

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

# **Behavioural responses to video and live presentations of females reveal a dissociation between performance and motivational aspects of birdsong**

Logan S. James\*, R. Fan\*, and J.T. Sakata#

Department of Biology, McGill University

\* equal contributions

# *corresponding author:*

Jon T. Sakata, PhD

Department of Biology

McGill University

Montreal, QC H3A 1B1

ph: 514-398-3636

sakatasongbirdlab@gmail.com; jon.sakata@mcgill.ca

29 **ABSTRACT**

30 Understanding the regulation of social behavioural expression requires insight into  
31 motivational and performance aspects of social behaviours. While a number of studies  
32 have independently investigated motivational or performance aspects of social  
33 behaviours, few have examined how these aspects relate to each other. By comparing  
34 behavioural variation in response to live or video presentations of conspecific females,  
35 we analysed how variation in the motivation to produce courtship song covaries with  
36 variation in performance aspects of courtship song in male zebra finches (*Taeniopygia*  
37 *guttata*). Consistent with previous reports, we observed that male zebra finches were  
38 less motivated to produce courtship songs to videos of females than to live  
39 presentations of females. However, we found that acoustic features that reflect song  
40 performance were indistinguishable between songs produced to videos of females and  
41 songs produced to live presentations of females. For example, songs directed at video  
42 presentations of females were just as fast and stereotyped as songs directed at live  
43 females. These experimental manipulations and correlational analyses reveal a  
44 dissociation between motivational and performance aspects of birdsong and suggest a  
45 refinement of neural models of song production and control. In addition, they support  
46 the efficacy of videos to study both motivational and performance aspects of social  
47 behaviours.

48

49 **Keywords:** social context, courtship, songbird, zebra finch, vocal consistency,  
50 stereotypy

51

## 52 INTRODUCTION

53 The extent and quality of various social displays, including communicative and  
54 courtship behaviours, reflect an individual's motivation and performance. Motivation  
55 refers to the "drive" to display a behaviour whereas performance refers to the fine  
56 motoric aspects of the behaviour. For example, internal and external states can affect  
57 the likelihood of displaying maternal behaviours (e.g., pup retrieval and grooming), and  
58 the latency and efficiency of pup-directed behaviours can vary between individuals as  
59 well as within individuals over time (Champagne et al., 2003; Clark et al., 2002;  
60 Stolzenberg et al., 2012). Both the motivation to engage in maternal behaviours and the  
61 performance of various components of maternal behaviour have important  
62 developmental consequences, and such findings highlight the importance of  
63 investigating both motivation and performance to gain a comprehensive understanding  
64 of social behaviour (Meaney, 2001; Rilling and Young, 2014). However, motivation and  
65 performance are often studied independently, and relatively little is known about the  
66 relationship between mechanisms regulating motivational and performance aspects of  
67 behaviour. In particular, little is known about the extent to which factors that affect the  
68 motivation to display a behaviour similarly affect the performance of the behaviour.

69  
70 Birdsong provides an excellent opportunity to assess the degree to which mechanisms  
71 underlying motivational and performance aspects of social behaviour are shared or  
72 independent. When presented with an adult female, adult male songbirds become  
73 motivated to produce courtship song, and individual differences in this motivation is  
74 important because female songbirds also tend to prefer males that display greater  
75 motivation to sing (Bradbury and Vehrencamp, 2011; Catchpole and Slater, 2008; Gil  
76 and Gahr, 2002; Sakata and Vehrencamp, 2012). Further, males alter a number of song  
77 performance features when producing courtship songs compared to non-courtship  
78 songs (Chen et al., 2016; Moser-Purdy and Mennill, 2016; Sakata and Vehrencamp,  
79 2012; Tocalino et al., 2016; Vignal et al., 2004; Woolley and Kao, 2015). For example,  
80 male zebra finches produce songs that are faster and more acoustically stereotyped  
81 when courting female conspecifics than when singing in isolation (Chen et al., 2016;

82 Cooper and Goller, 2006; Kao and Brainard, 2006; Sossinka and Böhner, 1980;  
83 Woolley et al., 2014). These performance-related song traits can affect a male's  
84 attractiveness and reproductive success, since female songbirds prefer the courtship  
85 version of an individual male's song, as well as males with song features that are  
86 generally characteristic of courtship song (e.g., faster and longer songs; Gil and Gahr,  
87 2002; Podos et al., 2009; Woolley and Doupe, 2008).

88  
89 Despite knowledge about the functional relevance of motivational and performance  
90 aspects of birdsong, little is known about how experimental variation in the motivation  
91 to produce courtship song relates to experimental variation in song performance. Brain  
92 areas that underlie the motivation to sing project to sensorimotor brain regions that  
93 regulate song performance, suggesting that song motivation could influence song  
94 performance (reviewed in Riters, 2012; Riters et al., 2004; Woolley and Kao, 2015). In  
95 addition, seasonal changes in the motivation to produce song have been found to  
96 covary with seasonal changes in song performance (Smith et al., 1997; Smith et al.,  
97 1995). On the other hand, some studies have found a dissociation between song  
98 motivation and performance (Alward et al., 2013; Ritschard et al., 2011; Toccalino et  
99 al., 2016).

100  
101 Here, we investigated variation in vocal performance across conditions that are known  
102 to modulate the motivation to produce courtship song. Video playbacks of social  
103 stimuli have been used to elicit a wide range social behaviours (Evans and Marler,  
104 1991; Fleishman and Endler, 2000; Gonçalves et al., 2000; Guillette and Healy, 2017;  
105 Oliveira et al., 1999; Ophir et al., 2005; Ord et al., 2002; Rosenthal, 1999; Uetz and  
106 Roberts, 2002; Ware et al., 2016), including courtship song in songbirds (Galoch and  
107 Bischof, 2007; Ikebuchi and Okanoya, 1999; Takahasi et al., 2005). Despite that videos  
108 can elicit courtship song, male songbirds have been found to sing less to videos of  
109 females than to live presentations of females (Ikebuchi and Okanoya, 1999). However,  
110 it is not known whether performance aspects of courtship song (e.g., tempo and  
111 stereotypy) are similarly reduced for songs produced to video presentations of females.

112 Previous studies of other social behaviours have found that behavioural performance  
113 can be distinct when individuals are presented with video or live presentations of  
114 conspecifics (Balshine-Earn and Lotem, 1998; Ord et al., 2002; Swaddle et al., 2006).  
115 Consequently, we analysed motivational and performance aspects of male zebra finch  
116 song in response to video and live presentations of females.

117

## 118 **MATERIALS AND METHODS**

### 119 *Animals*

120 Adult male zebra finches (*Taeniopygia guttata*; >4 months; n=13) were bred and raised  
121 in our colony at McGill University. Males were socially housed in same-sex group  
122 cages and visually isolated from females. Birds were kept on a 14L:10D photoperiod,  
123 with food and water provided *ad libitum*. All procedures were in accordance with  
124 McGill University Animal Care and Use Committee protocols, as well as guidelines  
125 from the Canadian Council on Animal Care.

126

### 127 *Video stimuli*

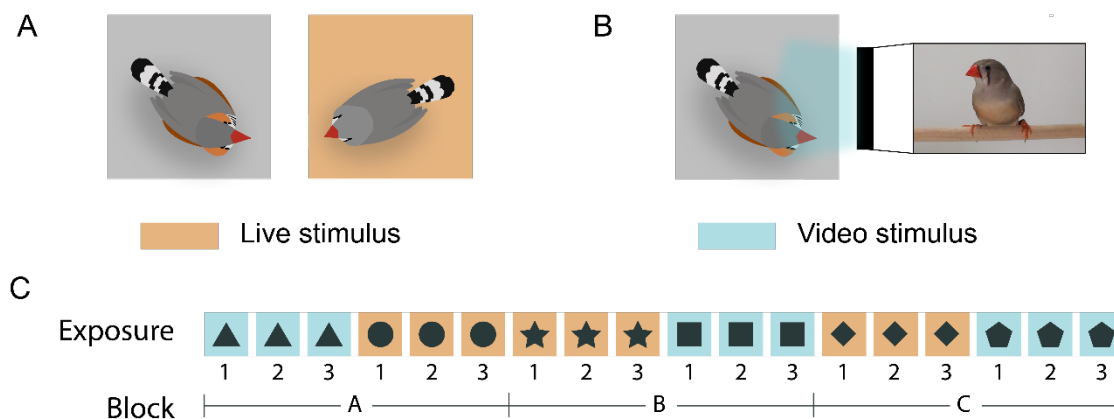
128 Stimulus females were videotaped using a SONY DCR-SR 220 HD camcorder at 60  
129 frames per second. We gathered footage of individual females perched at camera-level  
130 in front of a neutral background. Adobe Premiere 2017 was used for minor white  
131 balance corrections, cropping and trimming. Playback clips were 30 seconds in  
132 duration and featured a silent, perched female engaged in a moderate level of activity  
133 (e.g., movements of head and along perch but no flying; Movie 1). A total of six females  
134 were filmed with three clips created per individual.

135

### 136 *Behaviour testing and song collection*

137 Figure 1 illustrates the experimental setup used during song collection. Male finches  
138 were isolated in individual cages (20 x 20 x 20 cm) inside sound-attenuating chambers  
139 (“soundboxes”; TRA Acoustics, Ontario, Canada) from at least one day prior to  
140 experiments. All songs were recorded using an omnidirectional microphone  
141 (Countryman Associates, Inc., Menlo Park, CA, USA) positioned directly above the

142 male's cage. During experiments, song was detected and digitized using a sound-  
 143 activated system (Sound Analysis Pro v.1.04  
 144 [http://ofer.sci.ccny.cuny.edu/html/sound\\_analysis.html](http://ofer.sci.ccny.cuny.edu/html/sound_analysis.html), digitized at 44.1 kHz). A  
 145 Microsoft Surface Pro 3 tablet (2160 x 1440 pixels) was used to playback videos and  
 146 was fixed to a wall of the soundbox. The tablet was placed in the soundbox at least 10  
 147 minutes before the onset of testing and was positioned ~12 cm from the male's cage.  
 148 We sized the video playback window such that stimulus bird in the video was  
 149 approximately life-size at the distance between the cage and tablet. The screen was  
 150 blank (black) when not displaying video stimuli. A camera mounted above the tablet  
 151 provided a live stream of the experimental bird for monitoring. All experiments began  
 152 within 2 hours of lights turning on.



153

154 **Figure 1: Experimental designs for live and video presentations of females.**  
 155 *Schematics represent experimental setup during (A) exposure to live female stimulus*  
 156 *and (B) exposure to videos of females (not drawn to scale). (C) Example of stimulus*  
 157 *presentation order for one male. Males were tested in three blocks (A, B, C), with each*  
 158 *block consisting of three consecutive exposures to distinct videos of an individual*  
 159 *female and three consecutive live exposures to an individual female. All presentations*  
 160 *were separated by a five-minute interval. Shapes represent stimulus individuals.*  
 161

162 During experiments, we collected courtship songs from male zebra finches using a  
 163 design similar to that described by Toccalino et al. (2016). Specifically, each male was  
 164 briefly exposed to six different females, three via live presentations and three via video

165 presentations. During live female presentations, an experimenter opened the soundbox  
166 door and placed a cage housing a conspecific female next to the experimental male's  
167 cage. The soundbox door was then closed, and the female remained in the soundbox  
168 for 30 seconds. During video presentations of females, an experimenter opened the  
169 soundbox door, started a video of a female, and then closed the door to the soundbox.  
170 Videos played for 30 seconds and ended on a black screen.

