



Song is not a reliable signal of general cognitive ability in a songbird

Adrienne L. DuBois^a, Stephen Nowicki^{b, c}, Susan Peters^b, Karla D. Rivera-Cáceres^a, William A. Searcy^{a, *}

^a Department of Biology, University of Miami, Coral Gables, FL, USA

^b Department of Biology, Duke University, Durham, NC, USA

^c Department of Neurobiology, Duke University Medical School, Durham, NC, USA

ARTICLE INFO

Article history:

Received 18 August 2017

Initial acceptance 4 October 2017

Final acceptance 11 December 2017

MS. number: A17-00668

Keywords:

animal communication

assessment signal

birdsong

cognition

swamp sparrow

Learned aspects of song affect female mating preferences in a number of species of songbirds, including swamp sparrows, *Melospiza georgiana*. One explanation for why female songbirds attend to such song features is that these song attributes convey information on the general cognitive ability of singers. The fact that song attributes and cognitive ability are affected during development by the same stressors makes a connection between the two plausible. Here we test the hypothesis that song is a signal of cognitive ability by relating five measures of song quality to five measures of cognitive performance in 49 captive male swamp sparrows. The five song measures are repertoire size, mean and minimum vocal deviation (measures of vocal performance), and mean and maximum typicality (measures of song learning). Cognitive performance was measured as the speed with which five cognitive tasks were mastered: a novel foraging task, a colour association, a colour reversal, a spatial learning problem and a detour-reaching test. In general linear mixed models controlling for neophobia, none of the song measures were predictive of any of the cognitive performance measures. Thus the results do not support the hypothesis that song attributes signal general cognition in swamp sparrows.

© 2018 The Association for the Study of Animal Behaviour. Published by Elsevier Ltd. All rights reserved.

Humans across a wide range of cultures express a preference for more intelligent over less intelligent mates (Marlowe, 2004; Prokosch, Coss, Scheib, & Blozis, 2009; Shackleford, Schmitt, & Buss, 2005; Souza, Conroy-Beam, & Buss, 2016). This preference is often present in both sexes, but is particularly pronounced in females (Furnham, 2009; Li, Bailey, Kenrick, & Linsenmeier, 2002; Shackleford et al., 2005). Female preferences for more intelligent males have also been found in other animals, including fish (Shohet & Watt, 2009), rodents (Spritzer, Meikle, & Solomon, 2005) and birds (Keagy, Savard, & Borgia, 2009, 2011). Human females may use cues such as vocabulary size (Rosenberg & Tunney, 2008) or oral reading ability (Borkenau, Mauer, Riemann, Spinath, & Angleitner, 2004; Prokosch et al., 2009) to assess male intelligence, but how females in nonhuman animals might assess the intelligence of prospective mates is more mysterious. Presumably, many male display traits important to female choice are unlikely to contain information on male cognitive ability, such as (in birds) the length of tails (Andersson, 1982; Møller, 1988), the brightness of plumage (Hill, 1991; Whittingham & Dunn, 2016), or the colour of

bills (Simons & Verhulst, 2011). Learned displays constitute a possible exception, where a link between a display and intelligence seems much more likely (Peters, Searcy, & Nowicki, 2014). Birdsong is highly dependent on learning (Beecher & Brenowitz, 2005; Marler & Mundinger, 1971) and often affects female choice (Searcy & Andersson, 1986; Searcy & Yasukawa, 1996), and thus seems particularly likely to provide information on cognitive ability (Boogert, Giraldeau, & Lefebvre, 2008; Boogert, Anderson, Peters, Searcy, & Nowicki, 2011). Here we test the hypothesis that song signals cognitive ability in swamp sparrows, *Melospiza georgiana*, by relating several measures of cognitive performance to a recently developed measure of success in song learning (Lachlan, Anderson, Peters, Searcy, & Nowicki, 2014).

Previous studies of the relationship between song and cognition in songbirds have focused largely on what have been considered to be measures of song complexity, such as song duration, number of elements per song and number of unique elements per song in zebra finches, *Taeniopygia guttata* (Boogert et al., 2008; Templeton, Laland, & Boogert, 2014), song bout length in European starlings, *Sturnus vulgaris* (Farrell, Weaver, An, & MacDougall-Shackleton, 2012), and song repertoire size in song sparrows, *Melospiza melodia* (Boogert et al., 2011; Sewall, Soha, Peters, & Nowicki, 2013). If we equate song complexity with diversity of song types or song

* Correspondence: W. A. Searcy, Department of Biology, University of Miami, Coral Gables, FL 33124, USA.

E-mail address: wsearcy@miami.edu (W. A. Searcy).

elements, some of these metrics seem more appropriate as measures of complexity (number of unique elements per song, song repertoire size) than do others (song duration), but note that in European starlings song bout length is strongly correlated with song phrase repertoire size (Eens, 1997). These complexity measures are a logical focus for cognitive studies, in that one or more of them have been associated with female mating preferences in each of the above species (zebra finches: Riebel, 2009; Vyas, Harding, Borg, & Bogdan, 2009; starlings: Eens, Pinxten, & Verheyen, 1991; song sparrows: Reid et al., 2004; Searcy, 1984). Song complexity, however, is ambiguous as a measure of learning success. On the one hand, individual song elements and song types are definitely learned from models in these songbirds (Eales, 1985; Eens, Pinxten, & Verheyen, 1992; Marler & Peters, 1987; Nordby, Campbell, & Beecher, 1999), so measures such as number of unique elements per song or song repertoire size can be interpreted as measuring the amount of material that has been learned. On the other hand, in some songbird species (including swamp sparrows) individual males do not include in their adult repertoires all the songs that they have memorized (Marler & Peters, 1981; Nelson, 2000; Nordby, Campbell, & Beecher, 2007; Peters & Nowicki, 2017; Prather, Peters, Nowicki, & Mooney, 2010), so measures of adult song complexity may not accurately reflect the amount of learning.

Another approach to assessing song learning is to measure the accuracy with which adult birds reproduce the songs they are known to have heard early in life during the period when song models are memorized. This approach is much easier to implement with laboratory-reared birds than with free-living ones, because the set of model songs can be controlled experimentally in the laboratory, whereas for free-living birds that set may be both very large and very difficult to delimit. We have shown that laboratory-reared song sparrows differ in how accurately they copy model songs recorded locally, and that adult females from the same local population prefer well-copied songs to poorly copied ones (Nowicki, Searcy, & Peters, 2002b). We have also assessed the association of cognitive measures to success in copying model songs in hand-reared song sparrows (Anderson et al., 2017). This approach produces convincing measures of song learning, but sample sizes are typically limited because of the logistical difficulty of hand-rearing wild birds.

Lachlan, Verhagen, Peters, and ten Cate (2010) and Lachlan, Anderson, Peters, Searcy, & Nowicki (2014) suggested an alternative method for measuring accuracy of song learning that can be applied to free-living songbirds, so that quality of learning can be assessed in a larger sample of males than is practicable in hand-rearing studies and without the potential concern that learning in a laboratory setting differs from learning in the field. Rather than measuring how well a subject's song matches a specific song model, the method measures how well the subject's song matches the average features defining a category of songs. The method is therefore best applied to species of songbirds in which populations produce a limited set of song categories, or 'song types', each of which is sung by multiple males. The method requires first grouping a population of songs into song types using objective clustering procedures. Once clusters are defined, the 'typicality' of any one song can be measured by its proximity to the centre of the cluster to which it belongs (Lachlan et al., 2010, 2014). The typicality of a male's songs is then a measure of how well that male has succeeded in learning and reproducing the common features of particular song types (Lachlan & Nowicki, 2012). Typicality is distinct from stereotypy: the latter measures how consistently a male produces his version of a song type (Rivera-Gutierrez, Pinxten, & Eens, 2010; Smith, Brenowitz, Beecher, & Wingfield, 1997), whereas typicality measures how similar a male's version is to the versions produced by other males. The assumption that typicality

measures success in song learning is bolstered if it can be shown that a signaller benefits from producing more typical rather than less typical songs.

