

Subspecies discrimination based on song structure by Willow Flycatchers

Sean M. Mahoney,¹  Bret Pasch, and Tad C. Theimer

Department of Biological Sciences, Northern Arizona University, Flagstaff, Arizona 86011, USA

Received 21 September 2020; accepted 28 April 2021

ABSTRACT. Animals use acoustic signals to repel competitors and attract mates, and signal divergence among populations can promote reproductive isolation. *Empidonax* flycatchers are insectivorous songbirds distributed across North and Central America that are conservative in plumage, but often exhibit differences in songs both among and within species. Four subspecies of Willow Flycatchers (*Empidonax traillii*) have been recognized and previous analyses have revealed differences in song structure among a subset of these subspecies. Using reciprocal playback experiments in the field, we tested for subspecific song discrimination among these four putative subspecies of Willow Flycatchers. We found that three subspecies (*E. t. adastus*, *E. t. brewsteri*, and *E. t. traillii*) responded similarly to their own songs and those of each other, but all three subspecies responded significantly less aggressively to songs of the southwestern subspecies (*E. t. extimus*). In contrast, the southwestern subspecies (*E. t. extimus*) responded significantly more aggressively to its own song than to those of the other three subspecies. Our results indicate that behavioral responses reflect differences in song structure among subspecies; subspecies responded more strongly to songs of subspecies with similar structures, less strongly to songs most different in structure, and the subspecies with the most distinctive song (*E. t. extimus*) responded less to songs of the other three subspecies. If responses of males to songs reflect relative reproductive compatibility within and among subspecies, songs may contribute to reproductive isolation of the four subspecies of Willow Flycatchers.

RESUMEN. **Discriminación de subspecies de *Empidonax traillii* con base en la estructura del canto**

Los animales usan señales acústicas para repeler competidores y atraer parejas y la divergencia en la señal entre poblaciones puede promover aislamiento reproductivo. Atrapamoscas en el género *Empidonax* son aves insectívoras distribuidas a lo largo de Norte y Centro América que tienen plumaje conservado filogenéticamente, pero con frecuencia muestran diferencias en cantos dentro y entre especies. Se reconocen cuatro subspecies de *Empidonax traillii* y análisis previos han revelado diferencias en la estructura del canto entre un sub grupo de estas subspecies. Usando experimentos de playback recíprocos en el campo, evaluamos la discriminación subspecifica del canto entre estas cuatro subspecies putativas de *Empidonax traillii*. Encontramos que tres subspecies (*E. t. adastus*, *E. t. brewsteri*, y *E. t. traillii*) respondieron similarmente a su propio canto que al de otras subspecies, pero las tres subspecies respondieron significativamente menos agresivamente a cantos de la subespecie del suroeste (*E. t. extimus*). En contraste, la subespecie del suroeste (*E. t. extimus*) respondió significativamente más agresivamente a su propio canto que a los de las otras tres subspecies. Nuestros resultados indican que las respuestas comportamentales reflejan diferencias en la estructura del canto entre subspecies; subspecies responden más fuertemente a cantos de subspecies con estructura similar, menos fuertemente a cantos con mayores diferencias en la estructura y las subspecies con el canto más diferente (*E. t. extimus*) respondieron menos a los cantos de las otras tres subspecies. Si las respuestas de los machos reflejan la compatibilidad reproductiva relativa dentro y entre subspecies, los cantos pueden contribuir al aislamiento reproductivo de las cuatro subspecies de *Empidonax traillii*.

Key words: behavioral isolation, *Empidonax traillii*, playback experiment, reproductive isolation, subspecies

Signal divergence among populations of closely related species may result in different behavioral responses that contribute to reproductive isolation and subsequent speciation (Nowicki et al. 1998, Coyne and Orr 2004). Bird songs are complex signals

broadcast broadly and readily evaluated by potential mates and competitors, and thus have the potential to be behavioral barriers to gene flow (Collins 2004, Coyne and Orr 2004). Most birds produce structurally unique songs and typically respond strongest to songs of their own species, supporting the hypothesis that song contributes to behavioral isolation among species (Andersson 1994). Reproductive isolation and diversification through song differentiation are hypothesized

¹Corresponding author. Email: sm2275@nau.edu

[Correction added on June 7, 2021, after first Online publication: Figure 2 has been replaced with revised figure].

to have contributed to the species diversity exhibited by oscine songbirds (Nottebohm 1972, Fitzpatrick 1988, Edwards et al. 2005). Oscine birds learn their songs, and the greater phenotypic plasticity allowed by learning has been argued to lead to the evolution of reproductive barriers based on song more rapidly in this group through a process of cultural evolution (West-Eberhard 1983, Verzijden et al. 2012, Yeh and Servedio 2015). Recently, however, a test of this hypothesis revealed that several species of suboscines, where songs are innate, showed higher rates of divergence through song discrimination than their oscine counterparts (Freeman et al. 2017).

Tyrannid flycatchers in the genus *Empidonax* are suboscines where songs have been found to be an important reproductive isolating mechanism at the species level. Plumage of tyrannid flycatchers is evolutionarily conservative (Zink and Johnson 1984, Rheindt et al. 2008) and morphologically similar species have been differentiated by differences in song (Stein 1963, Johnson and Cicero 2002, Rheindt et al. 2008). For example, Alder (*Empidonax alnorum*) and Willow (*Empidonax traillii*) flycatchers were considered conspecifics based on plumage pattern, morphology, and ecology (Stein 1958, Johnson and Cicero 2002), but song analyses revealed structural differences and reciprocal playback experiments showed that birds recognized those song differences (Stein 1963).

A similar phenomenon may be occurring at the subspecific level for Willow Flycatchers. Four subspecies of Willow Flycatchers are currently recognized by the USFWS (1995), including *E. t. traillii* in the eastern United States, *E. t. brewsteri* in the northwestern United States, *E. t. adastus* in much of the interior western United States, and *E. t. extimus* in riparian areas in the southwestern United States. Original subspecies designations were based on qualitative plumage assessments (Phillips 1948, Aldrich 1951, Unitt 1987, Browning 1993) and subsequent quantitative studies of plumage of birds in the hand have revealed significant differences in mean plumage coloration values of the four subspecies, but with considerable overlap (Paxton et al. 2010). Although models including both plumage and morphological measurements were better able to correctly

classify subspecies (Paxton et al. 2011), examination of color differences of museum specimens using models that account for avian visual systems (i.e., tetrahedral colorspace, Stoddard and Prum 2008) revealed few consistent differences in plumage among subspecies (Mahoney et al. 2020).