171  
172 Males were exposed to a total of 6 randomly-chosen stimulus females from a pool of 6  
173 videotaped females and 12 live females. Females that were videotaped were distinct  
174 from those used for live presentations. Video and live presentations were grouped into  
175 three blocks (blocks A-C; Figure 1C), with each block consisting of three consecutive  
176 exposures to either a video or live presentation of an individual female (exposures 1-3),  
177 followed by three consecutive exposures to the other stimulus type. Within each block  
178 of video presentations, males were exposed to distinct video clips of the same female.  
179 All presentations were separated by five-minute intervals. The order of conditions  
180 (video vs. live presentation) within a block was pseudo-randomly determined to  
181 balance the order of conditions. The first condition presented in each block was  
182 determined by a coin flip, and, if the first conditions of the first two blocks were the  
183 same (e.g., video first for blocks A and B), the order was reversed for the last block,  
184 ensuring that no experimental session consisted of blocks that each started with the  
185 same condition.

186  
187 We categorized a male's song as directed toward the live or video presentation of a  
188 female if at least two of the following conditions were met during song production: (1)  
189 the male approached or oriented toward the stimulus females; (2) the male fluffed his  
190 plumage; and (3) the male pivoted his body from side to side (James and Sakata, 2015;  
191 Kao and Brainard, 2006; Morris, 1954; Toccalino et al., 2016). Typically, male zebra  
192 finches produce courtship song within a few seconds of stimulus presentation. Males  
193 in this study each produced a minimum of three courtship songs to video or live

194 presentations of females ( $18.0 \pm 3.3$  and  $10.0 \pm 2.1$  song bouts per male, respectively,  
195 toward live and video presentations of females).

196

197 We also collected non-courtship, or undirected (UD) songs (i.e., songs produced  
198 spontaneously when alone) during the experiment to contrast with courtship songs.

199 Undirected songs were generally produced during the five-minute intervals between

200 female exposures. In cases where few UD songs were produced between female

201 presentations, UD songs produced in the 30 minutes before and after the testing

202 period were used for analysis ( $16.2 \pm 2.9$  UD song bouts per male)(e.g. James and

203 Sakata, 2015; Sakata et al., 2008; Tocalino et al., 2016).

204

### 205 *Song analysis*

206 We used the following definitions for our analyses (Figure 2). “Song bouts” are defined

207 as epochs of singing that are separated by at least 1 s of silence (e.g. Johnson et al.,

208 2002; Poopatanapong et al., 2006). Each song bout consists of a stereotyped

209 sequence of vocalizations called a “motif” that is repeated throughout the bout

210 (Sossinka and Böhner, 1980; Zann, 1996). Motifs consist of distinct vocal elements

211 (“syllables”) that are separated by at least 5 ms of silence. The first motif of a bout is

212 preceded by repetitions of brief vocal elements called “introductory notes.”

213

214 Our primary measure of courtship song motivation was the total amount of time

215 (seconds) males engaged in courtship song across all exposures to live or video

216 presentations of females (“time spent singing”). We also deconstructed this measure

217 into various components, including the likelihood that males will produce courtship

218 song on a given exposure and the total duration of song during each exposure. We

219 also broke down the total song duration during each exposure into the number of

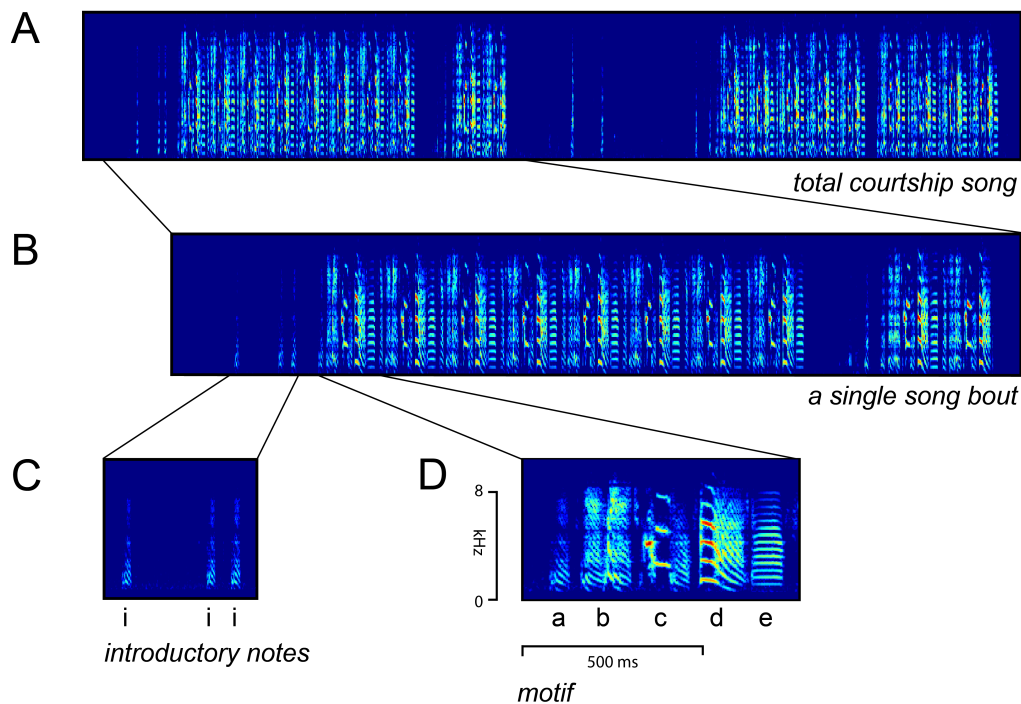
220 bouts produced during each exposure, and the duration of each of those bouts. Bout

221 durations were defined as the interval between the onset of the first syllable to the

222 onset of the last syllable of the bout.

223





224

225 **Figure 2:** Organization of zebra finch song. Spectrograms plot the frequency (kHz)  
226 against time (ms) with brightness reflecting amplitude. (A) An example of song  
227 produced during a 30-sec exposure to a video. In this example, the male produced two  
228 song bouts (bouts separated by  $\geq 1$  second of silence). (B) Zoomed-in image of a single  
229 song bout; song bouts begin with the repetition of introductory notes (C) and consist of  
230 a stereotyped sequence of syllables (“motif”) that is repeated throughout the bout (D).  
231 The motif of this bird consists of five separate syllables, each separated by  $\geq 5$  ms of  
232 silence.

233

234 We analysed song features that are consistently affected by social stimuli and that  
235 have been used as indices of song performance (Sakata and Vehrencamp, 2012). In  
236 particular, we measured the number of introductory notes preceding song, song  
237 tempo, and the variability of the fundamental frequency (FF) of syllables with flat,  
238 harmonic structure (Chen et al., 2016; Cooper and Goller, 2006; James and Sakata,  
239 2014; Kao and Brainard, 2006; Sakata et al., 2008; Stepanek and Doupe, 2010). For  
240 these analyses, we first manually labelled syllables and introductory notes following  
241 amplitude-based element segmentation using custom software written in MATLAB  
242 (The MathWorks, Natick, MA, USA). Introductory notes were quantified by starting with

243 the note immediately preceding the first syllable of the bout and counting backwards  
244 until we reached  $\geq 1$  second of silence. Motif duration was defined as the duration from  
245 the onset of the first syllable of the motif to the onset to the last syllable of the motif  
246 and was used as the metric for song tempo (e.g., James and Sakata, 2015; Kao and  
247 Brainard, 2006; Sakata et al., 2008). We restricted the analysis of song tempo to the  
248 first motif of the bout, because motif durations have been found to change across the  
249 song bout and because bout durations differ between live and video presentations (see  
250 Results; Chi and Margoliash, 2001; Cooper and Goller, 2006; Glaze and Troyer, 2006).  
251 Finally, we computed the fundamental frequency (FF) of syllables with flat, harmonic  
252 structure (e.g., syllables 'c,' 'd,' and 'e' in Figure 2D) by calculating the autocorrelation  
253 of a segment of the sound waveform and measuring the distance (in Hz) from the zero-  
254 offset peak to the highest peak in the autocorrelation function. We measured the FF for  
255 every rendition of the syllable, and then computed the coefficient of variation (CV:  
256 standard deviation/mean) per syllable per condition as an index of acoustic variability  
257 (Sakata et al., 2008; Toccalino et al., 2016). We also computed these same measures  
258 of song performance for UD songs to contrast with performance for songs directed at  
259 videos of females [video-directed (VD) song] and songs directed at live females [live-  
260 directed (LD) songs].

261

### 262 *Data Analysis*

263 We compared song motivation between experimental conditions for all 13 males.  
264 However, five males produced courtship songs only during live presentations of  
265 females; therefore, in our direct comparisons of song motivation and performance  
266 during live and video presentations of females, data were restricted to the eight males  
267 that produced songs during both live and video presentations of females. Data for a  
268 number of song features were computed for each exposure (total song duration per  
269 exposure, bout duration, first motif duration, introductory notes and fundamental  
270 frequency), and for these data, we only analysed exposures during which song was  
271 produced.

272

273 Statistical analyses were conducted in R 2.15.1. We used linear mixed models (LMMs)  
274 and generalized linear mixed models (GLMMs) within the ‘lme4’ library (Bates et al.,  
275 2015) to compare singing behaviour across experimental conditions. Our experimental  
276 design consisted of three testing blocks (blocks A-C), with each block consisting of  
277 three consecutive exposures to videos of a single female and three consecutive  
278 exposures to a live female (exposures 1-3; Figure 1C). Therefore, we ran three-way  
279 factorial models with Block (A-C; ordinal), Exposure (1-3; ordinal), Condition (live vs.  
280 video; nominal) and all possible interactions as fixed effects. Because of the repeated-  
281 measures nature of this design, we also included Bird ID as a random factor. Further,  
282 because birds can produce multiple bouts within an exposure and because bout  
283 number can affect some song features (see Results), we also ran four-way full-factorial  
284 models with the same three fixed effects plus Bout (1-3; ordinal).

285  
286 In the analysis of the likelihood to produce courtship song, we had one binary  
287 response variable (whether the bird produced at least one courtship song bout or not  
288 during each exposure); therefore, we ran this model as a GLMM with a binomial error  
289 family. The number of song bouts produced during each exposure and the number of  
290 introductory notes preceding song bouts (see above) were count responses;  
291 consequently, we ran these models as GLMMs with a Poisson error family. Total time  
292 spent singing, total song duration per exposure, and song bout durations were highly  
293 skewed; therefore these data were analysed with a gamma error family and a log link  
294 (data plotted following log-transformation for ease of presentation). Finally, the total  
295 number of exposures in which at least one song bout was produced and the duration  
296 of the first motif were analysed with LMMs with a Gaussian error family. Prior to  
297 running the statistical models, data were visually screened to assess model fit using Q-  
298 Q plots. To test the significance of each mixed model, we ran Type II Wald chi-square  
299 tests using the ‘car’ library (Fox et al., 2011).