The assumptions of the typicality method fit well with the biology of swamp sparrows. Swamp sparrow songs almost always consist of a single syllable repeated in a steady-rate trill. Syllables are constructed from two to five notes drawn from a limited set of note type categories (Lachlan & Nowicki, 2015; Marler & Pickert, 1984). Lachlan et al. (2014) used a dynamic time-warping algorithm to compare hundreds of songs drawn from one swamp sparrow population, and then clustered songs based on their acoustic properties (see below). The clusters they uncovered corresponded well with song types identified by human observers through visual inspection of spectrograms. When tested with songs varying in typicality, female swamp sparrows gave a stronger courtship response to songs of high typicality (close to the centroids of their clusters) than to songs of low typicality (far from their centroids) (Lachlan et al., 2014). In parallel tests, male swamp sparrows gave a stronger aggressive response to songs of high typicality than to songs of low typicality (Lachlan et al., 2014). An earlier study found that female swamp sparrows prefer songs of males with good developmental histories (Searcy, Peters, Kipper, & Nowicki, 2010), and Lachlan et al. (2014) showed that variation in song typicality could explain these preferences.

Here we test the hypothesis that song signals cognitive ability by relating speed of learning of five cognitive tasks to five measures of song quality in wild-caught swamp sparrows. Two of these song measures assess song typicality, which we have argued above is a good measure of the quality of song learning in swamp sparrows. A third measure is song repertoire size, which we use in parallel to other studies of song and cognition, as reviewed above, even though it is only equivocally tied to sexual selection in swamp sparrows (Searcy, Searcy, & Marler, 1982) or to song learning in songbirds in general. The final two song measures are mean and minimum vocal deviation, which assess how closely a song approaches the maximum achievable combination of trill rate and frequency bandwidth (Podos, 2001). Vocal deviation is not considered a measure of song learning, but rather of vocal performance, that is, of the ability to perform difficult motor skills in vocal production (Byers, Hebets, & Podos, 2010; Podos & Nowicki, 2004). Because vocal deviation is not expected to reflect song learning, it seems less likely to be associated with cognition than are the other song measures; we nevertheless include vocal deviation in this study, as there is evidence that it is another song parameter tied to sexual selection in swamp sparrows (Ballentine, Hyman, & Nowicki, 2004; DuBois, Nowicki, & Searcy, 2011).

METHODS

Study Sites and Subjects

The study was carried out during May–July in the years 2014–2016. Subjects were 49 adult male swamp sparrows from Conneaut Marsh (Pennsylvania State Gamelands 213), in Crawford County, Pennsylvania, U.S.A. The subjects were initially captured on their territories using mist nets, banded with unique combinations of coloured leg bands, and then released. Song repertoires of banded males were then recorded in the field using Marantz PMD 660 or 670 digital recorders, Shure SM-57 microphones and Sony PBR-330 parabolas. Recordings were examined visually using spectrograms made with Syrinx (Burt, Campbell, & Beecher, 2001). Ballentine (2006) found that male swamp sparrows typically cycle through their entire repertoire before repeating a song type, so that if a male is recorded continuously until he repeats a song type, the entire repertoire is obtained in over 90% of cases. We considered

the entire repertoire to be recorded if the number of recorded switches between song types exceeded the number of recorded song types, which requires recording one additional song type switch beyond what is implied by the cycling rule. We met our recording criterion for 48 of 49 subjects. The one exception was a male that we recorded on four different days but who switched song types only once within a day; we nevertheless recorded the modal number of song types (3) from this male. We exceeded our recording criterion for 44 of the remaining 48 subjects, by a mean of 5.3 switches (range 1–19); only one of the 44 (2.3%) added an additional song type after meeting criterion. The mean repertoire size for our 49 subjects was 3.1 song types (range 2–5), which is identical to the mean previously reported for our study population by Ballentine et al. (2004) and Ballentine (2006).

After they were recorded, subjects were again captured using mist nets and transported approximately 20 km to the Pymatuning Laboratory of Ecology in Linesville, Pennsylvania. Each subject was then housed in a separate 46 × 22 × 26 cm cage that was in turn housed within one of 10 individual sound attenuation chambers (Industrial Acoustics AC-1). Chamber doors were for the most part kept open so that subjects could interact vocally. All procedures for housing and testing were approved by the Institutional Animal Care and Use Committees of Duke University (Protocol A033-14-02) and the University of Miami (Protocol 12-073) and were in accord with the Guidelines for the treatment of animals in behavioural research and teaching of the Animal Behavior Society and the Association for the Study of Animal Behaviour. Permits to capture and temporarily hold swamp sparrows were obtained from the U.S. Fish and Wildlife Service and the Pennsylvania Game Commission. Subjects were held for approximately 3 weeks; after cognitive testing, all were released in good health at the site of capture.

Cognitive Tests

Subjects were presented first with a test for neophobia and then with five cognitive tests in a fixed order: (1) novel foraging, (2) colour association, (3) colour reversal, (4) spatial learning and (5) detour reaching. We tested 19 subjects in 2014, 20 in 2015 and 10 in 2016. Because of timing constraints we were not able to complete all the tests for all the subjects in 2014. Consequently the total sample sizes were 49 for the neophobia test and the first three cognitive tests, 48 for spatial learning and 46 for detour reaching. We tested subjects daily starting at 1000 hours after 4 h of food deprivation, which was imposed to increase motivation to work for food rewards. For all tasks the food rewards were larvae of the mealworm beetle, *Tenebrio molitor*, which are a preferred food for captive swamp sparrows. Continued motivation to eat was tested immediately after the last test of the day for each subject by reintroducing mixed seed to each subject's cage. Latency to feed was less than 2 min on 1047 of 1053 (99.4%) of these motivation trials, and ≤10 s on 949 of 1053 (90.1%).

Cognitive tests were administered to one subject at a time in the subject's home cage and viewed remotely via video. Subjects were visually isolated from each other during testing. For each task, cognitive performance was measured as the number of trials needed to reach a specified criterion for successful learning.

Neophobia

We measured neophobia not as a cognitive skill, but as a noncognitive covariate that might affect cognitive performance. In neophobia trials, subjects were presented for the first time with a foraging grid, a 13.5 × 9 × 2.5 cm block of plastic containing six wells 1.3 cm in diameter and 0.8 cm deep, which was placed inside the subject's home cage (see Anderson et al., 2017, for an illustration). One mealworm was placed in each of four wells, and

neophobia was assessed as the amount of time a subject took to remove the first mealworm.

Novel foraging

Subjects learned to remove a plastic lid covering a well in order to extract a mealworm from the well. Subjects learned this task in five stages: (1) mealworms were placed in four wells with no lids; (2) lids were placed next to the four baited wells; (3) lids half-covered the baited wells; (4) lids were placed to loosely cover the baited wells; and (5) lids were fitted snugly over the baited wells. A trial was deemed successful if the subject obtained at least two of the four available mealworms within 2 min. Subjects needed to succeed on three of four consecutive trials to pass stages 1–4 and to succeed on four of five consecutive trials to pass stage 5. The most common technique used by subjects to solve this task was to pry the lid up and off using their bills. This technique is novel for swamp sparrows in that they are not known to perform a similar foraging movement in nature.