In terms of song, comparison of two western subspecies revealed that the main advertisement songs of *E. t. adastus* and *E. t. extimus* differed, with the songs of *E. t. extimus* having longer notes and phrases at overall lower frequencies (Sedgwick 2001). Comparison of the songs of all four subspecies revealed that the songs of *E. t. extimus* were lower in frequency with fewer terminal frequency modulations than songs of the other three subspecies (Mahoney et al. 2020). Hierarchical clustering identified two song groups, one containing *E. t. extimus* songs and the other containing songs of the other three subspecies (Mahoney et al. 2020). These differences in the songs of the four subspecies mirrored earlier studies of variation in mtDNA that revealed less divergence among the three northern subspecies, but significant differences between those subspecies and the southwestern subspecies (Paxton 2000, Paxton et al. 2008). The possibility that differences in the songs of these four subspecies are recognized by the birds, potentially acting as an isolating mechanism contributing to genetic differentiation, has not been tested.

Identifying the mechanisms contributing to population divergence requires an understanding of the degree of character divergence and the strength of behavioral responses to the diverged traits (Seddon and Tobias 2010, Hudson and Price 2014). Reciprocal playback experiments with territorial males serve as useful tests of the strength of behavioral responses to vocalizations that have diverged in structure (e.g., Catchpole 1978, Prescott 1987, Lipshutz et al. 2017). Our objective was to quantify aggressive responses during reciprocal playback experiments to determine if differences in the songs of the four subspecies of Willow Flycatchers potentially contribute to behavioral isolation. We predicted that birds would respond more aggressively to the songs of their own subspecies and those of subspecies with songs that were similar in structure, and less aggressively to songs of subspecies that differed in structure from their own songs.

METHODS

We searched areas on state and federal public lands for singing flycatchers in the ranges of *E. t. adastus* and *E. t. brewsteri* during June and July 2018, and in the ranges of *E. t. adastus*, *E. t. extimus*, and *E. t. traillii* during June and July 2019 (Fig. 1, Table S1). Subspecies ranges were those delineated in Paxton (2000, Fig. 1) based on Unitt (1987) and Browning (1993). We determined approximate territory boundaries for each singing bird by identifying at least three perches that the individual used at least three times and then placed a speaker (Bose Soundlink II) 1.5-2-m high in a tree or shrub approximately equidistant from the three perches. Territories of focal individuals were separated by a minimum of 150 m.

At each site, we randomly presented five treatments to 2-8 focal individuals: (1) *E. t. adastus* song, (2) *E. t. brewsteri* song, (3) *E. t. extimus* song, (4) *E. t. traillii* song, and (5) a white-noise control. Thus, each focal individual received stimuli from its own subspecies, all other subspecies, and a control treatment. To minimize the effects of pseudoreplication (Kroodsma 1989, McGregor et al. 1992), we randomly selected stimulus files for each trial from a library of 60 songs (one song from 60 different individuals) (Table S2, 15 each of *E. t. adastus*, *E. t. brewsteri*, *E. t. extimus*, and *E. t. traillii*). We used white noise (i.e., random noise with equal energy at all frequencies) as our control because it covers the frequency bandwidth of Willow Flycatcher song (~ 1–7 kHz, Mahoney et al. 2020) and experiments have shown that birds respond

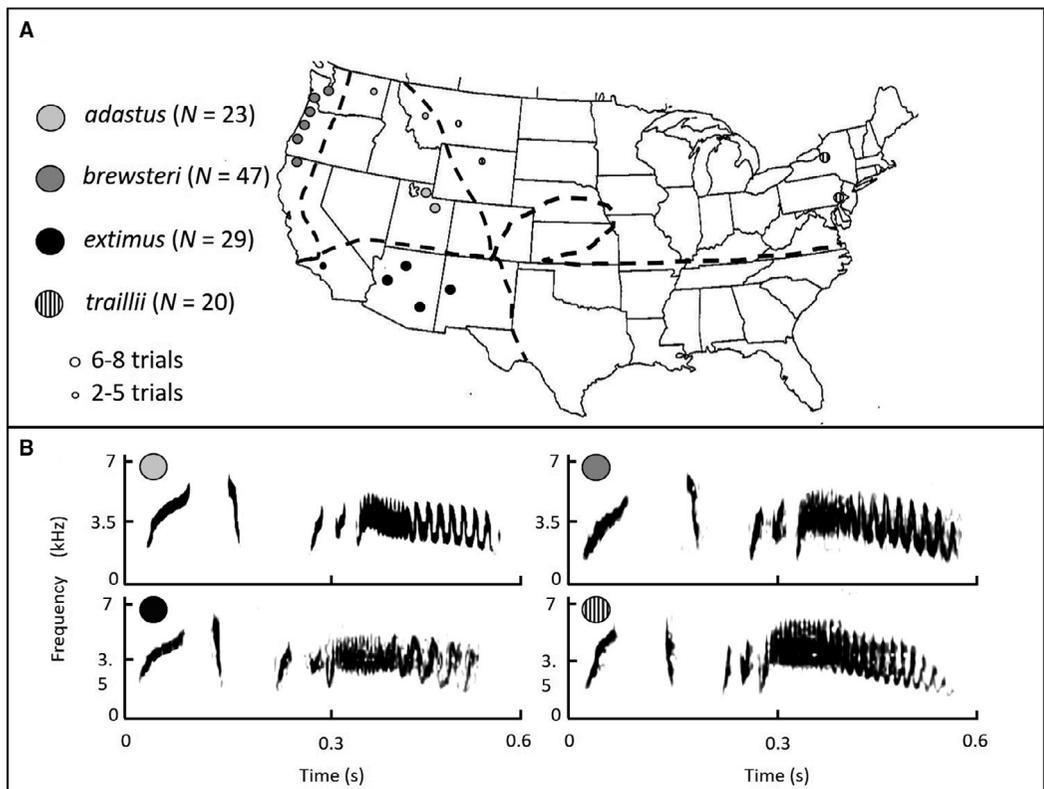


Fig. 1. (A) Locations and number of trials of playback experiments to test subspecific song discrimination by Willow Flycatchers. Dot color refers to putative subspecies and subspecies ranges are indicated by dashed lines. Dots are scaled to trial sample sizes at a particular location. (B) Representative spectrograms of song stimuli of putative subspecies of Willow Flycatchers used in playback experiments. Dot color refers to subspecies from panel A.