300  
301 We used a different statistical model to analyse experimental variation in the CV of FF.  
302 This is because the CV (standard deviation divided by the mean) cannot be computed

303 for a single rendition and needs to be computed across multiple renditions of syllable  
304 (or motif). To provide reliable estimates of the CV of the FF of a particular syllable, we  
305 measured the CV across each rendition of a syllable for all songs produced during all  
306 blocks and exposures. The statistical model to analyse variation in the CV of FF also  
307 differs from the ones described above because birds can produce multiple syllables for  
308 which we calculated the FF ( $n=13$  syllables across the eight males). Consequently, for  
309 the analysis of the CV of FF, we ran a LMM with a Gaussian error family, and with  
310 Condition as the fixed effect and Syllable ID nested in Bird ID as a random effect so  
311 that we could directly compare the CV of the same syllable across conditions.

312  
313 In addition to assessing differences in song performance across VD and LD songs, we  
314 also compared song performance of VD and LD songs with those of UD songs. For  
315 these analyses, we computed data for VD and LD song across all renditions of song  
316 (i.e., across blocks, exposures, and bouts) and compared these values to those for UD  
317 song. We ran one-way models with similar parameterization as before, with Condition  
318 as the sole independent variable. We used a Poisson error family for introductory notes  
319 and a Gaussian error family for first motif durations and CV of FF. Bird ID was a  
320 random variable for introductory notes and first motif durations, and Syllable ID nested  
321 in Bird ID was a random effect for CV of FF. For these analyses, we ran Tukey's tests  
322 with the Holm correction using the 'multcomp' library (Hothorn et al., 2008) for post-  
323 hoc contrasts across the three conditions. As with the previous analysis, we nested  
324 Syllable ID within Bird ID as the random effect in the analysis of the CV of FF.

325  
326 To gain further insight into the relationship between song changes across experimental  
327 conditions, we also analysed the extent to which motivational and performance  
328 changes driven by video presentations of females co-varied with motivational and  
329 performance changes driven by live presentations of females. We correlated the total  
330 amount of song a male produced to live presentations of females with the total amount  
331 of song a male produced to video presentations of females. In addition, we computed  
332 the percent change of song features from UD to VD song and from UD to LD song, and

333 correlated these percent changes. Because we measured the CV of FF of multiple  
334 syllables, and each syllable within a bird could change independently, we ran these  
335 correlations as LMMs with the Gaussian error family with Bird ID as a random effect. All  
336 other relationships were analysed using Pearson's product-moment correlations.

337  
338 Finally, we analysed the extent to which individual variation in the differential motivation  
339 to produce courtship song to video and live presentations of females covaried with  
340 individual variation in the differential modulation of song performance across video and  
341 live presentations. Specifically, we correlated the difference in time spent singing LD  
342 and VD song with the difference in the modulation of each song feature from UD to LD  
343 and UD to VD. We used Pearson's correlations for analyses of the number of  
344 introductory notes and the first motif duration. Because the differential motivation to  
345 produce courtship song to videos or live presentations of females is summarised by  
346 one value per bird, we calculated the average change in the CV of FF across all  
347 syllables produced by each bird to relate performance to motivation (i.e., each bird has  
348 only one data point representing the average percent change in the CV of FF of  
349 syllables).

350

## 351 **RESULTS**

352 *Differences in the motivation to produce courtship song to video vs. live presentations*  
353 *of females*

354 We first counted, for each male (n=13), the number of exposures to live females or  
355 videos of females in which a male produced at least one courtship song (out of nine  
356 exposures per male for each condition). We found that male zebra finches produced  
357 courtship song on significantly more exposures to live females ( $6.6 \pm 0.7$  (out of 9);  
358 mean  $\pm$  SEM) than to videos of females ( $3.6 \pm 0.9$  (out of 9);  $\chi^2_1=19.0$ ,  $p<0.0001$ ). Upon  
359 further inspection, we noted that, while most birds produced courtship songs to both  
360 live and video presentations of females (n=8), five birds sang exclusively towards live  
361 females (no males sang exclusively towards videos of females). However, the number  
362 of exposures with courtship song remained significantly higher for live presentations of

363 females even when analyses were restricted to males that produced courtship song to  
364 both video and live presentations of females (live:  $7.8 \pm 0.8$ ; video:  $5.9 \pm 0.8$ ;  $\chi^2_1=8.6$ ,  
365  $p=0.0034$ ).

366  
367 The total amount of time a male spends singing towards a female is widely considered  
368 a reliable measure of song motivation. As such, we compared the total amount of time  
369 male zebra finches sang to live and video presentations of females (i.e., total duration  
370 of song across all exposures). Because this study is focused on experimental  
371 differences in motivation and performance features and because performance features  
372 of video-directed songs cannot be computed for males that did not sing to videos of  
373 females, we limited our analyses to birds that sang in both conditions ( $n=8$ ). Overall, we  
374 found that birds produced significantly more song towards live females ( $116 \pm 18$   
375 seconds) than to videos of females ( $46 \pm 8$  seconds;  $\chi^2_1=54.5$ ,  $p<0.0001$ ; Figure 3A;  
376 Table S1).

377  
378 Differences in the total amount of song produced to live vs. video presentations of  
379 females could be caused by a number of factors, including variation in the probability  
380 of producing courtship song on each exposure, in the total amount of song on each  
381 exposure, in the number of song bouts produced on each exposure, and in bout  
382 durations. We first analysed whether males differed in their probability of producing  
383 courtship song towards live or video presentations of females on each individual  
384 exposure to a female stimulus. We performed a 3-way GLMM with Condition (live or  
385 video), Block (A-C) and Exposure (1-3; ordinal) as independent factors, Courtship (0 or  
386 1; binomial) as the response variable, and Bird ID as a random effect (all birds included  
387 in this analysis). We found significant effects of Condition ( $\chi^2_1=13.8$ ,  $p=0.0323$ ) and  
388 Block ( $\chi^2_2=21.1$ ,  $p=0.0122$ ), indicating that birds were significantly more likely to  
389 produce courtship song to live females than to videos of females and that the  
390 likelihood of a male producing courtship song decreased across blocks (Figure 3B). In  
391 addition, there was a marginally significant interaction between Condition and Block  
392 ( $\chi^2_2=7.5$ ,  $p=0.0588$ ), with differences across conditions being larger for later blocks.

393

394 To further reveal the factors that contributed to the overall difference in the amount of  
395 courtship song produced to live vs. video presentations of females, we examined the  
396 total amount of courtship song produced (in seconds) during each video or live  
397 exposure to a female (each exposure to a female stimulus was 30 sec in duration;  
398 Table S1). We found significant effects of Condition ( $\chi^2_1=69.6$ ,  $p<0.0001$ ), Block  
399 ( $\chi^2_2=52.0$ ,  $p<0.0001$ ), and Exposure ( $\chi^2_2=27.8$ ,  $p<0.0001$ ) on total song duration per  
400 exposure (Figure 3C). Overall, song durations per exposure were shorter in response to  
401 video presentations of females than live presentations, and durations decreased across  
402 blocks and exposures. In addition, there was a significant interaction between  
403 Condition and Block ( $\chi^2_2=20.3$ ,  $p<0.0001$ ), which was characterized by smaller changes  
404 in song durations across blocks for live presentations than for video presentations. As  
405 such, the difference in song durations between video and live exposures became larger  
406 over the blocks of testing.

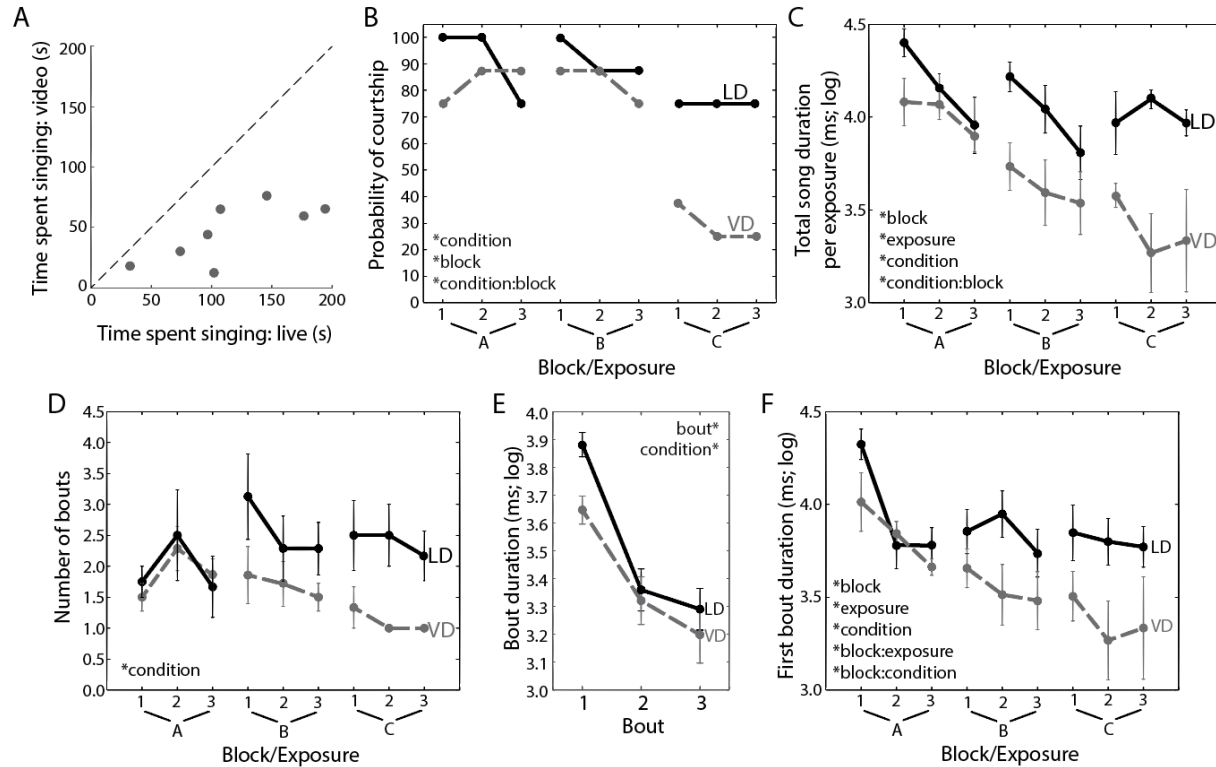
407

408 Because birds produce courtship songs in bouts (i.e., epochs of song separated by  $\geq 1$   
409 sec of silence), differences in courtship song duration per exposure could be due to  
410 differences in the number of song bouts produced during each exposure as well as  
411 differences in the lengths of song bouts. Consequently, we first analysed the number of  
412 bouts that male zebra finches produced on each exposure (Supplementary Table S1).  
413 We found a significant effect of Condition ( $\chi^2_1=5.2$ ,  $p=0.0221$ ; Figure 2D), with males  
414 producing fewer song bouts per exposure to videos of females (mean  $\pm$  SEM:  $1.70 \pm$   
415  $0.12$  bouts) than per exposure to a live female ( $2.32 \pm 0.18$  bouts). While the interaction  
416 between Condition and Block was not statistically significant, visual inspections of the  
417 data indicate a trend for the difference between video and live presentations to  
418 become larger over the blocks of testing.