Colour association

Subjects were tested with two adjacent six-well grids. Four of the 12 wells were covered with blue lids and four with yellow lids, with positions of the colours intermixed and changed arbitrarily for each trial. For each subject, one colour was chosen to be rewarded and one to be unrewarded, with choices balanced across subjects. For all trials, mealworms were placed only in the wells covered by lids of the rewarded colour. To pass this task a subject had to remove all four lids of the rewarded colour before removing any lids of the unrewarded colour or remove two to three lids of the rewarded colour and none of the unrewarded colour on six of seven consecutive trials.

Colour reversal

The set-up was the same as in the colour association task, except that the reward scheme was reversed for each subject: the unrewarded colour in the previous task now was rewarded and the previously rewarded colour was now unrewarded. To pass the task, the subject again had to remove all four lids of the rewarded colour before removing any lids of the unrewarded colour or two to three lids of the rewarded colour and none of the unrewarded on six of seven consecutive trials.

Spatial learning

This task employed four smaller plastic blocks each containing a single covered well and each placed in one corner of the subject's cage. As a preliminary test for spatial preferences, six trials were first run with the blocks in all four corners rewarded. In the subsequent spatial preference trials, we chose one corner to be rewarded from among those that were not the most preferred or the least preferred by the subject in these preliminary trials. The criterion for success on the spatial learning task was to choose the correct block first on six of seven consecutive trials. Subjects that reached a point at which they were unable to meet the criterion in 48 trials were considered to have failed and were assigned a score of 48. After a bird completed the spatial task, we conducted a probe trial in which we placed the baited block in a different corner and ran the trial as usual. Subjects chose the rewarded block first on only 2 of 35 completed probes (6%), significantly less than the random expectation of 25% ($\chi^2 = 6.94$, $P = 0.008$), indicating that the subjects were not able to locate the reward using direct sensory cues.

Detour reaching

Subjects learned to remove a mealworm from a clear plastic cylinder laid on its long axis by detouring to one of the open ends

without first pecking at the food through the solid walls of the cylinder (which was never successful). Each subject was first presented with a baited cylinder made from black opaque plastic. To advance from this stage, the subject had to remove the mealworm from the cylinder without pecking the sides in four of five consecutive trials. The subject was then tested with a clear plastic cylinder, where the criterion for success was to remove the mealworm without error in six of seven consecutive trials. Overall success was measured by the number of trials taken to reach criterion on the clear cylinder stage.

Interobserver agreement

A second observer scored the videos of just over 10% of all trials (604 of 5709), spread evenly over testing days and chosen randomly within days. Rescoring was done blind to the original scores. Total agreement (House, House, & Campbell, 1981) was 98.5% and Cohen's kappa (Cohen, 1960) was 0.970 (± 0.01 SE). Disagreements were rare (1–3) for all five cognitive tests.

Song Analysis

For each subject, we measured song repertoire size as well as the typicality and vocal deviation of each of his song types. Repertoire size was estimated as described above, by visual inspection of spectrograms. For measurement of vocal deviation, we chose up to 10 exemplar songs of each song type based on recording quality. We then measured vocal deviation using Signal v.4.0 Software (Engineering Design, Belmont, MA, U.S.A.) following the methods of Ballentine et al. (2004) and DuBois, Nowicki, and Searcy (2009) (256-point transform length, 5.8 ms time resolution, 172.3 Hz frequency resolution). Briefly, we measured the trill rate as the number of syllables per second, averaged across the entire song. We measured frequency bandwidth as the difference between the maximum and minimum frequencies at -36 dB relative to the peak amplitude frequency. In cases where fewer than 10 exemplars were available, we measured as many as possible (mean: 9.5; range 2–10). Vocal deviation was then estimated as the minimum orthogonal deviation from a performance boundary, using as the boundary the upper-bound regression line of frequency bandwidth on trill rate calculated for the Emberizidae by Podos (1997). Vocal deviation was then averaged across the exemplars of each song type for subsequent analysis. We used both a male's mean vocal deviation (the average vocal deviation across all song types in his repertoire) and his minimum vocal deviation (the vocal deviation of the song type in his repertoire with the lowest deviation from the performance boundary) when investigating relationships between vocal performance and cognitive ability.

We measured the typicality of song types using the methods of Lachlan et al. (2014). In this approach, the songs of a population are first clustered into song types based on multiple acoustic measurements, and typicality of a song is then measured based on how close it lies to the centroid of the cluster in which it falls (Fig. 1). For this analysis, we chose three song exemplars from the 10 used for the vocal performance analysis. We measured the signal-to-noise ratio of each of the three exemplars, and chose the one with the highest signal-to-noise ratio for analysis. Chosen songs ($N = 153$) were then added to a database of 658 swamp sparrow songs from the same Pennsylvania population (Lachlan et al., 2014) using the software package Luscinia (Robert Lachlan; <http://luscinia.sourceforge.net>). We compared each of the song types sung by the males in our sample to each other and to the other songs in the database for the population using the dynamic time-warping (DTW) algorithm in Luscinia. Dynamic time warping searches for an optimal alignment between two signals using the Euclidean distance between several acoustic features. This DTW analysis has

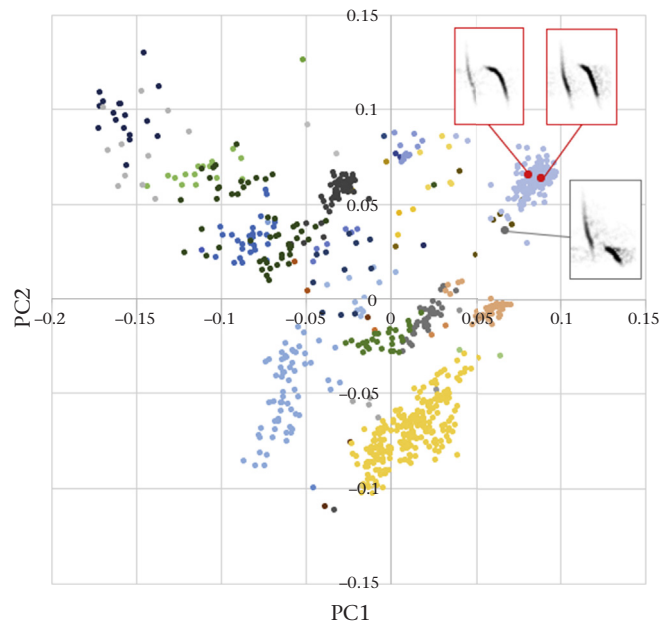


Figure 1. How song typicality is measured. Each point represents a swamp sparrow syllable recorded in the study population located in a space defined by the first two principal components from the dynamic time-warping algorithm in Luscinia. The different colours represent different clusters used to define song types; additional principal component axes not illustrated here are also used to distinguish clusters. Three syllables from one cluster are shown, two (delineated in red) of high typicality (near the cluster centroid) and one (delineated in dark grey) of low typicality.

been used previously to generate song and note type categories in swamp sparrows that match both the subjective assessments of human observers (Lachlan et al., 2010, 2014) and the categorical perception of the swamp sparrows themselves (Lachlan & Nowicki, 2015; Lachlan et al., 2010). The acoustic features we used for analysis were spectrograph measures of syllables: time (weight of 10), fundamental frequency (1), fundamental frequency change (1) and 'vibrato amplitude' (1). Additional parameter settings in Luscinia include compression factor (0.25), minimum element length (10), 's.d.' ratio (1), maximum warp (100%) and cost for alignment error (0.2). Syllable comparison was made by individual element, with weight by amplitude, log transform of frequencies, weight features with SDs, interpolate in DTW and dynamic time warping all selected. All these Luscinia settings were based on those used for analysis of swamp sparrow songs by Lachlan et al. (2014). The DTW analysis provides dissimilarity scores between each pair of songs in the full sample of 811 songs.