similarly to white noise and heterospecific songs (Blumstein et al. 2017) so it is a robust stimulus to use as a baseline for responses to playback. Most focal birds received all treatments on the same day ($N = 116$ of 119), with at least 30 min between successive trials (Prescott 1987). Most trials ($N = 550$ of 564) were conducted between 05:00 and 12:30 (Demko et al. 2019). However, we included time as a fixed effect and playback treatment order as a random effect in our models because birds may respond differently to territorial intrusions throughout the day (see below).

Our trials consisted of a 2.5-min pre-playback period (pre-stimulus behavior) and a 2.5-min playback period (treatment behavior). Trials began when a focal bird was observed singing ~ 20 m away. One observer sat 8–10 m from the speaker to record bird behavior. During trials, the observer recorded behavior documented in other playback studies of Willow Flycatchers (Prescott 1987, Lovell and Lein 2005), including latency to approach to 8 m from the speaker, time spent within 3 m of the speaker, closest approach distance to the speaker, number of fitz-bew songs, and number of wheeo, churr, breet, breet-bew, whitt, and pip vocalizations. Although the function of breets, breet-bews, and pips is unknown, whitts, wheeos, and churrs are thought to indicate increasing agitation (Stein 1958, 1963, Prescott 1987). We only recorded behaviors from the putative male of any pair, as identified by territorial behavior (mainly singing and chasing conspecifics). We defined more aggressive responses as shorter latency to 8 m from the speaker, more time spent within 3 m of the speaker, and closer approach distances to the speaker. Observers ($N = 1$ in 2018 and $N = 2$ in 2019) were not blind to the stimulus type, and trials were videotaped when possible to provide a permanent record.

Playback stimuli. Playback stimuli consisted of a Willow Flycatcher song repeated at the natural rate of 10 songs/min (Prescott 1987). Stimuli consisted of songs we either recorded in the field or obtained from publicly available sources (Table S2). To ensure that song files had high signal-to-noise ratios, we used AviSoft SASLab Pro (v. 5.2.12) to remove all background noise. The white-noise control was generated using Audacity

(v. 2.1.3) and was 0.5 s in duration (the approximate mean duration of Willow Flycatcher songs across all subspecies; Mahoney et al. 2020) and played at a rate of 10 per minute. All stimuli were normalized to a peak amplitude value of 1 in MATLAB (version 2018a) prior to use in playback experiments.

Statistical analyses. To assess the behavioral responses of Willow Flycatchers to subspecific songs, we used principal components analysis (PCA) to reduce the behavioral variables to an uncorrelated behavioral response variable (using R; R Core Development Team 2019) for all subspecies in one model for the pre-stimulus and another for the treatment period (pre-stimulus behavior PCA and treatment behavior PCA). To meet assumptions of linearity, we used logarithmic transformations via $\log(x + 1)$ of all data prior to the PCA. To assist in interpretation of the PCA results, we multiplied PCA scores by -1 (Vehrencamp et al. 2003). We then tested for differences in the pre-stimulus PC1 and treatment behavior PC1, which explains the most variation in the data, using separate linear mixed effects models for each subspecies. We used PC1 as the response variable, playback treatment (*E. t. adastus*, *E. t. brewsteri*, *E. t. extimus*, *E. t. traillii*, and control), site, ordinal date (day of year), time of day, observer, and the interaction between treatment and site as fixed effects, and bird identity and playback treatment order as random effects. We assessed assumptions of linearity and homoscedasticity by visually assessing model QQ and residual plots. We then assessed post hoc pairwise comparisons using the lsmeans package in R (Lenth and Love 2018). We treated playback experiments in each subspecies range as separate hypotheses. To minimize type I error risk, we evaluated our hypotheses using a Bonferroni-corrected alpha-level for within-species post hoc comparisons ($N = 10$ post hoc tests for each subspecies). Therefore, our corrected alpha was 0.005.

RESULTS

From our pre-stimulus behavior PCA, PC1 (Table 1, 56% of variation, eigenvalue = 2.69) was associated with more fitz-bew, breet-bew, and breet vocalizations. Pre-stimulus PC2 (Table 1, 15%, eigenvalue =

Table 1. Factor loadings for the first two principal components derived from behavioral responses of Willow Flycatchers during playback experiments. Loadings > 0.25 are shown. Vocalizations are indicated in italics, and movements in ordinary text.

Trial period	Behavior	PC1	PC2
Pre-stimulus	<i>Fitz-bew</i>	0.69	–
	<i>breet-bew</i>	0.51	–
	<i>breet</i>	0.46	–
	<i>whitt</i>	–	–0.96
	Variation	56%	15%
	Eigenvalue	2.69	0.71
Treatment	Time within 3 m	0.84	–
	Latency to 8 m	–0.28	–
	Closest distance	–0.39	–
	<i>Fitz-bew</i>	–	–0.62
	<i>breet</i>	–	–0.49
	<i>breet-bew</i>	–	–0.46
	Variation	43%	21%
	Eigenvalue	3.96	1.96

0.71) was associated with fewer whitt vocalizations. Pre-stimulus behavior did not vary with playback treatment type (Fig. S1, $F_{5,373.55} = 0.4$, $P = 0.85$), indicating that focal birds returned to pre-stimulus behavior between trials, regardless of the stimulus previously received. Pre-stimulus behavior varied by site (Fig. S2, $F_{19, 101.03} = 3.6$, $P < 0.0001$) and time of day ($F_{1, 444.2} = 133.2$, $P < 0.0001$). We found no observer effect ($F_{1, 91.87} = 0.1$, $P = 0.82$), no treatment x site interaction ($F_{76, 365.91} = 0.8$, $P = 0.93$), and no date effect ($F_{1, 227.78} = 2.0$, $P = 0.16$). Playback treatment order and individual explained little variation in pre-stimulus behavior (treatment order $\sigma^2 = 0.0$, individual $\sigma^2 = 0.83$).