419

420





421

422 **Figure 3: Male zebra finches produced courtship song to both video and live**  
 423 **presentations of females but were less motivated to produce song to video**  
 424 **presentations of females.** For panels (B-F), data for video-directed (VD) and live-  
 425 directed (LD) songs are plotted, respectively, with dashed grey and solid black lines. All  
 426 error bars are the standard error of the mean. (A) Birds produced more song towards  
 427 live females than towards videos of females. Each dot represents an individual bird, and  
 428 the dashed line depicts the line of unity. (B) The probability of courtship per exposure is  
 429 affected by condition, block and a marginally significant interaction between condition  
 430 and block. (C) Total song duration per exposure is affected by condition, block,  
 431 exposure and the interaction between condition and block. (D) The number of song  
 432 bouts produced during each exposure to a female is only affected by condition. (E-F)  
 433 Song bout duration is affected by bout, block, exposure, condition and the interaction  
 434 between block, exposure and bout (see Results). For ease of presentation, we first  
 435 depict the effects of bout and condition (E), followed by the effects of block, exposure  
 436 and condition on the duration of the first bout produced during each exposure (F).

437

438

439 To analyse song bout duration, we ran a four-way factorial mixed effects model with  
 440 the same fixed factors as above (Condition, Block and Exposure) as well as Bout (i.e.,  
 441 the serial order of bouts within each exposure; ordinal). Because birds rarely produced  
 442 more than three bouts in an exposure, we limited our analysis to the first three bouts



443 per exposure (models were rank deficiency when data for all bouts were included;  
444 Table S1). We found significant effects for all main factors (Condition:  $\chi^2_1=40.3$ ,  
445  $p<0.0001$ ; Block:  $\chi^2_2=46.5$ ,  $p<0.0001$ ; Exposure:  $\chi^2_2=21.0$ ,  $p<0.0001$ ; Bout:  $\chi^2_6=129.6$ ,  
446  $p<0.0001$ ), as well as a three-way interaction between Block, Exposure, and Bout  
447 ( $\chi^2_8=36.5$ ,  $p<0.0001$ ) and between Exposure, Condition and Bout ( $\chi^2_4=9.8$ ,  $p=0.0431$ ;  
448 Figure 3E). We also observed a significant interaction between Block and Exposure  
449 ( $\chi^2_4=10.6$ ,  $p=0.0318$ ) and a marginal interaction between Block and Bout ( $\chi^2_4=8.4$ ,  
450  $p=0.0795$ ). Overall, bout durations were longer for songs produced to live  
451 presentations of females than for songs produce to video presentations of females.  
452 Additionally, bout durations decreased across blocks, across exposures within blocks,  
453 and across bouts produced within each exposure to a female stimulus.

454  
455 Because of the complexity of the four-way model, we conducted another analysis  
456 limited to the data from the first bout (i.e., Bout not included as an effect in the model)  
457 to obtain a simplified depiction of variation in song bout duration (Table S1). We  
458 observed significant main effects for all three factors (Condition:  $\chi^2_1=38.7$ ,  $p<0.0001$ ;  
459 Block:  $\chi^2_2=34.1$ ,  $p<0.0001$ ; Exposure:  $\chi^2_2=30.1$ ,  $p=0.0001$ ) as well as significant  
460 interactions between Block and Exposure ( $\chi^2_4=18.3$ ,  $p=0.0011$ ) and between Block and  
461 Condition ( $\chi^2_2=9.9$ ,  $p=0.0071$ ; Figure 3F). Overall, the duration of the first bout of  
462 courtship song was longer for songs produced to live presentations of females than for  
463 songs produced to video female presentations, and bout durations became shorter  
464 across blocks and exposures. The interactions were characterized by larger decreases  
465 across exposures during block A than during blocks B and C, and by larger decreases  
466 across blocks for video presentations than for live presentations.

467  
468 Together, these analyses indicate that differences in total amount of courtship song in  
469 response to video and live presentations of females were due to differences in the  
470 likelihood of producing courtship song, the number of song bouts per exposure, and  
471 the duration of individual song bouts.

472

473 Despite differences in the amount of courtship song produced to live vs. video  
474 presentations of females, it is possible that individual variation in the motivation to  
475 produce courtship songs to videos of females is related to variation in the motivation to  
476 produce courtship songs to live presentations of females. Therefore, we correlated  
477 individual variation in the total amount of time males spent singing to live and video  
478 presentations of females. Consistent with the notion that motivation to court videos of  
479 females is related to the motivation to court live females, we found a significant  
480 correlation between the total amount of song produced to live versus video stimuli  
481 ( $n=8$ ;  $r=0.73$ ,  $p=0.0382$ ).

482

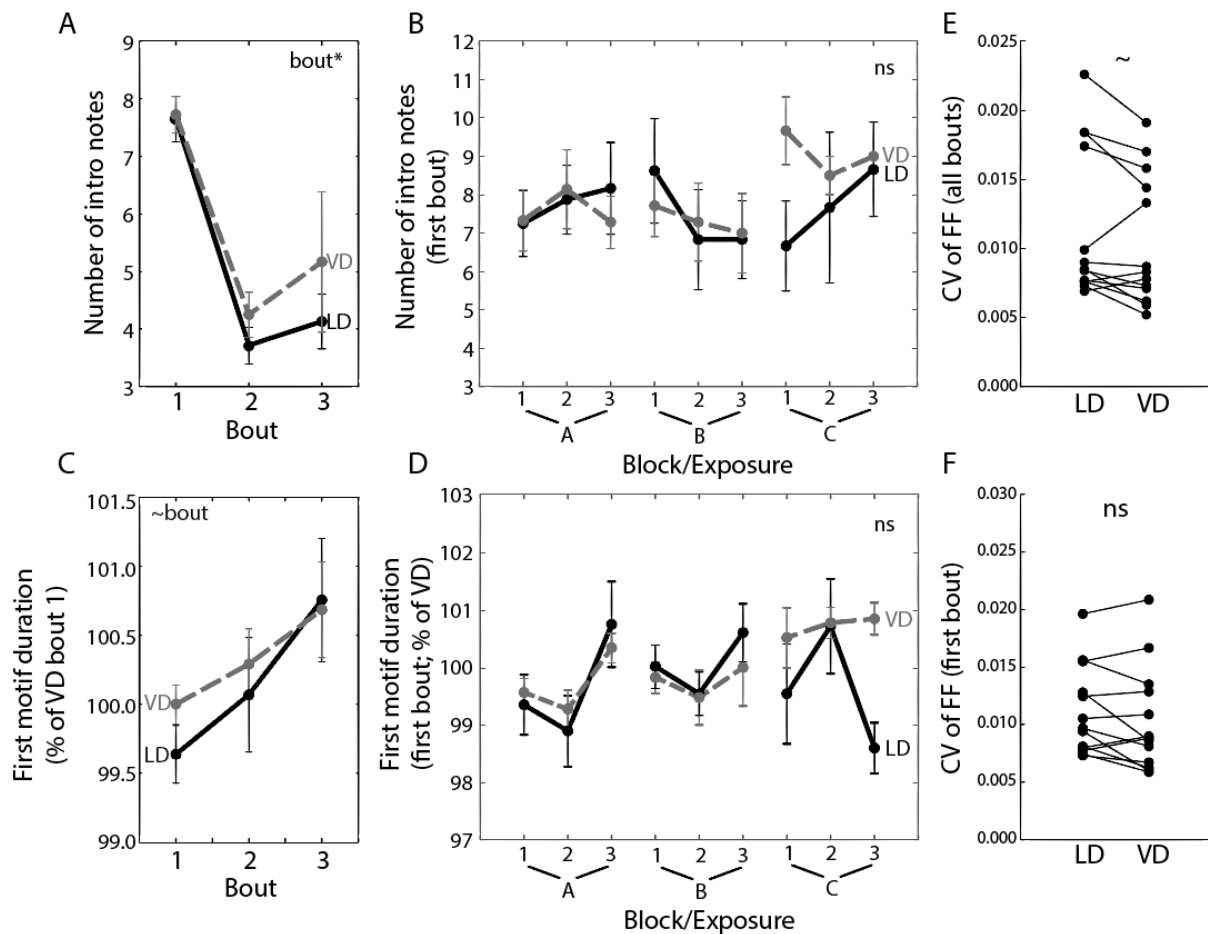
483 *Lack of differences in performance features of courtship songs produced to video vs.*  
484 *live presentations of females*

485 Our results support that zebra finches are less motivated to court videos of females  
486 than live females, and we next sought to determine whether performance aspects also  
487 varied across songs directed at live or video presentations of females. To this end, we  
488 compared various measures of song performance (see Methods) between video-  
489 directed (VD) and live-directed (LD) songs among males that produced both types of  
490 songs ( $n=8$  birds).

491

492 To analyse differences in the number of introductory notes preceding song, we first ran  
493 a four-way factorial model with Condition, Block, Exposure and Bout (limited to the first  
494 three bouts; see above) as fixed effects, Bird ID as a random factor, and the number of  
495 introductory notes before each bout of song as a Poisson response variable (Table S2).  
496 Importantly, we found no significant effect of Condition or interaction between  
497 Condition and other variables for the number of introductory notes. We only observed  
498 an effect of Bout ( $\chi^2_6=123.9$ ,  $p<0.0001$ ), with the number of introductory notes  
499 decreasing across consecutive bouts produced during an exposure to a stimulus  
500 (Figure 4A). We also ran a similar analysis with data limited to the first bout of song on  
501 each exposure (Bout excluded as a factor) and, again, found no significant variation  
502 across conditions, blocks, and exposures (Figure 4B).

503



504

505 **Figure 4: Male zebra finches produce video-directed (VD) and live-directed (LD)**  
 506 **courtship songs with similar song features.** For panels (A-D), data for VD and LD  
 507 songs are plotted with dashed grey and solid black lines, respectively. For ease of  
 508 presentation we first depict the effects of bout and condition for the first three bouts (A,  
 509 C), followed by the effects of block, exposure and condition on a data set limited to the  
 510 first bout produced (B, D). All error bars are the standard error of the mean. (A) The  
 511 number of introductory notes is affected by bout, with no significant difference between  
 512 VD and LD song. (B) The number of introductory notes before the first bout of song is  
 513 not affected by block, exposure or condition. (C) There was a marginal effect of bout on  
 514 the first motif duration, but no significant effect of condition. Data in the figure are  
 515 normalized to the mean duration of VD song for data visualization purposes (see  
 516 Methods). (D) The first motif duration of the first bouts of song produced during an  
 517 exposure was not significantly affected by block, exposure or condition. (E, F) The CV  
 518 of FF was marginally lower for VD song than LD song when all bouts were included in  
 519 the analysis (E), but there was no significant difference when only the first bouts per  
 520 exposure were analysed (F). For (E, F) each pair of connected dots depicts one syllable  
 521 within a bird (n=13 syllables across the eight birds).  
 522

523 To analyse variation in song tempo between VD and LD song, we calculated the  
524 duration of the first motif of each song bout and analysed experimental variation in first  
525 motif durations using the same four-way factorial model as above (only first three  
526 bouts). Only the first motif in each bout was analysed for this comparison because  
527 motif durations change as bout length increases (e.g., Chi and Margoliash, 2001; Glaze  
528 and Troyer, 2006; James and Sakata, 2014; James and Sakata, 2015) and because  
529 bout lengths differed between VD and LD song (Figure 3; Table S2). There was no  
530 significant effect of any factor, including Condition, on song tempo (Figure 4C). We  
531 also ran a three-way factorial model using only data from the first bout produced per  
532 exposure and, again, found no significant effects (Figure 4D).