Songs were then categorized into population-wide types using unweighted pair group method with arithmetic mean (UPGMA) hierarchical clustering of the dissimilarity scores from the DTW analysis. A peak in the Global Silhouette Index (Rousseeuw, 1987) was used to decide where to cut the resulting tree into different syllable types. In this analysis, we identified 40 overall song types (Fig. 1). This number of song types is lower than in previous descriptions of the population using this method (Lachlan et al., 2014) but is closer to the number of song types identified in subjective clustering by human observers.

To measure typicality of the songs in our sample, we quantified how typical each song was of its song type cluster. Following Anderson's (2006) method for measuring multivariate dispersion, we carried out a principal coordinates analysis of the matrix of song dissimilarities, and then calculated the multivariate centroid of each song type cluster. As our measure of typicality, we calculated the Euclidean distance, d_c , between each song and its song type

centroid. A low value of d_c means that a song is similar to the centroid of its song type and is therefore highly typical of its type. As with vocal deviation, we produced two measures of typicality: average typicality (the average d_c across all song types in a male's repertoire) and maximum typicality (the d_c of the song type in a male's repertoire with the shortest Euclidean distance to its song type centroid).

Statistical Analysis

All statistical tests are two tailed. Several of the cognitive performance measures and the song measures failed Shapiro–Wilk tests for normality ($P < 0.05$). For uniformity, we used nonparametric tests to analyse associations involving all of these measures. The five cognitive tests were each run in three different years, so before combining data across years we tested for between-year differences using nonparametric ANOVA. Two of the five tests showed significant between-year differences: novel foraging ($F_{2,46} = 8.91$, $P < 0.001$) and spatial learning ($F_{2,46} = 5.52$, $P = 0.007$). For the other three tasks, no evidence of between-year differences was found ($P > 0.70$ in each case). Further analysis showed that for the novel foraging results, 2015 was the year that differed from the other two, whereas for spatial learning, it was 2016 that differed. Accordingly, we exclude one of these years from some of the analyses involving novel foraging and spatial learning, specifically from tests that cannot handle year as a fixed effect.

Associations of cognition measures with each other and with neophobia were examined using Spearman rank correlations. We tested whether our song variables predicted each cognitive measure using general linear mixed models fit by PQL maximum likelihood with neophobia as a covariate and year as a fixed effect. GLMMs were run in R (v.3.3.1, R Foundation for Statistical Computing, Vienna, Austria) with the function `glmmPQL` of the MASS package (Venables & Ripley, 2002). The probability distribution used for each model was determined by a graphical method using the `qpp` function of the `car` package (Fox & Weisberg, 2011) in R (v.3.3.1), which creates a display of the observed data against the quantiles modelled by the distribution (Fox, 2008). The graphical method allows one to determine visually which distribution provides the best fit to the data, and is widely recommended (Bolker et al., 2009; Zuur, Ieno, Walker, Saveliev, & Smith, 2009). A Gaussian distribution with a logarithmic link was used for novel foraging and colour association, and a gamma distribution with an identity link was used for colour reversal, spatial learning and detour reaching.

RESULTS

Variation in Song and Cognition Measures

Four of our five song measures showed considerable variation: mean typicality (measured by mean d_c) ranged from 0.0396 to 0.0946, maximum typicality (measured by minimum d_c) ranged from 0.0240 to 0.0828, mean vocal performance (measured as mean vocal deviation) ranged from 3.71 to 17.12, and maximum performance (measured as minimum vocal deviation) ranged from 1.0 to 16.1. Repertoire size, by contrast, showed rather limited variation: 31 of the 49 males (63%) had a repertoire size of three song types (range 2–5). This limited amount of variation in repertoire size is typical for this species (Mowbray, 1997). Four of the five cognition measures showed considerable variation: colour association (range 9–37), colour reversal (range 10–44), spatial learning (range 7–48) and inhibitory control (range 15–53). Only the novel foraging task showed limited variation (range 16–25).

Associations among Cognitive Measures

The five cognitive measures tended to show low associations with each other (Table 1). The only statistically significant association ($P < 0.001$) was a moderately strong positive correlation between the colour association score and the colour reversal score, meaning that individuals that performed well on colour association also performed well on colour reversal. None of the other pairwise correlations approached statistical significance ($P > 0.25$ in all cases). Of a total of 10 correlations, only three had positive signs, while the remaining seven were negative.

Song as a Predictor of Cognitive Performance

In each of five linear mixed effects models, we tested whether the set of song variables we measured (repertoire size, average typicality, maximum typicality, average performance, maximum performance) could predict one of the five cognitive measures, with neophobia as a covariate and year as a random factor. For novel foraging (Table 2), neophobia showed a significant positive association ($P = 0.003$), meaning that individuals that were more neophobic took more trials to reach criterion on the cognitive task. None of the song measures was significantly associated with success on the novel foraging task, although the association with repertoire size approached significance ($P = 0.062$).

For colour association (Table 3), neophobia was not significantly associated with learning success, nor were any of the five song measures. Similarly, neither neophobia nor any of the five song measures was associated with learning success on the colour reversal task (Table 4), the spatial learning task (Table 5) or the inhibitory control task (Table 6).

DISCUSSION

The goal of this study was to test the hypothesis that learned features of song in male songbirds serve as a signal of male cognitive ability, a signal that would presumably be directed principally at conspecific females. The results of the study do not support the predictions of that hypothesis, in that none of the song features that we measured showed a statistical association with any of the measures of cognitive performance we produced for male swamp sparrows. The case for expecting such an association was strongest for the two measures of song typicality (mean and maximum typicality), as typicality is both a logical measure of success in song learning (Lachlan & Nowicki, 2012) and a feature known to be associated with female preferences in swamp sparrows (Lachlan et al., 2014). The case for expecting associations with cognitive performance was weaker for the two performance measures (mean and maximum performance), because although song performance is associated with female preferences in swamp sparrows (Ballentine et al., 2004), differences in the ability of males to sing physically challenging songs are not an obvious outcome of differences in song learning in this or other songbird species. The case for expecting associations with cognition was weaker still for repertoire size, which is neither well associated with female preferences in swamp sparrows (Searcy et al., 1982) nor necessarily a measure of song learning. Regardless of the strength of the a priori expectations, however, the evidence for associations between the song measures and cognitive performance was negative in all cases.