Based on our treatment behavior PCA, PC1 (Table 1, 43% of variation, eigenvalue = 3.96) was associated with more time spent within 3 m of the speaker, closer approach distances to the speaker, and shorter latency to approach within 8 m of the speaker. Treatment behavior PC2 (Table 1, 21% of variation, eigenvalue = 1.96) was associated with fewer fitz-bew, breet, and breet-bew vocalizations. We found no effect of site ($F_{19, 98.35} = 1.1$, $P = 0.41$), date ($F_{1, 119.57} = 0.1$, $P = 0.74$), time of day ($F_{1, 29.75} = 0.3$, $P = 0.59$), or observer ($F_{1, 90.87} = 0.4$, $P = 0.53$) on behavior PC1 (Table 2). Individual and treatment order explained little variation in treatment behavior (Table 2,

individual $\sigma^2 = 0.84$, treatment order $\sigma^2 = 0.03$). We found a significant effect of treatment ($F_{4, 369.66} = 24.7$, $P < 0.0001$; Table 2), suggesting that subspecies responded differently based on subspecific stimuli, and an interaction between site and treatment ($F_{76, 367.93} = 2.0$, $P < 0.0001$) (Table 2, Fig. S3). This interaction was driven primarily by the fact that sites were associated with different subspecies, and responses at sites in the range of *E. t. extimus* differed from those at sites in the ranges of the other three subspecies (Fig. S3).

Individuals generally responded more aggressively to the song of their own subspecies and to heterosubspecific songs that were similar in structure than to songs of subspecies with a different song structure (Fig. 2A-D). For three subspecies, *E. t. adastus*, *E. t. breusteri*, and *E. t. traillii*, aggressive responses to their own songs and those of the other two heterosubspecifics with songs similar in structure, as reflected in PC1 scores, did not differ (Fig. 2A-C, Table 2, all post hoc comparisons $P > 0.05$). However, all three of these subspecies responded significantly less aggressively to the song of the subspecies with a more different structure (*E. t. extimus*) (Fig. 2A-C, Table 2, all post hoc comparisons $P < 0.005$). The subspecies *E. t. extimus* responded significantly more aggressively to the song of their own subspecies than to the songs of the other three

Table 2. Results of the full mixed effects model of male Willow Flycatcher behavior PC1 in response to subspecific songs and white-noise playback. Post hoc comparisons between two treatment stimuli presented to focal individuals are shown below. Bold indicates within-subspecies Bonferroni-corrected significant differences between two trial stimuli (playback stimulus comparison, $\alpha = 0.005$).

Full model	<i>F</i>	df	<i>P</i> value	Individual σ^2	Playback order σ^2
Treatment	26.7	4, 369.66	< 0.0001	0.84	0.03
Site	1.1	19, 98.35	0.41		
Date	0.1	1, 119.57	0.74		
Time	0.3	1, 29.75	0.59		
Observer	0.4	1, 90.87	0.53		
Treatment \times site	2.0	76, 367.93	< 0.0001		

Focal subspecies	Playback stimulus comparison	Estimate	SE	df	<i>t</i>	<i>P</i> value
<i>E. t. adastus</i>	<i>adastus</i> – <i>brewsteri</i>	–0.30	0.47	67.61	–0.6	0.97
	<i>adastus</i> – control	–2.78	0.52	72.78	–5.3	< 0.0001
	<i>adastus</i> – <i>extimus</i>	–2.69	0.45	70.57	–5.9	< 0.0001
	<i>adastus</i> – <i>traillii</i>	–0.56	0.45	70.52	–1.2	0.73
	<i>brewsteri</i> – control	–2.48	0.54	72.48	–4.6	0.0002
	<i>brewsteri</i> – <i>extimus</i>	–2.38	0.45	70.92	–5.3	< 0.0001
	<i>brewsteri</i> – <i>traillii</i>	–0.26	0.46	70.74	–0.6	0.98
	<i>control</i> – <i>extimus</i>	0.10	0.53	73.55	0.2	0.99
	<i>control</i> – <i>traillii</i>	2.22	0.53	72.40	4.2	0.0006
	<i>extimus</i> – <i>traillii</i>	2.12	0.45	70.90	4.7	0.0001
<i>E. t. brewsteri</i>	<i>adastus</i> – <i>brewsteri</i>	0.69	0.33	139.25	2.1	0.23
	<i>adastus</i> – control	–2.06	0.43	154.64	–4.8	< 0.0001
	<i>adastus</i> – <i>extimus</i>	–1.24	0.34	141.51	–3.7	0.003
	<i>adastus</i> – <i>traillii</i>	0.26	0.33	140.75	0.8	0.94
	<i>brewsteri</i> – control	–2.75	0.43	154.47	–6.4	< 0.0001
	<i>brewsteri</i> – <i>extimus</i>	–1.93	0.34	139.88	–5.7	< 0.0001
	<i>brewsteri</i> – <i>traillii</i>	–0.43	0.33	141.81	–1.3	0.69
	<i>control</i> – <i>extimus</i>	0.82	0.44	153.86	1.9	0.33
	<i>control</i> – <i>traillii</i>	2.32	0.44	148.18	5.3	< 0.0001
	<i>extimus</i> – <i>traillii</i>	1.50	0.33	141.85	4.5	0.0001
<i>E. t. extimus</i>	<i>adastus</i> – <i>brewsteri</i>	–0.02	0.50	92.90	–0.04	1.00
	<i>adastus</i> – control	–0.92	0.54	96.10	–1.7	0.45
	<i>adastus</i> – <i>extimus</i>	2.34	0.51	93.90	4.6	0.0001
	<i>adastus</i> – <i>traillii</i>	–0.26	0.50	91.75	–0.5	0.98
	<i>brewsteri</i> – control	–0.90	0.55	96.80	–1.6	0.49
	<i>brewsteri</i> – <i>extimus</i>	2.36	0.52	94.20	4.6	0.0002
	<i>brewsteri</i> – <i>traillii</i>	–0.24	0.51	93.70	–0.5	0.99
	<i>control</i> – <i>extimus</i>	3.25	0.54	95.40	6.1	< 0.0001
	<i>control</i> – <i>traillii</i>	0.65	0.54	95.40	1.2	0.74
	<i>extimus</i> – <i>traillii</i>	–2.60	0.51	94.0	–5.1	< 0.0001
<i>E. t. traillii</i>	<i>adastus</i> – <i>brewsteri</i>	–0.13	0.57	56.86	–0.2	1.00
	<i>adastus</i> – control	–1.77	0.61	56.74	–2.9	0.04
	<i>adastus</i> – <i>extimus</i>	–1.53	0.57	56.39	–2.7	0.07
	<i>adastus</i> – <i>traillii</i>	0.85	0.57	56.57	1.5	0.58
	<i>brewsteri</i> – control	–1.64	0.62	57.66	–2.7	0.07
	<i>brewsteri</i> – <i>extimus</i>	–1.41	0.54	55.02	–2.6	0.09
	<i>brewsteri</i> – <i>traillii</i>	0.97	0.53	53.91	1.9	0.36
	<i>control</i> – <i>extimus</i>	0.24	0.60	56.39	0.4	0.99
	<i>control</i> – <i>traillii</i>	2.62	0.62	57.60	4.2	0.0008
	<i>extimus</i> – <i>traillii</i>	2.38	0.55	55.17	4.4	0.0005