533  
534 The FF of syllables with flat, harmonic structure is less variable from rendition-to-  
535 rendition when males direct song at females (Sakata and Vehrencamp, 2012; Woolley  
536 and Kao, 2015). We calculated the CV of FF across all syllable renditions in every bout  
537 of song and compared this variability between conditions (i.e., Condition is the only  
538 independent variable). We found a marginally significant difference between Conditions  
539 ( $\chi^2_1=3.7$ ,  $p=0.0545$ ; Figure 4E) with VD song tending to have lower CVs than LD song.  
540 However, no significant difference between VD and LD song was observed when only  
541 data from the first bout were analysed ( $\chi^2_1=2.3$ ,  $p=0.1266$ ; Figure 4F). This difference in  
542 the magnitude of differences between VD and LD song is primarily due to a decrease in  
543 the CV of FF for LD song.

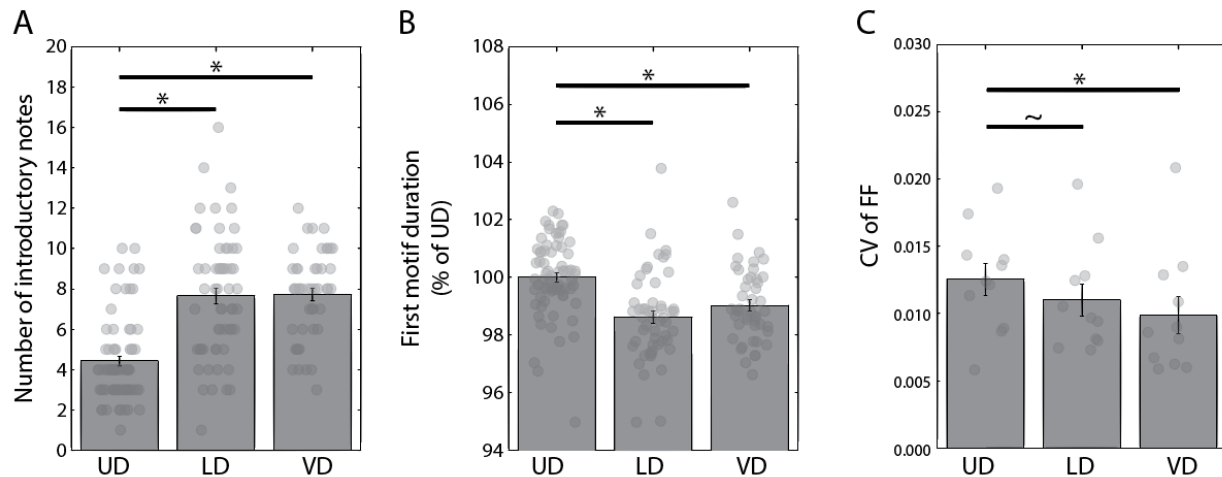
544  
545 *Courtship songs produced to video or live presentations of females are distinct in*  
546 *performance from undirected song*

547 Overall, the preceding analyses indicate a lack of difference between VD and LD songs  
548 for three performance measures: the number of introductory notes, song tempo, and  
549 spectral variability. However, these analyses do not indicate whether males alter VD  
550 songs to be distinct from non-courtship songs (undirected or UD songs) in the same  
551 way that LD songs differ from UD songs. We therefore compared performance  
552 measures of every VD and LD song of a male to all his UD songs. We found a

553 significant effect of Condition for introductory notes ( $\chi^2_2=51.3$ ,  $p<0.0001$ ) with post-hoc  
554 contrasts indicating that both VD and LD songs were preceded by more introductory  
555 notes than UD songs and that VD songs were preceded by more introductory notes  
556 than LD songs ( $p<0.003$  for all). We also found a significant effect of Condition on first  
557 motif duration ( $\chi^2_2=39.7$ ,  $p<0.0001$ ), with post-hoc contrasts indicating that motifs were  
558 shorter in VD and LD songs relative to UD songs ( $p<0.0001$  for both). Finally, we found  
559 a significant effect of Condition on the CV of FF ( $\chi^2_2=8.8$ ,  $p=0.01201$ ) with post-hoc  
560 contrasts indicating that the CV of FF was lower for VD songs compared to UD songs  
561 ( $p=0.0089$ ).

562  
563 The preceding analyses indicated that VD songs were distinct from UD songs, with  
564 mixed results regarding LD songs. However, our analyses above highlight how  
565 performance features can change across bouts and how the number of bouts  
566 produced per exposure differed between video and live presentations of females  
567 (Figure 4); consequently, the previous results are confounded by experimental variation  
568 in the number of bouts per exposure. To examine variation without this confound, we  
569 conducted the same analyses with data restricted to the first bout of song per  
570 exposure. In addition, we limited our UD song data to songs preceded by at least 30 s  
571 of silence to approximate the first bout restriction for VD and LD songs (see Methods).  
572 The number of introductory notes was significantly affected by Condition ( $\chi^2_2=67.4$ ,  
573  $p<0.0001$ ; Figure 5A), and this difference was due to the fact that VD and LD songs  
574 were preceded by more introductory notes than UD song ( $p<0.0001$  for each). The  
575 duration of the first motif was significantly different across Conditions ( $\chi^2_2=34.3$ ,  
576  $p<0.0001$ ; Figure 5B) and was shorter during VD and LD songs than during UD song  
577 ( $p<0.0002$  for each). Finally, the CV of FF was also affected by Condition ( $\chi^2_2=12.2$ ,  
578  $p=0.0023$ ; Figure 5C), with the CV of FF being significantly lower during VD song than  
579 during UD song ( $p=0.0015$ ) and with the same trend for LD song ( $p=0.0875$ ).  
580 Consistent with the analysis described above, all these song features were not  
581 significantly different between VD and LD songs. Taken together, these data indicate  
582 that males change their song performance when directing songs at video presentations

583 of females and that the nature and degree of these changes are comparable between  
584 VD and LD songs. These data also suggest that shared mechanisms could mediate the  
585 modulation of courtship song in response to video and live presentations of females.  
586



587

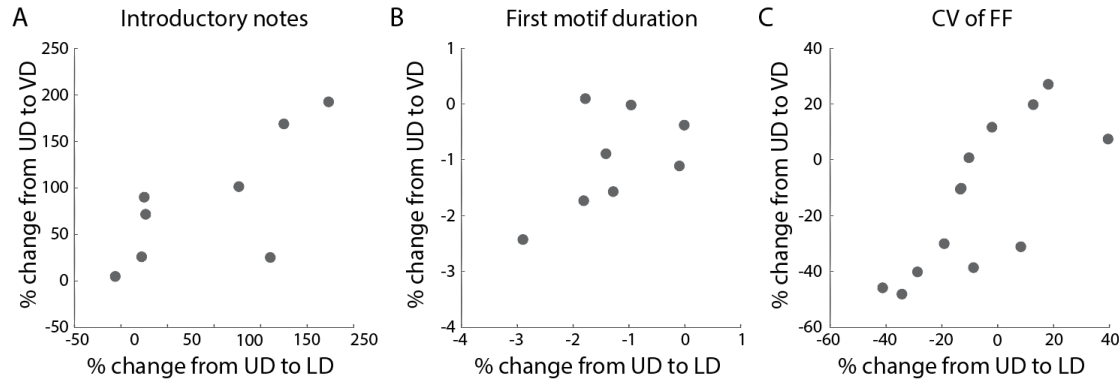
588 **Figure 5: Temporal and spectral features of video-directed (VD) and live-directed**  
589 **(LD) songs were significantly different from those of undirected (UD) songs but**  
590 **not different from each other.** (A) The number of introductory notes was significantly  
591 greater before VD and LD song than before UD song. (B) First motif durations were  
592 significantly shorter during both VD and LD song than during UD song. (C) The  
593 coefficient of variation (CV) of fundamental frequency (FF) was lower during VD and LD  
594 song than during UD song. Data were taken from the first bout of song for VD and LD  
595 song. “\*” denotes  $p < 0.05$ , and “~” denotes  $p < 0.10$  (Tukey’s HSD test with Holm  
596 correction).

597

598 To further investigate whether the modulation of VD and LD songs could be mediated  
599 by similar neural mechanisms, we correlated individual variation in the magnitude of  
600 change in song performance from UD song to VD song with individual variation in the  
601 magnitude of change from UD song to LD song (data from first bout only). If shared  
602 mechanisms underlie the modulation of VD and LD song, we should observe significant  
603 and positive relationships between the magnitudes of song modulation for VD and LD  
604 song. The relationship was significantly positive for the number of introductory notes  
605 ( $r=0.72$ ,  $p=0.0430$ ; Figure 6A) and the CV of FF ( $\chi^2_1=14.4$ ,  $p=0.0015$ ; Figure 6C). The  
606 relationship was positive but not statistically significant for first motif duration ( $r=0.53$ ,  
607  $p=0.1766$ ; Figure 6B).



608



609

610 **Figure 6:** Relationships between the magnitudes of song changes from undirected (UD)  
611 to live-directed (LD) song (x-axis) and from UD to video-directed (VD) song (y-axis).  
612 Correlations were positive for all song features and were statistically significant for the  
613 number of introductory notes (A) and the CV of FF (C) but not for the duration of the  
614 first motif (B). For (C), each dot represents data for one syllable within a bird ( $n=13$   
615 syllables across the eight birds).