The negative evidence on associations between song and cognitive performance in this study is based on sample sizes of 46–49 subjects, depending on the cognitive test. In studies of other songbirds, a few significant associations between song and cognitive performance have been found, based on smaller numbers of subjects. Boogert et al. (2008) found in a sample of 26 zebra finches

Table 1
Spearman rank correlations between the different cognitive measures

	Colour association	Colour reversal	Spatial learning	Inhibitory control
Novel foraging	–0.213 <i>P</i> =0.266 <i>N</i> =29	–0.104 <i>P</i> =0.593 <i>N</i> =29	0.076 <i>P</i> =0.750 <i>N</i> =20	–0.050 <i>P</i> =0.809 <i>N</i> =26
Colour association		0.464 <i>P</i> =0.001 <i>N</i> =49	–0.020 <i>P</i> =0.904 <i>N</i> =38	–0.043 <i>P</i> =0.776 <i>N</i> =46
Colour reversal			0.171 <i>P</i> =0.304 <i>N</i> =38	–0.063 <i>P</i> =0.679 <i>N</i> =46
Spatial learning				–0.121 <i>P</i> =0.484 <i>N</i> =36

Table 2
Linear mixed effects model predicting success on the novel foraging task from song measures

	Estimate	SE	<i>df</i>	<i>t</i>	<i>P</i>
Intercept	2.96	0.14	42	21.2	<0.0001
Neophobia	0.00088	0.00028	42	3.126	0.003
Repertoire size	–0.063	0.033	42	–1.914	0.062
Average typicality	0.32	1.93	42	0.167	0.868
Maximum typicality	–0.62	2.11	42	–0.294	0.770
Average performance	0.012	0.011	42	1.021	0.313
Maximum performance	–0.003	0.010	42	–0.301	0.765

Neophobia is entered as a covariate and year as a random factor.

Table 3
Linear mixed effects model predicting success on the colour association task from song measures

	Estimate	SE	<i>df</i>	<i>t</i>	<i>P</i>
Intercept	2.793	0.473	42	5.908	<0.0001
Neophobia	–0.0002	0.001	42	–0.159	0.874
Repertoire size	0.086	0.112	42	0.771	0.445
Average typicality	–1.025	6.521	42	–0.157	0.876
Maximum typicality	–2.614	7.137	42	–0.366	0.716
Average performance	–0.010	0.038	42	–0.250	0.804
Maximum performance	0.011	0.034	42	0.307	0.760

Neophobia is entered as a covariate and year as a random factor.

Table 4
Linear mixed effects model predicting success on the colour reversal task from song measures

	Estimate	SE	<i>df</i>	<i>t</i>	<i>P</i>
Intercept	11.85	8.68	42	1.365	0.180
Neophobia	0.043	0.024	42	1.776	0.083
Repertoire size	0.152	2.082	42	0.073	0.942
Average typicality	164.9	123.0	42	1.340	0.187
Maximum typicality	–23.38	140.1	42	–0.167	0.868
Average performance	–0.217	0.718	42	–0.302	0.764
Maximum performance	0.376	0.633	42	0.595	0.555

Neophobia is entered as a covariate and year as a random factor.

Table 5
Linear mixed effects model predicting success on the spatial learning task from song measures

	Estimate	SE	<i>df</i>	<i>t</i>	<i>P</i>
Intercept	6.204	14.042	41	0.442	0.661
Neophobia	0.0666	0.0399	41	1.672	0.102
Repertoire size	–2.464	2.991	41	–0.824	0.415
Average typicality	220.1	189.4	41	1.162	0.252
Maximum typicality	58.83	198.4	41	0.296	0.768
Average performance	0.329	1.224	41	0.268	0.790
Maximum performance	–0.048	1.077	41	–0.044	0.965

Neophobia is entered as a covariate and year as a random factor.

Table 6
Linear mixed effects model predicting success on the inhibitory control task from song measures

	Estimate	SE	<i>df</i>	<i>t</i>	<i>P</i>
Intercept	17.08	11.63	39	1.469	0.150
Neophobia	–0.027	0.017	39	–1.555	0.128
Repertoire size	–1.55	2.73	39	–0.569	0.573
Average typicality	106.9	161.1	39	0.664	0.511
Maximum typicality	–100.4	170.8	39	–0.588	0.560
Average performance	1.75	0.992	39	1.764	0.086
Maximum performance	–0.878	0.936	39	–0.939	0.354

Neophobia is entered as a covariate and year as a random factor.

that success in learning a novel foraging task was positively correlated with one measure of song complexity, the number of elements per song, but not with two others, the number of unique elements per song and song duration. In a second zebra finch study with a larger sample (35 subjects), these measures of song complexity were not significantly associated with the same measure of cognitive performance (Templeton et al., 2014). In song sparrows, Boogert et al. (2011) found that repertoire size of wild-caught adults was significantly positively associated with performance on a detour-reaching task in a sample of 22 subjects, whereas in tests with larger sample sizes (51–52), repertoire size was not significantly associated with performance on motor, colour association and colour reversal tasks. The relationship with performance on a detour-reaching task was not supported in a subsequent study of hand-reared males (*N* = 19) of the same species (Anderson et al., 2017). Sewall et al. (2013) found a significant negative association between repertoire size and spatial learning performance in a sample of 14 song sparrows, but again this relationship was not confirmed in the study of hand-reared birds (Anderson et al., 2017). Analyses of the data from the hand-rearing study that controlled for neophobia revealed some significant associations between accuracy of song learning and cognitive performance, but these associations were negative as often as positive (Anderson et al., 2017). In European starlings, a measure of song complexity (song bout length) was significantly positively associated with performance on a spatial learning task (*N* = 19) but not with performance on a social learning task (*N* = 19) (Farrell et al., 2012). Overall, the tests with the largest samples (Boogert et al., 2011; Templeton et al., 2014; this study) have been particularly likely to find no significant association between song and cognitive performance.

Why has the expectation that song quality signals cognitive ability received such inconsistent support? Two main arguments have been used to justify the expectation in the first place. The first is that both song and cognition are negatively affected by the same developmental stresses, so that individuals experiencing little stress during development should have both good song and good cognition, while those experiencing greater stress should show the opposite (Boogert et al., 2008; Nowicki, Hasselquist, Bensch, & Peters, 2000; Peters et al., 2014; Searcy & Nowicki, 2009). A variety of experimental evidence has supported negative effects on song of developmental stressors such as under-nutrition and parasitism (Peters et al., 2014; Spencer & MacDougall-Shackleton, 2011). This evidence includes results from swamp sparrows showing that nutritional restriction lowers song learning accuracy (Nowicki, Searcy, & Peters, 2002a) and that females prefer songs of males with good growth histories (Searcy et al., 2010). It should be noted, however, that there are studies of the effects of early stresses on song development in songbirds that produced entirely negative results (Gil, Naguib, Riebel, Rutstein, & Gahr, 2006; Müller, Vergauwen, & Eens, 2010), while even the studies with some positive results typically found that some song parameters are not

affected by particular developmental stresses (Peters et al., 2014). Effects of developmental stress on cognitive performance have also been variable. In western scrub-jays, *Aphelocoma californica*, early nutritional stress had a negative effect on performance in two spatial tasks in adulthood but no effect on colour association tasks (Pravosudov, Lavenex, & Omanska, 2005). In European starlings, early nutritional stress resulted in poorer adult performance on a spatial learning task but in better performance on a social learning task (Farrell et al., 2012). In a second starling experiment, early nutritional stress negatively affected learning in a colour association task and in some auditory learning tasks but not in others (Farrell, Morgan, & MacDougall-Shackleton, 2016). In zebra finches, early nutritional restriction impaired adult spatial learning but enhanced spatial associative learning (Kriengwatana, Farrell, Aitken, Garcia, & MacDougall-Shackleton, 2015). Given the variability of effects of developmental stresses on song and especially on cognition, it would not be surprising if this pathway produced somewhat inconsistent associations between song and cognition.