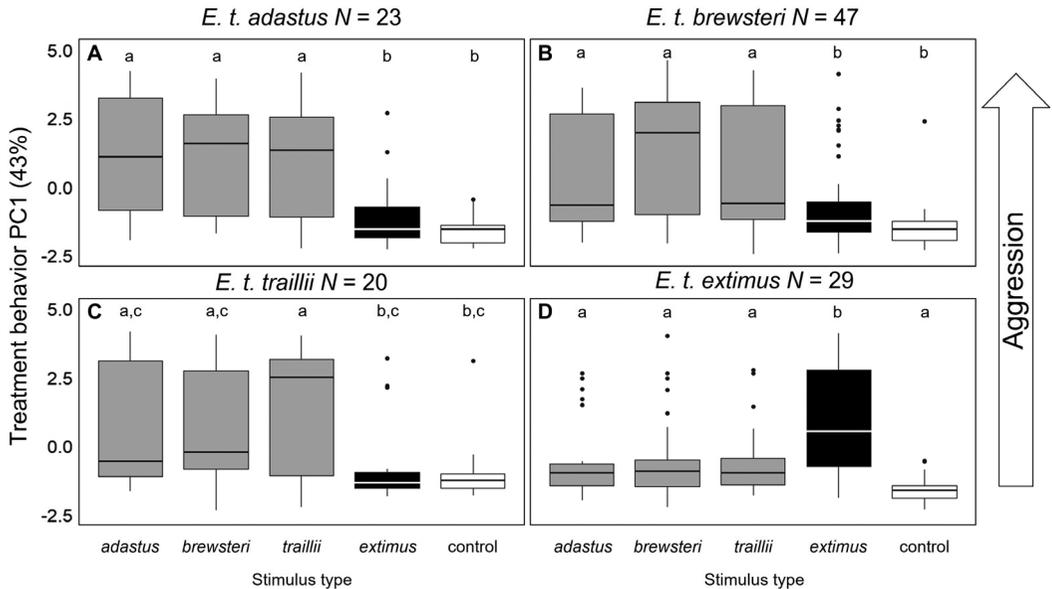


Fig. 2. Boxplots of behavioral response (PC1 [43% of variation], y -axis) to subspecific song stimuli and white-noise control (x -axis) during playback experiment trials with subspecies of Willow Flycatchers (*Empidonax traillii*), including (A) *E. t. adustus*, (B) *E. t. brewsteri*, (C) *E. t. traillii*, and (D) *E. t. extimus*. Positive PC1 scores represent more aggressive responses to playback as indicated by the arrow. Boxes are colored based on song structure groups (song group 1 [black] and song group 2 [gray]) from Mahoney et al. (2020) and lower-case letters represent Bonferroni-corrected differences ($P < 0.005$). Subspecies responded more aggressively to songs that were structurally similar to their own subspecies song ($F_{4, 369.66} = 24.7$, $P < 0.0001$), and the interaction between treatment type and study site was significant ($F_{7, 367.93} = 2.0$, $P < 0.0001$). Boxes show 1st and 3rd quartiles, horizontal line represents median, values within 1.5 times the interquartile range are represented by whiskers and values outside this range are represented by dots.

subspecies (Fig. 2D, Table 2, all post hoc comparisons $P < 0.0001$).

DISCUSSION

Our results are consistent with the hypothesis that subspecific differences in songs (Sedgwick 2001, Mahoney et al. 2020) may act as a behavioral isolating mechanism between *E. t. extimus* and the other three subspecies. Comparison of songs across all four subspecies revealed that the song of *E. t. extimus* was most different from those of the other three subspecies. The songs of *E. t. extimus* had fewer frequency modulations in the terminal portion of the song and were generally lower in frequency; songs of the other subspecies were higher in frequency with more terminal frequency modulations (Mahoney et al. 2020). The behavioral responses we

documented mirrored that pattern; the subspecies with the most distinct song (*E. t. extimus*) responded more strongly to its own song than to those of the other three subspecies, and the three subspecies with similar songs responded to each other's songs more strongly than to the song of *E. t. extimus*.