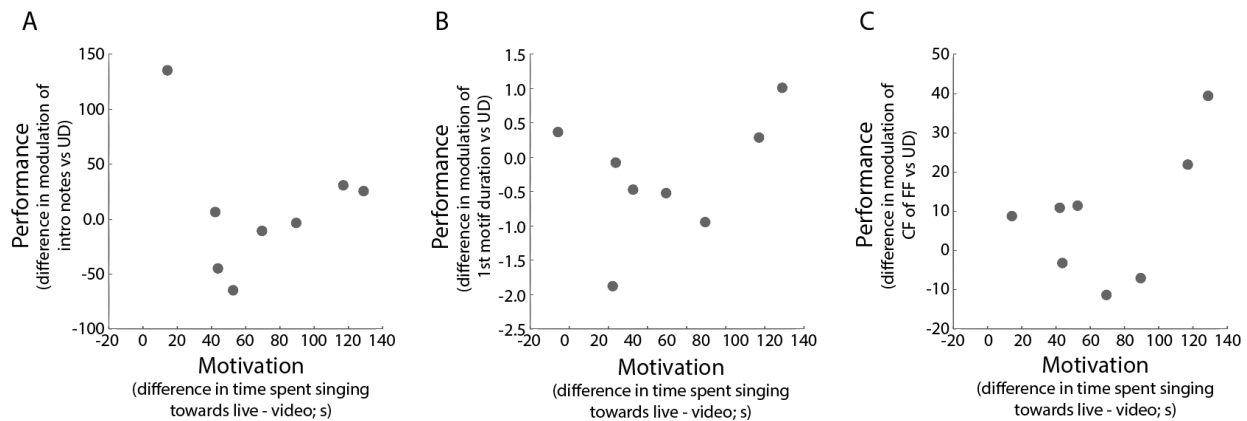
616

617

618 *Lack of relationship between experimental variation in motivational and performance*  
619 *aspects of song*

620 The lack of difference between various aspects of LD and VD song performance  
621 contrasts with the difference in the motivation to produce LD and VD song. This  
622 suggests that the motivation to produce songs to live versus video presentations of  
623 females is independent of song performance. To further investigate the relationship  
624 between motivational and performance aspects of song, we assessed whether  
625 individual variation in the differential motivation to produce courtship songs to video  
626 versus live presentations of females correlated with individual variation in the  
627 differential modulation of performance features from UD (baseline) song to VD or LD  
628 song (first bouts only). Specifically, we calculated the difference in motivation as the  
629 difference in time total time spent singing LD and VD song and correlated this  
630 difference with the difference in performance modulation, measured as the difference  
631 in percent change from UD to LD (modulation when singing LD song) and from UD to  
632 VD song (modulation when singing VD song). Overall, we observed no significant  
633 correlations between experimental variation in motivation and performance (Figure 7;

634 introductory notes:  $r=-0.15$ ,  $p=0.7219$ ; first motif duration:  $r=0.39$ ,  $p=0.3380$ ; FF of CV:  
635  $r=0.51$ ,  $p=0.1975$ ).  
636



637

638

639 **Figure 7: Relationships between measures of motivation and performance.** For all  
640 panels, individual variation in motivation (difference in time spent singing LD and VD  
641 song; x-axis) is compared to individual variation in the modulation of song performance  
642 (differences in percent change of features from UD to VD or LD; y-axis). Because each  
643 bird only has one value for difference in motivation and because we measured the CV  
644 of FF for multiple syllables within a bird's song, for this analysis we averaged the  
645 percent change in the CV of FF across syllables within each bird (i.e., one value per bird  
646 for CV of FF). Correlations were not significant for introductory notes (A), first motif  
647 durations (B) or the CV of FF (C).

648

## 649 DISCUSSION

650 Male songbirds direct songs at females as part of their courtship ritual to secure  
651 copulations. This aspect of courtship can be analysed from both motivational and  
652 performance perspectives, with the former referring to the “drive” to produce courtship  
653 song and the latter referring to the acoustic features of courtship song (e.g., song  
654 tempo and stereotypy). Both aspects of courtship song are important because deficits  
655 in either component can affect attractiveness and mating success (Gil and Gahr, 2002;  
656 Heinig et al., 2014; Sakata and Vehrencamp, 2012; Woolley and Doupe, 2008).  
657 However, little is known about the extent to which these aspects of courtship song are  
658 regulated by similar or distinct mechanisms. Indeed, because neural circuits regulating  
659 the motivation to sing project to brain areas that regulate song performance (Riters,



660 2012), it is possible that motivational and performance aspects of courtship song are  
661 linked.

662

663 Here, we took advantage of previous studies that outline experimental manipulations  
664 that affect the motivation to produce courtship song and assessed the degree to which  
665 such motivational variation was associated with variation in vocal performance.

666 Previous studies indicate that male songbirds will produce courtship songs to videos of  
667 females but tend to be less motivated to sing to videos of females than to live  
668 presentations of females (Galoch and Bischof, 2007; Ikebuchi and Okanoya, 1999;  
669 Takahasi et al., 2005). Consequently, we analysed whether the vocal performance of  
670 courtship songs produced to videos of females was distinct from songs produced to  
671 live females. Consistent with previous studies, we found that male zebra finches  
672 produced courtship songs to video presentations of female conspecifics but were less  
673 motivated to produce courtship songs to video presentations of females than live  
674 presentations of females. Specifically, males produced less than half the amount  
675 courtship song during video presentations of females than during live presentations,  
676 and this difference was due to males being less likely to produce courtship songs  
677 during video presentations and producing shorter songs when courting videos of  
678 females (Figure 3). However, courtship songs that male zebra finches produced to  
679 videos of females were structurally indistinguishable in most ways from courtship  
680 songs produced towards live females. In particular, the number of introductory notes  
681 preceding song, song tempo, and the variability of the fundamental frequency of  
682 syllables with flat, harmonic structure were not significantly different between video-  
683 directed (VD) and live-directed (LD) songs (Figure 5). Consequently, these data support  
684 the notion that the motivation to produce courtship song is controlled by independent  
685 mechanisms than the regulation of song performance, a notion that is further  
686 supported by the finding that individual variation in motivation to produce VD and LD  
687 songs was not related to individual variation in the modulation of VD and LD  
688 performance.

689

690 Such a dissociation between motivational and performance aspects of song has also  
691 been reported in previous studies (Cornil and Ball, 2010; Ritschard et al., 2011;  
692 Tocalino et al., 2016). For example, Tocalino et al. (2016) document that the  
693 familiarity of a female (i.e., repeated presentations of the same female) decreases the  
694 motivation of a male to direct courtship song to that female but does not affect  
695 performance aspects of his courtship song. In addition, Alward et al. (2013) found that  
696 testosterone implants into the medial preoptic area increased the amount of songs that  
697 male canaries produced to females but did not affect song performance measures  
698 such as song stereotypy. Collectively, these data suggest that distinct mechanisms  
699 contribute to motivational and performance aspects of birdsong and encourage  
700 experiments that further tease apart these aspects. Indeed, studies that revealed a  
701 dissociation between appetitive and consummatory aspects of copulatory behaviour  
702 (Moses et al., 1995; Pfaus et al., 1990; Ritters et al., 1998; Seredynski et al., 2013)  
703 deeply shaped perspectives on social behavioural control and inspired a range of  
704 different experiments (Balthazart and Ball, 2007; Cornil et al., 2018).

705

706 Additionally, this interpretation suggests a need to revisit or build upon existing models  
707 of song motivation and control. Catecholamine (e.g., dopamine) release from midbrain  
708 and hindbrain circuits is hypothesized to contribute to the motivation to produce  
709 courtship song. For example, individual variation in the motivation to produce courtship  
710 song is correlated with variation in the number of dopamine-synthesizing neurons in  
711 the ventral tegmental area (VTA) of male zebra finches (Goodson et al., 2009), and  
712 manipulations of catecholaminergic neurons affect the likelihood that male zebra  
713 finches will produce courtship song to females (Barclay et al., 1996; Vahaba et al.,  
714 2013). Dopaminergic neurons in the VTA and periaqueductal grey (PAG) and  
715 noradrenergic neurons in the locus coeruleus (LC) project to various brain areas that  
716 regulate song control, include the avian basal ganglia nucleus Area X and the  
717 sensorimotor nucleus HVC (Appeltants et al., 2000; Castelino and Schmidt, 2010;  
718 Hamaguchi and Mooney, 2012; Maney, 2013; Tanaka et al., 2018), and dopamine or  
719 norepinephrine release into these areas affects neural activity and song performance

720 (Cardin and Schmidt, 2004; Castelino and Ball, 2005; Ding and Perkel, 2002; Ihle et al.,  
721 2015; Leblois and Perkel, 2012; Leblois et al., 2010; Matheson and Sakata, 2015;  
722 Sasaki et al., 2006; Sizemore and Perkel, 2008; Solis and Perkel, 2006; Woolley, 2019).  
723 Taken together, this model suggests that variation in motivation should lead to  
724 variation in the amount of dopamine or norepinephrine released into areas like Area X  
725 or HVC, which should lead to variation in song performance. Our data do not support  
726 this model, suggesting that modifications or additional data are required. For example,  
727 further knowledge about the precise neural populations that regulate the repetition of  
728 introductory notes (e.g. Rajan and Doupe, 2013), song tempo (Long and Fee, 2008;  
729 Zhang et al., 2017) and the variability of fundamental frequency (reviewed in Woolley  
730 and Kao, 2015), and about the extent to which these populations receive  
731 catecholaminergic inputs would allow us to refine these models that link motivation  
732 and performance. Further, discovery of neurochemical systems that independently  
733 modulate song motivation or performance would greatly contribute to our  
734 understanding of this dissociation.

735  
736 In addition to addressing models of vocal communication and social behaviour in  
737 songbirds, our results also extend previous studies in important ways by  
738 demonstrating that video presentations of female conspecifics lead to comparable  
739 changes to song performance as live presentations of females. The lack of significant  
740 differences in acoustic features between VD and LD song (Figure 5) and the  
741 correlations in the degree of vocal modulations when males directed songs at live or  
742 video presentations of females (Figure 6) indicate that videos of females are effective at  
743 eliciting the same suite of vocal performance changes as live presentations of females.  
744 From a mechanistic perspective, these data also suggest that videos of females  
745 engage the neural circuits for song performance to a comparable extent as live  
746 presentations of females. Neural activity in the anterior forebrain pathway (AFP)  
747 regulates context-dependent changes in the variability of fundamental frequency  
748 (reviewed in Brainard and Doupe, 2013; Murphy et al., 2017; Sakata and Vehrencamp,  
749 2012; Woolley and Kao, 2015), whereas neural activity in the vocal motor pathway

750 (VMP) has been proposed regulate context-dependent changes to temporal features of  
751 songs such as song tempo or the number of introductory notes before song (Aronov  
752 and Fee, 2012; Hampton et al., 2009; Matheson et al., 2016; Murphy et al., 2017; Rajan  
753 and Doupe, 2013; Stepanek and Doupe, 2010). Our data suggest that videos of  
754 females modulate neural activity in these circuits in the same way and to the same  
755 extent as live females.

756

757 The reason for differences in the motivation to produce courtship song to live vs. video  
758 presentations of females remains unknown. One possibility is that variation in female  
759 behaviour across conditions could account for this difference. For example, females in  
760 the videos were quiet and provided no real-time feedback to courting males. In  
761 contrast, although the behaviour of female stimulus animals was not quantified, live  
762 stimulus females can occasionally vocalize or posture during exposures to males.  
763 These behaviours could serve as feedback signals to the male and affect his  
764 motivation to produce song. As such, it is possible that male zebra finches perceived  
765 the females in the video as inattentive or uninterested in the male, which could have led  
766 to the male producing fewer and shorter songs towards these females (e.g. Ware et al.,  
767 2016). Given these differences, a useful next step would be to assess how female  
768 vocalizations and movements (e.g., using different types of videos of females) influence  
769 the motivation to produce courtship song in male zebra finches.