The second argument for expecting associations between song and cognitive performance is that song learning is itself a cognitive task, and that performance on all cognitive tasks tends to be positively correlated. Cognition has been defined as ‘mechanisms by which animals acquire, process, store, and act on information from the environment’ (Shettleworth, 2010, p. 4). Swamp sparrows acquire information on the structure of conspecific song during their first few months of life by listening to the songs of older birds (Marler & Peters, 1988), store that information in memory for about 8 months (Marler & Peters, 1982a), and then use the information to shape the structure of their own songs as they begin singing prior to their first breeding season (Marler & Peters, 1982b). Swamp sparrows thus acquire, store and act on information from the environment during song learning, fully satisfying at least this one definition of cognition. That performance on all cognitive tasks tends to be positively correlated is true at least in humans (Carroll, 1993; Plomin, 2001; Spearman, 1904), which are undoubtedly the best-studied species in this regard, and perhaps also in rodents (Anderson, 1993; Matzel et al., 2003). This argument is weakened, however, by recent results suggesting that measurements of performance on different cognitive tasks are not consistently positively correlated in birds (Anderson et al., 2017; Farrell et al., 2016; Guillette, Hahn, Hoeschele, Przyszlupski, & Sturdy, 2015; Keagy et al., 2011; Nettle et al., 2015; Shaw, Boogert, Clayton, & Burns, 2015). In the present study, only one out of 10 correlations between pairs of cognitive performance measures was found to be statistically significant: that between colour association and colour reversal (Table 1). Similarly, a positive correlation between performance on a colour association task and a colour reversal task was the only significant association out of six such correlations assessed in wild-caught male song sparrows (Boogert et al., 2011) and out of 10 assessed in hand-reared male and female song sparrows (Anderson et al., 2017). That these two tasks produce the strongest positive correlations is not surprising, in that both require learning the same kind of colour association under very similar conditions. The colour reversal task, however, also requires abandoning a previously learned colour association, and Bebus, Small, Jones, Elderbrock, and Schoech (2016) actually found a significant negative correlation between performance on very similar colour association and colour reversal tasks in Florida scrub-jays, *Aphelocoma coerulescens*.

Given that cognitive performance is not consistently positively correlated across different tasks in songbirds, the second argument for expecting associations between learned features of song and cognitive performance in general breaks down. In addition, the fact that correlations between various aspects of cognitive performance are inconsistent in songbirds means that if one particular measure

of song is positively associated with one particular measure of cognition, that song measure is not likely to be positively associated with additional measures of cognition.

Present evidence thus suggests that song is not a reliable indicator of general cognitive ability in swamp sparrows (this study), nor in the closely related song sparrow (Anderson et al., 2017), nor perhaps in any songbird. Evidence from other studies indicates that various features of birdsong are associated with other types of information important to receivers, such as (in swamp sparrows and song sparrows) age and size (Ballentine, 2009), early growth history (Searcy et al., 2010), site of natal origin (Stewart & MacDougall-Shackleton, 2008), degree of inbreeding (Reid, Arcese, Cassidy, Marr et al., 2005), immunocompetence (Pfaff, Zanette, MacDougall-Shackleton, & MacDougall-Shackleton, 2007) and overall fitness (Reid, Arcese, Cassidy, Hiebert et al., 2005). Thus females should be selected to attend to song in mate choice, even if song does not reliably signal cognitive ability.

Acknowledgments

We thank Rob Lachlan for his generous help with the song analysis and Caitlin Cantrell, Rory Eggleston and Philippa Tanford for assistance with the cognitive tests. We thank the Pennsylvania Game Commission for access to field sites and the Pymatuning Laboratory of Ecology of the University of Pittsburgh for logistical support. We gratefully acknowledge the National Science Foundation for financial support through grants to S.N. (IOS-1144991) and W.A.S. (IOS-1144995).

References

- Anderson, B. (1993). Evidence from the rat for a general factor that underlies cognitive performance and that relates to brain size: Intelligence? *Neuroscience Letters*, *153*, 98–102.
- Anderson, M. J. (2006). Distance-based tests for homogeneity of multivariate dispersions. *Biometrics*, *62*, 245–253.
- Anderson, R. C., Searcy, W. A., Peters, S., Hughes, M., DuBois, A. L., & Nowicki, S. (2017). Song learning and cognitive ability are not consistently related in a songbird. *Animal Cognition*, *20*, 309–320.
- Andersson, M. (1982). Female choice selects for extreme tail length in a widowbird. *Nature*, *299*, 818–820.
- Ballentine, B. (2006). Morphological adaptation influences the evolution of a mating signal. *Evolution*, *60*, 1936–1944.
- Ballentine, B. (2009). The ability to perform physically challenging songs predicts age and size in male swamp sparrows, *Melospiza georgiana*. *Animal Behaviour*, *77*, 973–978.
- Ballentine, B., Hyman, J., & Nowicki, S. (2004). Vocal performance influences female response to male bird song: An experimental test. *Behavioral Ecology*, *15*, 163–168.
- Bebus, S. E., Small, T. W., Jones, B. C., Elderbrock, E. K., & Schoech, S. J. (2016). Associative learning is inversely related to reversal learning and varies with nestling corticosterone exposure. *Animal Behaviour*, *111*, 251–260.
- Beecher, M. D., & Brenowitz, E. A. (2005). Functional aspects of song learning in songbirds. *Trends in Ecology & Evolution*, *20*, 143–149.
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., et al. (2009). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*, 127–135.
- Boogert, N. J., Anderson, R. C., Peters, S., Searcy, W. A., & Nowicki, S. (2011). Song repertoire size in male song sparrows correlates with detour reaching, but not with other cognitive measures. *Animal Behaviour*, *81*, 1209–1216.
- Boogert, N. J., Giraldeau, L.-A., & Lefebvre, L. (2008). Song complexity correlates with learning ability in zebra finch males. *Animal Behaviour*, *76*, 1735–1741.
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, *86*, 599–614.
- Burt, J. M., Campbell, S. E., & Beecher, M. D. (2001). Song type matching as threat: A test using interactive playback. *Animal Behaviour*, *62*, 1163–1170.
- Byers, J., Hebets, E., & Podos, J. (2010). Female choice based on male motor performance. *Animal Behaviour*, *79*, 771–778.
- Carroll, J. B. (1993). *Human cognitive abilities*. New York, NY: Cambridge University Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.
- DuBois, A. L., Nowicki, S., & Searcy, W. A. (2009). Swamp sparrows modulate vocal performance in an aggressive context. *Biology Letters*, *5*, 163–165.

