, red deer stags are capable of modifying the formant dispersion of their roars by actively altering the configuration of their vocal tracts (Fitch and Reby 2001, Reby et al. 2005). Changes in formant dispersion occur at a level salient to receivers, at least in the case of female receivers (Charlton et al. 2007), and are therefore likely to influence the outcome of aggressive interactions (Reby et al. 2005). In contrast to both these examples, swamp sparrow males are able to exaggerate their signal by lowering vocal deviation (DuBois et al. 2009), but they do not alter the signal sufficiently to affect receiver response, meaning that modulation is unlikely to affect the outcome of an interaction. Swamp sparrows, therefore, may be more strictly constrained in their ability to modulate vocal deviation than are signalers in these other systems. Vocal deviation perhaps provides a better example of an index signal, since it cannot be effectively cheated.

Acknowledgments The authors would like to thank Barbara Ballentine for providing stimuli for the inter-male discrimination experiment, Susan Peters for her gracious and invaluable help in creating artificial stimuli for the intra-male discrimination experiment, and Kennon Todd for assistance in the field during pilot work for this study. We would like to thank two anonymous reviewers for helpful suggestions in revising the manuscript. We would also like to thank the Pymatuning Laboratory of Ecology for logistical support and the Pennsylvania Game Commission for access to study sites. This work was funded by the University of Miami College of Arts and Sciences and the University of Pittsburgh's Arthur and Barbara Pape Endowment.

Ethical standards The experimental methods for this study complied with current rules and regulations within the USA.

Conflicts of interest The authors declare that they have no conflict of interest.

References

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

Ballentine B, Searcy WA, Nowicki S (2008) Reliable aggressive signaling in swamp sparrows. Anim Behav 75:693–703

Bee MA, Perrill SA (1996) Responses to conspecific advertisement calls in the green frog (*Rana clamitans*) and their role in male-male communication. Behav 133:283–301

Bee MA, Perrill SA, Owen PC (1999) Size assessment in simulated territorial encounters between male green frogs (*Rana clamitans*). Behav Ecol Sociobiol 45:177–184

Beebee MD (2004) Variation in vocal performance in the songs of a wood-warbler: evidence for the function of distinct singing modes. Ethol 110:531–542

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, 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

Charlton BD, Reby D, McComb K (2007) Female perception of size-related formant shifts in red deer, *Cervus elaphus*. Behav 74:707–714

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

Davies NB, Halliday TR (1978) Deep croaks and fighting assessment in toads *Bufo bufo*. Nature 274:683–685

Dawkins R, Krebs JR (1978) Animal signals: information or manipulation? In: Krebs JR, Davies NB (eds) Behavioural ecology. Blackwell, Oxford, pp 282–309

Draganoiu TI, Nagle L, Kreutzer M (2002) Directional female preference for an exaggerated male trait in canary (*Serinus canaria*) song. Proc R Soc Lond B 269:2525–2531

DuBois AL, Nowicki S, Searcy WA (2009) Swamp sparrows modulate vocal performance in an aggressive context. Biol Lett 5:163–165

Fitch WT (1997) Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques. J Acoust Soc Am 102:1213–1222

Fitch WT, Reby D (2001) The descended larynx is not uniquely human. Proc R Soc Lond B 268:1669–1675

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

Grafen A (1990) Biological signals as handicaps. J Theor Biol 144:517–546

Hoese WJ, Podos J, Boetticher NC, Nowicki S (2000) Vocal tract function in birdsong production: experimental manipulation of beak movements. J Exp Biol 203:1845–1855

Holveck MJ, Riebel K (2007) Preferred songs predict preferred males: consistency and repeatability of zebra finch females across three test contexts. Behav 74:297–309

Hurd PL, Enquist M (2005) A strategic taxonomy of biological communication. Behav 70:1155–1170

Illes AE, Hall MH, Vehrencamp S (2006) Vocal performance influences male receiver response in the banded wren. Proc R Soc Lond B 273:1907–1912

Koivula K, Lahti K, Orell M, Rytkönen S (1993) Prior residency as a key determinant of social dominance in the willow tit (*Parus montanus*). Behav Ecol Sociobiol 33:283–287

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

Martin WF (1972) Evolution of vocalization in the genus *Bufo*. In: Blair WF (ed) Evolution in the Genus *Bufo*. University of Texas, Austin, pp 279–309

Maynard Smith J, Harper D (2003) Animal signals. New York, Oxford

Nowicki S (1987) Vocal tract resonances in oscine bird sound production: evidence from birdsongs in a helium atmosphere. Nature 325:53–55

Nowicki S, Marler P (1988) How do birds sing? Music Percept 5:391–426

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, Nowicki S (2004a) Beaks, adaptation, and vocal evolution in Darwin's finches. BioSci 54:501–510

Podos J, Nowicki S (2004b) Performance limits on birdsong. In: Marler P, Slabbekorn H (eds) Nature's music. Elsevier, San Diego, pp 318–342

Reby D, McComb K (2003) Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags. Anim Behav 65:519–530

Reby D, McComb K, Cargnelutti B, Darwin C, Fitch WT, Clutton-Brock TH (2005) Red deer stags use formants as assessment cues during intrasexual agonistic interactions. Proc R Soc Lond B 272:941–947

Richner H (1989) Phenotypic correlates of dominance in carrion crows and their effects on access to food. Anim Behav 38:606–612

Ryan MJ (1985) The Tungara frog: a study in sexual selection and communication. University of Chicago, Chicago

Sandell M, Smith HG (1991) Dominance, prior occupancy, and winter residency in the great tit (Parus major). Behav Ecol Sociobiol 29:147–152

Searcy WA (1979) Morphological correlates of dominance in captive male red-winged blackbirds. Condor 81:417–420

Searcy WA, Beecher MD (2009) Song as an aggressive signal in songbirds. Anim Behav 78:1281–1292

Searcy WA, Nowicki S (2005) The evolution of animal communication: reliability and deception in signaling systems. Princeton University Press, Princeton

Searcy WA, Podos J, Peters S, Nowicki S (1995) Discrimination of song types and variants in song sparrows. Anim Behav 49:1219–1226

Searcy WA, Nowicki S, Peters S (1999) Song types as fundamental units in vocal repertoires. Anim Behav 58:37–44

Vehrencamp S (2000) Handicap, index, and conventional signal elements of bird song. In: Espmark Y, Amundsen T, Rosenqvist G (eds) Signalling and signal design in animal communication. Tapir, Trondheim, pp 301–315

Wagner WE (1989a) Fighting, assessment, and frequency alteration in Blanchard's cricket frog. Behav Ecol Sociobiol 25:429–436

Wagner WE (1989b) Social correlates of variation in male calling behavior in Blanchard's cricket frog, *Acris crepitans blanchardi*. Ethol 82:27–45

Wagner W (1992) Deceptive or honest signaling of fighting ability? A test of alternative hypotheses for the function of changes in call dominant frequency by male cricket frogs. Anim Behav 44:449–462

Westneat MW, Long JH, Hoese W, Nowicki S (1993) Kinematics of birdsong: functional correlation of cranial movements and acoustic features in sparrows. J Exp Biol 182:141–171

Yamaguchi N, Kawano KK (2001) Effect of body size on the resource holding potential of male varied tits *Parus varius*. Jpn J Ornithol 50:65–70