770

771 Broadly speaking, our results support the notion that video playbacks are a powerful  
772 tool to reveal the mechanisms by which individuals alter evolutionarily important  
773 behaviours, including vocal performance (Heinig et al., 2014; Podos et al., 2009;  
774 Sakata and Vehrencamp, 2012; Woolley et al., 2014). These findings also suggest that  
775 a standardized set of video stimuli can be used to reveal neural mechanisms  
776 underlying song motivation and performance (see Supplementary Information) and  
777 provide additional impetus to evaluate how specific visual and/or auditory information  
778 regulate song motivation and performance.

779

780 **ACKNOWLEDGEMENTS**

781 We thank A. Jalayer for assistance with data analysis, and S.C. Woolley for  
782 constructive input and feedback throughout the experiment.

783

784 **COMPETING INTERESTS**

785 No competing interests declared.

786

787 **FUNDING**

788 This work was supported by funding from the National Science and Engineering  
789 Research Council (#05016 to J.T.S.); McGill University (Faculty of Science to R.F.),  
790 Canada Graduate Scholarship-Master's (CGS-M to R.F.), Fonds de recherche du  
791 Québec – Nature et technologies, Master's research scholarship (B1 to R.F.), and a  
792 Heller award (L.S.J.).

## 793 REFERENCES

- 794 **Alward, B. A., Balthazart, J. and Ball, G. F.** (2013). Differential effects of global versus  
795 local testosterone on singing behavior and its underlying neural substrate. *Proceedings of the*  
796 *National Academy of Sciences* **110**, 19573-19578.
- 797 **Appeltants, D., Absil, P., Balthazart, J. and Ball, G. F.** (2000). Identification of the  
798 origin of catecholaminergic inputs to HVC in canaries by retrograde tract tracing combined with  
799 tyrosine hydroxylase immunocytochemistry. *Journal of Chemical Neuroanatomy* **18**, 117-133.
- 800 **Aronov, D. and Fee MS.** (2012). Natural changes in brain temperature underlie  
801 variations in song tempo during a mating behavior. *PLoS ONE* **10**, e47856.
- 802 **Balshine-Earn, S. and Lotem, A.** (1998). Individual Recognition in a Cooperatively  
803 Breeding Cichlid: Evidence from Video Playback Experiments. *Behaviour* **135**, 369-386.
- 804 **Balthazart, J. and Ball, G. F.** (2007). Topography in the preoptic region: Differential  
805 regulation of appetitive and consummatory male sexual behaviors. *Frontiers in*  
806 *neuroendocrinology* **28**, 161-178.
- 807 **Barclay, S. R., Harding, C. F. and Waterman, S. A.** (1996). Central DSP-4 treatment  
808 decreases norepinephrine levels and courtship behavior in male zebra finches. *Pharmacology*  
809 *Biochemistry and Behavior* **53**, 213-220.
- 810 **Bates, D., Mächler, M., Bolker, B. and Walker, S.** (2015). Fitting Linear Mixed-Effects  
811 Models Using lme4. *Journal of Statistical Software; Vol 1, Issue 1 (2015)*.
- 812 **Bradbury, J. W. and Vehrencamp, S. L.** (2011). Principles of Animal Communication  
813 Sunderland, U.K.: Sinauer Associates.
- 814 **Brainard, M. S. and Doupe, A. J.** (2013). Translating birdsong: songbirds as a model  
815 for basic and applied medical research. *Annu Rev Neurosci* **36**, 489-517.
- 816 **Cardin, J. A. and Schmidt, M. F.** (2004). Noradrenergic Inputs Mediate State  
817 Dependence of Auditory Responses in the Avian Song System. *The Journal of Neuroscience*  
818 **24**, 7745-7753.
- 819 **Castelino, C. B. and Ball, G. F.** (2005). A role for norepinephrine in the regulation of  
820 context-dependent ZENK expression in male zebra finches (*Taeniopygia guttata*). *European*  
821 *Journal of Neuroscience* **21**, 1962-1972.
- 822 **Castelino, C. B. and Schmidt, M. F.** (2010). What birdsong can teach us about the  
823 central noradrenergic system. *Journal of Chemical Neuroanatomy* **39**, 96-111.
- 824 **Catchpole, C. K. and Slater, P. J. B.** (2008). Bird song: Biological themes and  
825 variations, 2nd ed. New York, NY, US: Cambridge University Press.
- 826 **Champagne, F. A., Francis, D. D., Mar, A. and Meaney, M. J.** (2003). Variations in  
827 maternal care in the rat as a mediating influence for the effects of environment on  
828 development. *Physiology & Behavior* **79**, 359-371.
- 829 **Chen, Y., Matheson, L. E. and Sakata, J. T.** (2016). Mechanisms underlying the social  
830 enhancement of vocal learning in songbirds. *Proceedings of the National Academy of Sciences*  
831 *of the United States of America* **113**, 6641-6646.
- 832 **Chi, Z. and Margoliash, D.** (2001). Temporal precision and temporal drift in brain and  
833 behavior of zebra finch song. *Neuron* **32**, 899-910.
- 834 **Clark, M. M., Moghaddas, M. and Galef, B. G.** (2002). Age at first mating affects  
835 parental effort and fecundity of female Mongolian gerbils. *Animal Behaviour* **63**, 1129-1134.
- 836 **Cooper, B. G. and Goller, F.** (2006). Physiological insights into the social-context-  
837 dependent changes in the rhythm of the song motor program. *J Neurophysiol* **95**, 3798-809.
- 838 **Cornil, C. A. and Ball, G. F.** (2010). Effects of social experience on subsequent sexual  
839 performance in naive male Japanese quail (*Coturnix japonica*). *Horm Behav* **57**, 515-22.



- 840 **Cornil, C. A., Ball, G. F. and Balthazart, J.** (2018). Differential control of appetitive and  
841 consummatory sexual behavior by neuroestrogens in male quail. *Hormones and Behavior* **104**,  
842 15-31.
- 843 **Ding, L. and Perkel, D. J.** (2002). Dopamine Modulates Excitability of Spiny Neurons in  
844 the Avian Basal Ganglia. *The Journal of Neuroscience* **22**, 5210-5218.
- 845 **Evans, C. S. and Marler, P.** (1991). On the use of video images as social stimuli in  
846 birds: audience effects on alarm calling. *Animal Behaviour* **41**, 17-26.
- 847 **Fleishman, L. J. and Endler, J. A.** (2000). Some comments on visual perception and  
848 the use of video playback in animal behavior studies. *Acta ethologica* **3**, 15-27.
- 849 **Fox, J., Weisberg, S. and Fox, J.** (2011). An R companion to applied regression. Los  
850 Angeles: SAGE.
- 851 **Galoch, Z. and Bischof, H.-J.** (2007). Behavioural responses to video playbacks by  
852 zebra finch males. *Behavioural Processes* **74**, 21-26.
- 853 **Gil, D. and Gahr, M.** (2002). The honesty of bird song: Multiple constraints for multiple  
854 traits.
- 855 **Glaze, C. M. and Troyer, T. W.** (2006). Temporal Structure in Zebra Finch Song:  
856 Implications for Motor Coding. *The Journal of Neuroscience* **26**, 991-1005.
- 857 **Gonçalves, D. M., Oliveira, R. F., Korner, K., Poschadel, J. R. and Schlupp, I. I.**  
858 (2000). Using video playbacks to study visual communication in a marine fish, *Salarias pavo*.  
859 *Animal Behaviour* **60**, 351-357.
- 860 **Goodson, J. L., Kabelik, D., Kelly, A. M., Rinaldi, J. and Klatt, J. D.** (2009). Midbrain  
861 dopamine neurons reflect affiliation phenotypes in finches and are tightly coupled to courtship.  
862 *Proceedings of the National Academy of Sciences of the United States of America* **106**, 8737-  
863 8742.
- 864 **Guillette, L. M. and Healy, S. D.** (2017). The roles of vocal and visual interactions in  
865 social learning zebra finches: A video playback experiment. *Behavioural Processes* **139**, 43-49.
- 866 **Hamaguchi, K. and Mooney, R.** (2012). Recurrent interactions between the input and  
867 output of a songbird cortico-basal ganglia pathway are implicated in vocal sequence variability.  
868 *The Journal of neuroscience : the official journal of the Society for Neuroscience* **32**, 11671-  
869 11687.
- 870 **Hampton, C. M., Sakata, J. T. and Brainard, M. S.** (2009). An avian basal ganglia-  
871 forebrain circuit contributes differentially to syllable versus sequence variability of adult  
872 Bengalese finch song. *J Neurophysiol* **101**, 3235-45.
- 873 **Heinig, A., Pant, S., Dunning, J., Bass, A., Coburn, Z. and Prather, J. F.** (2014). Male  
874 mate preferences in mutual mate choice: finches modulate their songs across and within male-  
875 female interactions. *Animal Behaviour* **97**, 1-12.
- 876 **Hothorn, T., Bretz, F. and Westfall, P.** (2008). Simultaneous Inference in General  
877 Parametric Models. *Biometrical Journal* **50**, 346-363.
- 878 **Ihle, E. C., van der Hart, M., Jongsma, M., Tecott, L. H. and Doupe, A. J.** (2015).  
879 Dopamine physiology in the basal ganglia of male zebra finches during social stimulation. *The*  
880 *European journal of neuroscience* **41**, 1506-1514.
- 881 **Ikebuchi, M. and Okanoya, K.** (1999). Male Zebra Finches and Bengalese Finches  
882 Emit Directed Songs to the Video Images of Conspecific Females Projected onto a TFT  
883 Display: SPIE.
- 884 **James, L. S. and Sakata, J. T.** (2014). Vocal motor changes beyond the sensitive  
885 period for song plasticity. *Journal of Neurophysiology* **112**, 2040-2052.