- DuBois, A. L., Nowicki, S., & Searcy, W. A. (2011). Discrimination of vocal performance by male swamp sparrows. *Behavioral Ecology and Sociobiology*, *65*, 717–726.
- Eales, L. A. (1985). Song learning in zebra finches: Some effects of song model availability on what is learnt and when. *Animal Behaviour*, *33*, 1293–1300.
- Eens, M. (1997). Understanding the complex song of the European starling: An integrated ethological approach. *Advances in the Study of Behavior*, *26*, 355–434.
- Eens, M., Pinxten, R., & Verheyen, R. F. (1991). Male song as a cue for mate choice in the European starling. *Behaviour*, *116*, 210–238.
- Eens, M., Pinxten, R., & Verheyen, R. F. (1992). Song learning in captive European starlings, *Sturnus vulgaris*. *Animal Behaviour*, *44*, 1131–1143.
- Farrell, T. M., Morgan, A., & MacDougall-Shackleton, S. A. (2016). Developmental stress impairs performance on an association task in male and female songbirds, but impairs auditory learning in females only. *Animal Cognition*, *19*, 1–14.
- Farrell, T. M., Weaver, K., An, Y. S., & MacDougall-Shackleton, S. A. (2012). Song bout length is indicative of spatial learning in European starlings. *Behavioral Ecology*, *23*, 101–111.
- Fox, J. (2008). *Applied regression analysis and generalized linear models* (2nd ed.). Los Angeles, CA: Sage.
- Fox, J., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Thousand Oaks, CA: Sage <http://socserv.socsci.mcmaster.ca/jfox/Books/Companion>.
- Furnham, A. (2009). Sex differences in mate selection preferences. *Personality and Individual Differences*, *47*, 262–267.
- Gil, D., Naguib, M., Riebel, K., Rutstein, A., & Gahr, M. (2006). Early condition, song learning, and the volume of song brain nuclei in the zebra finch (*Taeniopygia guttata*). *Journal of Neurobiology*, *66*, 1602–1612.
- Guillette, L. M., Hahn, A. H., Hoeschele, M., Przyslupski, A. M., & Sturdy, C. B. (2015). Individual differences in learning speed, performance accuracy and exploratory behaviour in black-capped chickadees. *Animal Cognition*, *18*, 165–178.
- Hill, G. E. (1991). Plumage coloration is a sexually selected indicator of male quality. *Nature*, *350*, 337–339.
- House, A. E., House, B. J., & Campbell, M. B. (1981). Measures of interobserver agreement: Calculation formulas and distribution effects. *Journal of Behavioral Assessment*, *3*, 37–57.
- Keagy, J., Savard, J. F., & Borgia, G. (2009). Male satin bowerbird problem-solving ability predicts mating success. *Animal Behaviour*, *78*, 809–817.
- Keagy, J., Savard, J. F., & Borgia, G. (2011). Complex relationship between multiple measures of cognitive ability and male mating success in satin bowerbirds, *Ptilonorhynchus violaceus*. *Animal Behaviour*, *81*, 1063–1070.
- Kriengwatanana, B., Farrell, T. M., Aitken, S. D. T., Garcia, L., & MacDougall-Shackleton, S. A. (2015). Early-life nutritional stress affects associative learning and spatial memory but not performance on a novel object test. *Behaviour*, *152*, 195–218.
- Lachlan, R. F., Anderson, R. C., Peters, S., Searcy, W. A., & Nowicki, S. (2014). Typical versions of learned swamp sparrow song types are more effective signals than are less typical versions. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20140252.
- Lachlan, R. F., & Nowicki, S. (2012). How reliable is song learning accuracy as a signal of male early condition? *American Naturalist*, *180*, 751–761.
- Lachlan, R. F., & Nowicki, S. (2015). Context-dependent categorical perception in a songbird. *Proceedings of the National Academy of Sciences of the United States of America*, *112*, 1892–1897.
- Lachlan, R. F., Verhagen, L., Peters, S., & ten Cate, C. (2010). Are there species-universal categories in bird song phonology and syntax? A comparative study of chaffinches (*Fringilla coelebs*), zebra finches (*Taeniopygia guttata*), and swamp sparrows (*Melospiza georgiana*). *Journal of Comparative Psychology*, *124*, 92–108.
- Li, N. P., Bailey, J. M., Kenrick, D. T., & Linsenmeier, J. A. W. (2002). The necessities and luxuries of mate preferences: Testing the tradeoffs. *Journal of Personality and Social Psychology*, *82*, 947–955.
- Marler, P., & Mundinger, P. (1971). Vocal learning in birds. In H. Moltz (Ed.), *Ontogeny of vertebrate behavior* (pp. 389–450). New York, NY: Academic Press.
- Marler, P., & Peters, S. (1981). Sparrows learn adult song and more from memory. *Science*, *213*, 780–782.
- Marler, P., & Peters, S. (1982a). Long-term storage of learned birdsong prior to production. *Animal Behaviour*, *30*, 479–482.
- Marler, P., & Peters, S. (1982b). Structural changes in song ontogeny in the swamp sparrow *Melospiza georgiana*. *Auk*, *99*, 446–458.
- Marler, P., & Peters, S. (1987). A sensitive period for song acquisition in the song sparrow, *Melospiza melodia*: A case of age-limited learning. *Ethology*, *76*, 89–100.
- Marler, P., & Peters, S. (1988). Sensitive periods for song acquisition from tape recordings and live tutors in the swamp sparrow, *Melospiza georgiana*. *Ethology*, *77*, 76–84.
- Marler, P., & Pickert, R. (1984). Species-universal microstructure in the learned song of the swamp sparrow (*Melospiza georgiana*). *Animal Behaviour*, *32*, 673–689.
- Marlowe, F. W. (2004). Mate preferences among Hadza hunter-gatherers. *Human Nature*, *15*, 365–376.
- Matzel, L. D., Han, Y. R., Grossman, H., Karnik, M. S., Patel, D., Scott, N., et al. (2003). Individual differences in the expression of a 'general' learning ability in mice. *Journal of Neuroscience*, *23*, 6423–6433.
- Møller, A. P. (1988). Female choice selects for male sexual tail ornaments in the monogamous swallow. *Nature*, *332*, 640–642.
- Mowbray, T. B. (1997). Swamp sparrow (*Melospiza georgiana*). In A. Poole, & F. Gill (Eds.), *The birds of North America* (Vol. 279, pp. 1–23). Philadelphia, PA: Academy of Natural Sciences; Washington, D.C.: American Ornithologists' Union.
- Müller, W., Vergauwen, J., & Eens, M. (2010). Testing the developmental stress hypothesis in canaries: Consequences of nutritional stress on adult song phenotype and mate attractiveness. *Behavioral Ecology and Sociobiology*, *64*, 1767–1777.
- Nelson, D. A. (2000). Song overproduction, selective attrition and song dialects in the white-crowned sparrow. *Animal Behaviour*, *60*, 887–898.
- Nettle, D., Andrews, C. P., Monaghan, P., Brilot, B. O., Bedford, T., Gillespie, R., et al. (2015). Developmental and familial predictors of adult cognitive traits in the European starling. *Animal Behaviour*, *107*, 239–248.
- Nordby, J. C., Campbell, S. E., & Beecher, M. D. (1999). Ecological correlates of song learning in song sparrows. *Behavioral Ecology*, *10*, 287–297.
- Nordby, J. C., Campbell, S. E., & Beecher, M. D. (2007). Selective attrition and individual song repertoire development in song sparrows. *Animal Behaviour*, *74*, 1413–1418.
- Nowicki, 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. *Proceedings of the Royal Society B: Biological Sciences*, *267*, 2419–2424.
- Nowicki, S., Searcy, W. A., & Peters, S. (2002a). Brain development, song learning and mate choice in birds: A review and experimental test of the 'nutritional stress hypothesis'. *Journal of Comparative Physiology A*, *188*, 1003–1014.
- Nowicki, S., Searcy, W. A., & Peters, S. (2002b). Quality of song learning affects female response to male bird song. *Proceedings of the Royal Society B: Biological Sciences*, *269*, 1949–1954.
- Peters, S., & Nowicki, S. (2017). Overproduction and attrition: The fates of songs memorized during song learning in songbirds. *Animal Behaviour*, *124*, 255–261.
- Peters, S., Searcy, W. A., & Nowicki, S. (2014). Developmental stress, song learning and cognition. *Integrative and Comparative Biology*, *54*, 555–567.
- Pfaff, J. A., Zanette, L., MacDougall-Shackleton, S. A., & MacDougall-Shackleton, E. A. (2007). Song repertoire size varies with HVC volume and is indicative of male quality in song sparrows (*Melospiza melodia*). *Proceedings of the Royal Society B: Biological Sciences*, *274*, 2035–2040.
- Plomin, R. (2001). The genetics of G in humans and mouse. *Nature Reviews Neuroscience*, *2*, 136–141.
- 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. (2004). Performance limits on birdsong. In P. Marler, & H. Slabbekoorn (Eds.), *Nature's music: The science of birdsong* (pp. 318–342). New York, NY: Academic Press.
- Prather, J. F., Peters, S., Nowicki, S., & Mooney, R. (2010). Persistent representation of juvenile experience in the adult songbird brain. *Journal of Neuroscience*, *30*, 10586–10598.
- Pravosudov, V. V., Lavenex, P., & Omanska, A. (2005). Nutritional deficits during early development affect hippocampal structure and spatial memory later in life. *Behavioral Neuroscience*, *119*, 1368–1374.
- Prokosh, M. D., Coss, R. G., Scheib, J. E., & Blozis, S. A. (2009). Intelligence and mate choice: Intelligent men are always appealing. *Evolution and Human Behavior*, *30*, 11–20.
- Reid, J. M., Arcese, P., Cassidy, A. L. E. V., Hiebert, S. M., Smith, J. N. M., Stoddard, P. K., et al. (2004). Song repertoire size predicts initial mating success in male song sparrows, *Melospiza melodia*. *Animal Behaviour*, *68*, 1055–1063.
- Reid, J. M., Arcese, P., Cassidy, A. L. E. V., Hiebert, S. M., Smith, J. N. M., Stoddard, P. K., et al. (2005). Fitness correlates of song repertoire size in free-living song sparrows (*Melospiza melodia*). *American Naturalist*, *165*, 299–310.
- Reid, J. M., Arcese, P., Cassidy, A. L. E. V., Marr, A. B., Smith, J. N. M., & Keller, L. F. (2005). Hamilton and Zuk meet heterozygosity? Song repertoire size indicates inbreeding and immunity in song sparrows (*Melospiza melodia*). *Proceedings of the Royal Society B: Biological Sciences*, *272*, 481–487.
- Riebel, K. (2009). Song and female mate choice in zebra finches: A review. *Advances in the Study of Behavior*, *40*, 197–238.
- Rivera-Gutierrez, H. F., Pinxten, R., & Eens, M. (2010). Multiple signals for multiple messages: Great tit, *Parus major*, song signals age and survival. *Animal Behaviour*, *80*, 451–459.
- Rosenberg, J., & Tunney, R. J. (2008). Human vocabulary use as display. *Evolutionary Psychology*, *6*, 538–549.
- Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*, 53–65.
- Searcy, W. A. (1984). Song repertoire size and female preferences in song sparrows. *Behavioral Ecology and Sociobiology*, *14*, 281–286.
- Searcy, W. A., & Andersson, M. (1986). Sexual selection and the evolution of song. *Annual Review of Ecology and Systematics*, *17*, 507–533.
- Searcy, W. A., & Nowicki, S. (2009). Consequences of brain development for sexual signaling in songbirds. In R. Dukas, & J. M. Ratcliffe (Eds.), *Cognitive ecology II* (pp. 71–87). Chicago, IL: University of Chicago Press.
- Searcy, W. A., Peters, S., Kipper, S., & Nowicki, S. (2010). Female response to song reflects male developmental history in swamp sparrows. *Behavioral Ecology and Sociobiology*, *64*, 1343–1349.
- Searcy, W. A., Searcy, M. H., & Marler, P. (1982). The response of swamp sparrows to acoustically distinct song types. *Behaviour*, *80*, 70–83.