Willow Flycatchers cannot be sexed in the field based on plumage and we assumed that the birds we observed were males, but females also sing (Seutin 1987, Yard and Brown 2003) and respond aggressively to song (Sogge et al. 2007). If the responses we recorded were predominately by males, the potential for song to act as a behavioral reproductive isolating mechanism must be interpreted with caution because it is ultimately the effect on female mate choice that determines gene flow. Females of several species respond differentially to subtle differences

in song when males do not (e.g., Searcy and Brenowitz 1988, Seddon and Tobias 2010, Danner et al. 2011). As a result, reproductive barriers based on song may still exist for those subspecies where we detected no difference (e.g., between *E. t. brewsteri* and *E. t. adastus*) if responses we recorded were primarily by males, but females in those areas discriminate among potential mates based on subtler differences in song. Further studies of behavioral responses to songs of different subspecies in populations where the sex of banded birds is known would clarify this issue.

Another factor that could shape responses to subspecific songs by Willow Flycatchers is competition in wintering areas. In this species, both sexes maintain long-term, mutually exclusive winter territories and use song to defend those territories (Koronkiewicz et al. 2006). Playback experiments using songs recorded in the breeding range of *E. t. extimus* elicited equally aggressive responses by both males and females wintering in Costa Rica (Sogge et al. 2007). Using plumage, morphological measurements, and genetics to identify putative subspecies, Paxton et al. (2011) found that the four subspecies occupied distinct, but overlapping, wintering areas. The eastern subspecies *E. t. traillii* wintered generally south of the three western subspecies in southern Central America and northern South America, whereas *E. t. extimus* wintered in Central America (especially Costa Rica) and *E. t. adastus* and *E. t. brewsteri* wintered in areas from Mexico to Panama. Based on those distributions, the potential for subspecies to compete in wintering areas would appear to be greater for the three western subspecies and, given the relative rarity of *E. t. extimus*, competition would be more likely between *E. t. brewsteri* and *E. t. adastus* (M. Whitfield, pers. comm.). If so, the greater behavioral response by the two western subspecies to each other's songs in breeding areas in our study could be due in part to competition in wintering areas.

One caveat of our study was that we did not intentionally test different responses to song among subspecies in areas of sympatry. Responses to playback of heterospecific songs may vary based on a receiver's prior experience with con- and heterospecifics (Catchpole 1978, Richards 1979), and birds in sympatry could exhibit a heightened response to both

con- and heterospecifics due to resource competition (Orians and Willson 1964). Diminished responses to songs of heterospecifics in allopatry have been demonstrated in many systems, for example, buntings (*Passerina*, Emlen et al. 1975), reed warblers (*Acrocephalus*, Catchpole and Leisler 1986) African tinkerbirds (*Pogoniulus*, Kirschel et al. 2009), Rufous-capped Warblers (*Basileuterus rufifrons*, Demko et al. 2019) and closely related Willow and Alder flycatchers also responded to the songs of heterospecifics in sympatry (Prescott 1987), but not in allopatry (Stein 1963). The two western-most *E. t. traillii* sites and the *E. t. adastus* site in Montana in our study, although within the range of their respective subspecies as delineated by Paxton (2000), are areas that have not been extensively studied and subspecies boundaries in this area may be imprecise. Although responses of birds at the *E. t. adastus* site in Montana were consistent with those of birds at other sites in the range of *E. t. adastus*, the pattern of response to heterosubspecific songs differed between the two western-most *E. t. traillii* sites. At the Wyoming site, birds responded most strongly to *E. t. traillii* song and secondarily to *E. t. adastus* song whereas, at the *E. t. traillii* site in Montana, birds unexpectedly responded more strongly to the song of *E. t. brewsteri* than to any other stimulus. An added complexity in areas of potential sympatry is the potential for song structure to change due to hybridization at subspecies boundaries. Songs of flycatchers are innate rather than learned (Kroodsma 1984), and Sedgwick (2001) found evidence of hybrid songs at sites along the boundary between the ranges of *E. t. extimus* and *E. t. adastus*. More extensive sampling along boundaries of the ranges of the different subspecies could clarify both the potential for changes in behavioral response in areas of sympatry and the potential for changes in song structure as the result of hybridization.

Finally, our results are consistent with the need to continue to recognize the southwestern subspecies under the Endangered Species Act (USFWS 1995, Zink 2015, Theimer et al. 2016). Our results demonstrate that differences in songs of *E. t. extimus* and the other three subspecies (Sedgwick 2001, Mahoney et al. 2020) are recognized by territorial males in breeding areas and, when

played back, elicit different behavioral responses. These results are important given the current and predicted impacts of climate change on the southwestern United States and the implications that will have for Willow Flycatchers in those habitats (Ruegg et al. 2018).

ACKNOWLEDGMENTS

This work was supported by the T & E Inc. Conservation grant (SMM), the NAU John Prather Conservation award (SMM), and the NAU Landscape Conservation Initiative Fellowship (SMM). Field assistance from S. I. Gonzalez was invaluable. Tom and Allison Mickel (Eugene, OR) provided crucial field site consultation for playback studies. M. J. Whitfield, M. A. McLeod, A. Pellegrini, G. C. Cummins, M. J. Johnson, and D. Moore provided critical field assistance. We thank C. E. Aslan and S. M. Shuster for reviewing early drafts of this manuscript. Reviews by one anonymous reviewer, M. J. Whitfield, E. Paxton, and G. Ritchison substantively improved this manuscript. Research was conducted in compliance with NAU IACUC and the USFWS. Video files, song stimuli used in playback experiments, and raw data are available in the Dryad Repository.

CONFLICT OF INTEREST

The authors declare no conflict of interest. Research was conducted in compliance with NAU IACUC and USFWS. Video files, song stimuli used in playback experiments, and raw data are available on the Dryad Repository.

AUTHOR CONTRIBUTIONS

SMM and TCT conceived of project. SMM conducted field work, analyzed data, and wrote initial manuscript draft. All authors designed project and wrote revised manuscript.