- 886 **James, L. S. and Sakata, J. T.** (2015). Predicting plasticity: acute context-dependent  
887 changes to vocal performance predict long-term age-dependent changes. *J Neurophysiol* **114**,  
888 2328-39.
- 889 **Johnson, F., Soderstrom, K. and Whitney, O.** (2002). Quantifying song bout  
890 production during zebra finch sensory-motor learning suggests a sensitive period for vocal  
891 practice. *Behavioural Brain Research* **131**, 57-65.
- 892 **Kao, M. H. and Brainard, M. S.** (2006). Lesions of an avian basal ganglia circuit  
893 prevent context-dependent changes to song variability. *J Neurophysiol* **96**, 1441-55.
- 894 **Leblois, A. and Perkel, D. J.** (2012). Striatal dopamine modulates song spectral but not  
895 temporal features through D1 receptors. *The European journal of neuroscience* **35**, 1771-1781.
- 896 **Leblois, A., Wendel, B. J. and Perkel, D. J.** (2010). Striatal dopamine modulates basal  
897 ganglia output and regulates social context-dependent behavioral variability through D1  
898 receptors. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **30**,  
899 5730-5743.
- 900 **Long, M. A. and Fee, M. S.** (2008). Using temperature to analyse temporal dynamics in  
901 the songbird motor pathway. *Nature* **456**, 189.
- 902 **Maney, D. L.** (2013). The incentive salience of courtship vocalizations: Hormone-  
903 mediated 'wanting' in the auditory system. *Hearing Research* **305**, 19-30.
- 904 **Matheson, A. M. M. and Sakata, J. T.** (2015). Relationship between the Sequencing  
905 and Timing of Vocal Motor Elements in Birdsong. *PLOS ONE* **10**, e0143203.
- 906 **Matheson, L. E., Sun, H. and Sakata, J. T.** (2016). Forebrain circuits underlying the  
907 social modulation of vocal communication signals. *Dev Neurobiol* **76**, 47-63.
- 908 **Meaney, M. J.** (2001). Maternal Care, Gene Expression, and the Transmission of  
909 Individual Differences in Stress Reactivity Across Generations. *Annual Review of Neuroscience*  
910 **24**, 1161-1192.
- 911 **Morris, D.** (1954). The Reproductive Behaviour of the Zebra Finch (*Poephila guttata*),  
912 with Special Reference to Pseudofemale Behaviour and Displacement Activities. *Behaviour* **6**,  
913 271-322.
- 914 **Moser-Purdy, C. and Mennill, D. J.** (2016). Large vocal repertoires do not constrain  
915 the dear enemy effect: a playback experiment and comparative study of songbirds. *Animal*  
916 *Behaviour* **118**, 55-64.
- 917 **Moses, J., Loucks, J. A., Watson, H. L., Matuszewich, L. and Hull, E. M.** (1995).  
918 Dopaminergic drugs in the medial preoptic area and nucleus accumbens: Effects on motor  
919 activity, sexual motivation, and sexual performance. *Pharmacology Biochemistry and Behavior*  
920 **51**, 681-686.
- 921 **Murphy, K., James, L. S., Sakata, J. T. and Prather, J. F.** (2017). Advantages of  
922 comparative studies in songbirds to understand the neural basis of sensorimotor integration. *J*  
923 *Neurophysiol* **118**, 800-816.
- 924 **Oliveira, R. F., Rosenthal, G. G., Schlupp, I. I., McGregor, P. K., Cuthill, I. C., Endler,**  
925 **J. A., Fleishman, L. J., Zeil, J., Barata, E., Burford, F. et al.** (1999). Considerations on the use  
926 of video playbacks as visual stimuli: the Lisbon workshop consensus. *Acta ethologica* **3**, 61-65.
- 927 **Ophir, A. G., Persaud, K. N. and Galef, B. G., Jr.** (2005). Avoidance of relatively  
928 aggressive male Japanese Quail (*Coturnix japonica*) by sexually experienced conspecific  
929 females. *J Comp Psychol* **119**, 3-7.
- 930 **Ord, T. J., Peters, R. A., Evans, C. S. and Taylor, A. J.** (2002). Digital video playback  
931 and visual communication in lizards. *Animal Behaviour* **63**, 879-890.



- 932 **Pfaus, J. G., Damsma, G., Nomikos, G. G., Wenkstern, D. G., Blaha, C. D., Phillips,**  
933 **A. G. and Fibiger, H. C.** (1990). Sexual behavior enhances central dopamine transmission in  
934 the male rat. *Brain Research* **530**, 345-348.
- 935 **Podos, J., Lahti, D. C. and Moseley, D. L.** (2009). Chapter 5 Vocal Performance and  
936 Sensorimotor Learning in Songbirds. **40**, 159-195.
- 937 **Poopatanapong, A., Teramitsu, I., Byun, J. S., Vician, L. J., Herschman, H. R. and**  
938 **White, S. A.** (2006). Singing, but not seizure, induces synaptotagmin IV in zebra finch song  
939 circuit nuclei. *Journal of Neurobiology* **66**, 1613-1629.
- 940 **Rajan, R. and Doupe, A. J.** (2013). Behavioral and neural signatures of readiness to  
941 initiate a learned motor sequence. *Current biology : CB* **23**, 87-93.
- 942 **Rilling, J. K. and Young, L. J.** (2014). The biology of mammalian parenting and its  
943 effect on offspring social development. *Science (New York, N.Y.)* **345**, 771-776.
- 944 **Riters, L. V.** (2012). The role of motivation and reward neural systems in vocal  
945 communication in songbirds. *Frontiers in neuroendocrinology* **33**, 194-209.
- 946 **Riters, L. V., Absil, P. and Balthazart, J.** (1998). Effects of brain testosterone implants  
947 on appetitive and consummatory components of male sexual behavior in Japanese quail. *Brain*  
948 *Research Bulletin* **47**, 69-79.
- 949 **Riters, L. V., Teague, D. P., Schroeder, M. B. and Cummings, S. E.** (2004). Vocal  
950 production in different social contexts relates to variation in immediate early gene  
951 immunoreactivity within and outside of the song control system. *Behavioural Brain Research*  
952 **155**, 307-318.
- 953 **Ritschard, M., Laucht, S., Dale, J. and Brumm, H.** (2011). Enhanced testosterone  
954 levels affect singing motivation but not song structure and amplitude in Bengalese finches.  
955 *Physiology & Behavior* **102**, 30-35.
- 956 **Rosenthal, G. G.** (1999). Using video playback to study sexual communication.  
957 *Environmental Biology of Fishes* **56**, 307-316.
- 958 **Sakata, J. T., Hampton, C. M. and Brainard, M. S.** (2008). Social modulation of  
959 sequence and syllable variability in adult birdsong. *J Neurophysiol* **99**, 1700-11.
- 960 **Sakata, J. T. and Vehrencamp, S. L.** (2012). Integrating perspectives on vocal  
961 performance and consistency. *J Exp Biol* **215**, 201-9.
- 962 **Sasaki, A., Sotnikova, T. D., Gainetdinov, R. R. and Jarvis, E. D.** (2006). Social  
963 context-dependent singing-regulated dopamine. *The Journal of neuroscience : the official*  
964 *journal of the Society for Neuroscience* **26**, 9010-9014.
- 965 **Seredynski, A. L., Balthazart, J., Christophe, V. J., Ball, G. F. and Cornil, C. A.**  
966 (2013). Neuroestrogens rapidly regulate sexual motivation but not performance. *The Journal of*  
967 *neuroscience : the official journal of the Society for Neuroscience* **33**, 164-174.
- 968 **Sizemore, M. and Perkel, D. J.** (2008). Noradrenergic and GABA B receptor activation  
969 differentially modulate inputs to the premotor nucleus RA in zebra finches. *Journal of*  
970 *Neurophysiology* **100**, 8-18.
- 971 **Smith, G. T., Brenowitz, E. A., Beecher, M. D. and Wingfield, J. C.** (1997). Seasonal  
972 Changes in Testosterone, Neural Attributes of Song Control Nuclei, and Song Structure in Wild  
973 Songbirds. *The Journal of Neuroscience* **17**, 6001-6010.
- 974 **Smith, T. G., Brenowitz, E. A., Wingfield, J. C. and Baptista, L. F.** (1995). Seasonal  
975 changes in song nuclei and song behavior in Gambel's white-crowned sparrows. *Journal of*  
976 *Neurobiology* **28**, 114-125.
- 977 **Solis, M. M. and Perkel, D. J.** (2006). Noradrenergic modulation of activity in a vocal  
978 control nucleus in vitro. *Journal of Neurophysiology* **95**, 2265-2276.

- 979 **Sossinka, R. and Böhner, J.** (1980). Song Types in the Zebra Finch (*Poephila guttata*  
980 *castanotis*). *Z. Tierpsychol* **53**, 123-132.
- 981 **Stepanek, L. and Doupe, A. J.** (2010). Activity in a cortical-basal ganglia circuit for  
982 song is required for social context-dependent vocal variability. *J Neurophysiol* **104**, 2474-86.
- 983 **Stolzenberg, D. S., Stevens, J. S. and Rissman, E. F.** (2012). Experience-facilitated  
984 improvements in pup retrieval; evidence for an epigenetic effect. *Hormones and Behavior* **62**,  
985 128-135.
- 986 **Swaddle, J. P., McBride, L. and Malhotra, S.** (2006). Female zebra finches prefer  
987 unfamiliar males but not when watching noninteractive video. *Animal Behaviour* **72**, 161-167.
- 988 **Takahasi, M., Ikebuchi, M. and Okanoya, K.** (2005). Spatiotemporal properties of  
989 visual stimuli for song induction in Bengalese finches. *NeuroReport* **16**, 1339-1343.
- 990 **Tanaka, M., Sun, F., Li, Y. and Mooney, R.** (2018). A mesocortical dopamine circuit  
991 enables the cultural transmission of vocal behaviour. *Nature* **563**, 117-120.
- 992 **Toccalino, D. C., Sun, H. and Sakata, J. T.** (2016). Social Memory Formation Rapidly  
993 and Differentially Affects the Motivation and Performance of Vocal Communication Signals in  
994 the Bengalese Finch (*Lonchura striata* var. *domestica*). *Front Behav Neurosci* **10**, 14.
- 995 **Uetz, G. W. and Roberts, A.** (2002). Multisensory Cues and Multimodal  
996 Communication in Spiders: Insights from Video/Audio Playback Studies. *Brain, Behavior and*  
997 *Evolution* **59**, 222-230.
- 998 **Vahaba, D. M., Lacey, W. H. and Tomaszycski, M. L.** (2013). DSP-4, a noradrenergic  
999 neurotoxin, produces sex-specific effects on pairing and courtship behavior in zebra finches.  
1000 *Behavioural Brain Research* **252**, 164-175.
- 1001 **Vignal, C., Mathevon, N. and Mottin, S.** (2004). Audience drives male songbird  
1002 response to partner's voice. *Nature* **430**, 448-51.
- 1003 **Ware, E. L. R., Saunders, D. R. and Troje, N. F.** (2016). Social interactivity in pigeon  
1004 courtship behavior. *Current Zoology* **63**, 85-95.
- 1005 **Woolley, S. C.** (2019). Dopaminergic regulation of vocal-motor plasticity and  
1006 performance. *Current Opinion in Neurobiology* **54**, 127-133.
- 1007 **Woolley, S. C. and Doupe, A. J.** (2008). Social Context-Induced Song Variation Affects  
1008 Female Behavior and Gene Expression. *PLOS Biology* **6**, e62.
- 1009 **Woolley, S. C. and Kao, M. H.** (2015). Variability in action: Contributions of a songbird  
1010 cortical-basal ganglia circuit to vocal motor learning and control. *Neuroscience* **296**, 39-47.
- 1011 **Woolley, S. C., Rajan, R., Joshua, M. and Doupe, A. J.** (2014). Emergence of context-  
1012 dependent variability across a basal ganglia network. *Neuron* **82**, 208-23.
- 1013 **Zann, R. A.** (1996). The zebra finch : a synthesis of field and laboratory studies. Oxford ;  
1014 New York: Oxford University Press.
- 1015 **Zhang, Y. S., Wittenbach, J. D., Jin, D. Z. and Kozhevnikov, A. A.** (2017).  
1016 Temperature Manipulation in Songbird Brain Implicates the Premotor Nucleus HVC in Birdsong  
1017 Syntax. *The Journal of Neuroscience* **37**, 2600-2611.
- 1018