- Searcy, W. A., & Yasukawa, K. (1996). Song and female choice. In D. E. Kroodsmas, & E. H. Miller (Eds.), *Ecology and evolution of acoustic communication in birds* (pp. 454–473). Ithaca, NY: Cornell University Press.
- Sewall, K. B., Soha, J. A., Peters, S., & Nowicki, S. (2013). Potential trade-off between vocal ornamentation and spatial ability in a songbird. *Biology Letters*, *9*, 20130344.
- Shackleford, T. K., Schmitt, D. P., & Buss, D. M. (2005). Universal dimensions of human mate preferences. *Personality and Individual Differences*, *39*, 447–458.
- Shaw, R. C., Boogert, N. J., Clayton, N. S., & Burns, K. C. (2015). Wild psychometrics: Evidence for 'general' cognitive performance in wild New Zealand robins, *Petroica longipes*. *Animal Behaviour*, *109*, 101–111.
- Shettleworth, S. J. (2010). *Cognition, evolution, and behavior* (2nd ed.). Oxford, U.K.: Oxford University Press.
- Shohet, A. J., & Watt, P. J. (2009). Female guppies *Poecilia reticulata* prefer males that can learn fast. *Journal of Fish Biology*, *75*, 1323–1330.
- Simons, M. J. P., & Verhulst, S. (2011). Zebra finch females prefer males with redder bills independent of song rate: A meta-analysis. *Behavioral Ecology*, *22*, 755–762.
- Smith, G. T., Brenowitz, E. A., Beecher, M. D., & Wingfield, J. C. (1997). Seasonal changes in testosterone, neural attributes of song control nuclei, and song structure in wild songbirds. *Journal of Neuroscience*, *17*, 6001–6010.
- Souza, A. L., Conroy-Beam, D., & Buss, D. M. (2016). Mate preferences in Brazil: Evolved desires and cultural evolution over three decades. *Personality and Individual Differences*, *95*, 45–49.
- Spearman, C. (1904). 'General intelligence', objectively determined and measured. *American Journal of Psychology*, *15*, 201–292.
- Spencer, K. A., & MacDougall-Shackleton, S. A. (2011). Indicators of development as sexually selected traits: The developmental stress hypothesis in context. *Behavioral Ecology*, *22*, 1–9.
- Spritzer, M. D., Meikle, D. B., & Solomon, N. G. (2005). Female choice based on male spatial ability and aggressiveness among meadow voles. *Animal Behaviour*, *69*, 1121–1130.
- Stewart, K. A., & MacDougall-Shackleton, E. A. (2008). Local song elements indicate local genotypes and predict physiological condition in song sparrows *Melospiza melodia*. *Biology Letters*, *4*, 240–242.
- Templeton, C. N., Laland, K. N., & Boogert, N. J. (2014). Does song complexity correlate with problem-solving performance in flocks of zebra finches? *Animal Behaviour*, *92*, 63–71.
- Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York, NY: Springer, ISBN 0-387-95457-0.
- Vyas, A., Harding, C., Borg, L., & Bogdan, D. (2009). Acoustic characteristics, early experience, and endocrine status interact to modulate female zebra finches' behavioral responses to songs. *Hormones and Behavior*, *55*, 50–59.
- Whittingham, L. A., & Dunn, P. O. (2016). Experimental evidence that brighter males sire more extra-pair young in tree swallows. *Molecular Ecology*, *25*, 3706–3715.
- Zuur, A., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. New York, NY: Springer.