LITERATURE CITED

- ALDRICH, J. W. 1951. A review of the races of the Traill's Flycatcher. *Wilson Bulletin* 63: 192–197.
- ANDERSSON, M. B. 1994. Sexual selection. Princeton University Press, Princeton, NJ.
- BLUMSTEIN, D. T., J. WHITAKER, J. KENNEN, AND G. A. BRYANT. 2017. Do birds differentiate between white noise and deterministic chaos? *Ethology* 123: 966–973.
- BROWNING, M. R. 1993. Comments on the taxonomy of *Empidonax traillii* (Willow Flycatcher). *Western Birds* 24: 241–257.
- CATCHPOLE, C. K. 1978. Interspecific territorialism and competition in *Acrocephalus* warblers as revealed by playback experiments in areas of sympatry and allopatry. *Animal Behaviour* 26: 1072–1080.
- , and B. LEISLER. 1986. Interspecific territorialism in Reed Warblers: a local effect revealed by playback experiments. *Animal Behaviour* 34: 299–300.
- COLLINS, S. 2004. Vocal fighting and flirting: the functions of birdsong. In: *Nature's music: the science of birdsong* (P. Marler and H. Slabbekoorn, eds.), pp. 39–79. Elsevier Academic Press, San Diego, CA.
- COYNE, J. A., AND H. A. ORR. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.
- DANNER, J. E., R. M. DANNER, F. BONIER, P. R. MARTIN, T. W. SMALL, AND I. T. MOORE. 2011. Female, but not male, tropical sparrows respond more strongly to the local song dialect: implications for population divergence. *American Naturalist* 178: 53–63.
- DEMKO, A. D., J. R. SOSA-LÓPEZ, AND D. J. MENNILL. 2019. Subspecies discrimination on the basis of acoustic signals: a playback experiment in a Neotropical songbird. *Animal Behaviour* 157: 77–85.
- EDWARDS, S. V., S. B. KINGAN, J. D. CALKINS, C. N. BALAKRISHNAN, W. B. JENNINGS, W. J. SWANSON, AND M. D. SORENSON. 2005. Speciation in birds: genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences USA* 102S: 6550–6557.
- EMLEN, S. T., J. D. RISING, AND W. L. THOMPSON. 1975. A behavioral and morphological study of sympatry in the Indigo and Lazuli buntings of the Great Plains. *Wilson Bulletin* 87: 145–177.
- FITZPATRICK, J. W. 1988. Why so many passerine birds? A response to Raikow. *Systematic Zoology* 37: 71.
- FREEMAN, B. G., G. A. MONTGOMERY, AND D. SCHLUTER. 2017. Evolution and plasticity: divergence of song discrimination is faster in birds with innate song than in song learners in Neotropical passerine birds. *Evolution* 71: 2230–2242.
- HUDSON, E. J., AND T. D. PRICE. 2014. Pervasive reinforcement and the role of sexual selection in biological speciation. *Journal of Heredity* 105: 821–833.
- JOHNSON, N. K., AND C. CICERO. 2002. The role of ecologic diversification in sibling speciation of *Empidonax* flycatchers (Tyrannidae): multigene evidence from mtDNA. *Molecular Ecology* 11: 2065–2081.
- KIRSCHER, A. N., D. T. BLUMSTEIN, AND T. B. SMITH. 2009. Character displacement of song and morphology in African tinkerbirds. *Proceedings of the National Academy of Sciences USA* 106: 8256–8261.
- KORONKIEWICZ, T. J., M. K. SOGGE, C. VAN RIPER III, AND E. H. PAXTON. 2006. Territoriality, site fidelity, and survivorship of Willow Flycatchers wintering in Costa Rica. *Condor* 108: 558–570.

- KROODSMA, D. E. 1984. Songs of the Alder Flycatcher (*Empidonax alnorum*) and Willow Flycatcher (*Empidonax traillii*) are innate. *Auk* 101: 13–24.
- . 1989. Suggested experimental designs for song playbacks. *Animal Behaviour* 37: 600–609.
- LENTH, R., AND J. LOVE. 2018. Package 'lsmeans'. *American Statistician* 34: 216–221.
- LIPSHUTZ, S. E., I. A. OVERCAST, M. J. HICKERSON, R. T. BRUMFIELD, AND E. P. DERRYBERRY. 2017. Behavioural response to song and genetic divergence in two subspecies of White-crowned Sparrows (*Zonotrichia leucophrys*). *Molecular Ecology* 26: 3011–3027.
- LOVELL, S. F., AND M. R. LEIN. 2005. Individual recognition of neighbors by song in a subsocial bird, the Alder Flycatcher (*Empidonax alnorum*). *Behavioral Ecology and Sociobiology* 57: 623–630.
- MAHONEY, S. M., M. W. REUDINK, B. PASCH, AND T. C. THEIMER. 2020. Song but not plumage varies geographically among Willow Flycatcher (*Empidonax traillii*) subspecies. *Journal of Avian Biology* 51: e02621.
- MCGREGOR, P. K., T. DABELSTEEN, M. SHEPHERD, AND S. B. PEDERSEN. 1992. The signal value of matched singing in Great Tits: evidence from interactive playback experiments. *Animal Behaviour* 43: 987–998.
- NOTTEBOHM, F. 1972. The origins of vocal learning. *American Naturalist* 106: 116–140.
- NOWICKI, S., S. PETERS, AND J. PODOS. 1998. Song learning, early nutrition and sexual selection in songbirds. *American Zoologist* 38: 179–190.
- ORIANI, G. H., AND M. F. WILLSON. 1964. Interspecific territories of birds. *Ecology* 45: 736–745.
- PAXTON, E. H. 2000. Molecular genetic structuring and demographic history of the Willow Flycatcher (*Empidonax traillii*). M. S. thesis, Northern Arizona University, Flagstaff, AZ.
- , M. K. SOGGE, T. J. KORONKIEWICZ, M. A. MCLEOD, AND T. C. THEIMER. 2010. Geographic variation in the plumage coloration of Willow Flycatchers *Empidonax traillii*. *Journal of Avian Biology* 41: 128–138.
- , M. K. SOGGE, T. C. THEIMER, J. GIRARD, AND P. KEIM. 2008. Using molecular markers to resolve a subspecies boundary: the northern boundary of the southwestern Willow Flycatcher in the Four Corner states. U.S. Geological Survey Open File Report 2008-1117, Flagstaff, AZ.
- PAXTON, E. H., P. UNITT, M. K. SOGGE, M. J. WHITFIELD, AND P. KEIM. 2011. Winter distribution of Willow Flycatcher subspecies. *Condor* 113: 608–618.
- PHILLIPS, A. R. 1948. Variation in *Empidonax traillii*. *Auk* 65: 507–514.
- PRESCOTT, D. R. 1987. Territorial responses to song playback in allopatric and sympatric populations of Alder (*Empidonax alnorum*) and Willow (*E. traillii*) flycatchers. *Wilson Bulletin* 99: 611–619.
- R CORE DEVELOPMENT TEAM. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- RHEINDT, F. E., J. A. NORMAN, AND L. CHRISTIDIS. 2008. DNA evidence show vocalizations to be a better indicator of taxonomic limits than plumage patterns in *Zimmerius* tyrant-Flycatchers. *Molecular Phylogenetics and Evolution* 48: 150–156.
- RICHARDS, D. G. 1979. Recognition of neighbors by associative learning in Rufous-sided Towhees. *Auk* 96: 688–693.
- RUEGG, K., R. A. BAY, E. C. ANDERSON, J. F. SARACCO, R. J. HARRIGAN, M. WHITFIELD, E. H. PAXTON, AND T. B. SMITH. 2018. Ecological genomics predicts climate vulnerability in an endangered southwestern songbird. *Ecology Letters* 21: 1085–1096.
- SEARCY, W. A., AND E. A. BRENOWITZ. 1988. Sexual differences in species recognition of avian song. *Nature* 332: 152–154.
- SEDDON, N., AND J. A. TOBIAS. 2010. Character displacement from the receiver's perspective: species and mate recognition despite convergent signals in subsocial birds. *Proceedings of the Royal Society B* 277: 2475–2483.
- SEDGWICK, J. A. 2001. Geographic variation in the song of Willow Flycatchers: differentiation between *Empidonax traillii adastus* and *E. t. extimus*. *Auk* 118: 366–379.
- SEUTIN, G. 1987. Female song in Willow Flycatchers. *Auk* 104: 329–330.
- SOGGE, M. K., T. J. KORONKIEWICZ, C. VAN RIPER III, AND S. L. DURST. 2007. Willow Flycatcher nonbreeding territory defense behavior in Costa Rica. *Condor* 109: 475–480.
- STEIN, R. C. 1958. The behavioral, ecological and morphological characteristics of two populations of the Alder Flycatcher: *Empidonax traillii* (Audubon). New York State Museum Science Service Bulletin, Albany, NY.
- . 1963. Isolating mechanisms between populations of Traill's Flycatchers. *Proceedings of the American Philosophical Society* 107: 21–31.
- STODDARD, M. C., AND R. O. PRUM. 2008. Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. *American Naturalist* 171: 755–776.
- THEIMER, T. C., A. D. SMITH, S. M. MAHONEY, AND K. E. IRONSIDE. 2016. Available data support protection of the Southwestern Willow Flycatcher under the Endangered Species Act. *Condor* 118: 289–299.
- UNITT, P. 1987. *Empidonax traillii extimus*. an endangered subspecies. *Western Birds* 18: 137–162.
- USFWS (U.S. FISH AND WILDLIFE SERVICE). 1995. Final rule determining endangered status for the Southwestern Willow Flycatcher. *Federal Register* 60: 10694–10715.
- VEHRENCAMP, S. L., A. F. RITTER, M. KEEVER, AND J. W. BRADBURY. 2003. Responses to playback of local vs. distant contact calls in the Orange-fronted Conure, *Aratinga canicularis*. *Ethology* 109: 37–54.
- VERZIJDEN, M. N., C. TEN CATE, M. R. SERVEDIO, G. M. KOZAK, J. W. BOUGHMAN, AND E. I. SVENSSON. 2012. The impact of learning on

- sexual selection and speciation. *Trends in Ecology & Evolution* 27: 511–519.
- WEST-EBERHARD, M. J. 1983. Sexual selection, social competition, and speciation. *Quarterly Review of Biology* 58: 155–183.
- YARD, H. K., AND B. T. BROWN. 2003. Singing behavior of Southwestern Willow Flycatchers. *Studies in Avian Biology* 26: 125–130.
- YEH, D. J., AND M. R. SERVEDIO. 2015. Reproductive isolation with a learned trait in a structured population. *Evolution* 69: 1938–1947.
- . 2015. Genetics, morphology, and ecological niche modeling do not support the subspecies status of the endangered Southwestern Willow Flycatcher (*Empidonax traillii extimus*). *Condor* 117: 76–86.
- ZINK, R. M., AND N. K. JOHNSON. 1984. Evolutionary genetics of flycatchers. I. Sibling species in the genera *Empidonax* and *Contopus*. *Systematic Biology* 33: 205–216.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article at the publisher's website.

Table S1. Locations of playback experiments in the United States (locality, county, state, latitude, and longitude), year of

sampling, number of experimental trials (N), and putative range of Willow Flycatcher subspecies.

Table S2. Song stimulus metadata used in playback experiments to test subspecific song discrimination in Willow Flycatchers.

Fig. S1. Boxplots of pre-stimulus behavioral response (PC1 (56% of variation), y-axis) observed prior to subspecific song stimuli and white noise control (x-axis) during playback experiment trials in male Willow Flycatcher (*Empidonax traillii*) subspecies (*E. t. adastus*, *E. t. brewsteri*, *E. t. traillii*, *E. t. extimus*).

Fig. S2. Boxplots of pre-stimulus behavioral response in each study site (PC1 (56% of variation), y-axis) observed prior to subspecific song stimuli and white noise control (x-axis) during playback experiment trials in male Willow Flycatcher (*Empidonax traillii*) subspecies (*E. t. adastus*, *E. t. brewsteri*, *E. t. traillii*, *E. t. extimus*).

Fig. S3. Boxplots of behavioral responses in each study site (PC1 (43% of variation), y-axis) to subspecific song stimuli and white noise control (x-axis) during playback experiment trials in male Willow Flycatcher (*Empidonax traillii*) subspecies (*E. t. adastus*, *E. t. brewsteri*, *E. t. traillii*, *E. t. extimus*).