



HAL
open science

Auditory perception of self and others in zebra finches: evidence from an operant discrimination task

Nicole Geberzahn, Sándor Zsebők, Sébastien Derégnaucourt

► **To cite this version:**

Nicole Geberzahn, Sándor Zsebők, Sébastien Derégnaucourt. Auditory perception of self and others in zebra finches: evidence from an operant discrimination task. *Journal of Experimental Biology*, 2021, 224 (6), pp.jeb233817. <10.1242/jeb.233817>. <hal-03177576>

HAL Id: hal-03177576

<https://hal.science/hal-03177576v1>

Submitted on 3 Jul 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

RESEARCH ARTICLE

Auditory perception of self and others in zebra finches: evidence from an operant discrimination task

Nicole Geberzahn^{1,*}, Sándor Zsebők^{2,3} and Sébastien Derégnaucourt^{1,4}

ABSTRACT

Vocal communication is essential for social interactions in many animal species. For this purpose, an animal has to perceive vocal signals of conspecifics and is often also required to discriminate conspecifics. The capacity to discriminate conspecifics is particularly important in social species in which individuals interact repeatedly. In addition, auditory perception of self plays an important role for vocal learners. A vocal learner has to memorise vocalisations of conspecifics and to subsequently modify its own vocalisations in order to match the memorised vocalisations. Here, we investigated auditory perception of self and others in zebra finches (*Taeniopygia guttata*), a highly gregarious songbird species and vocal learner. We used laboratory colonies in which founder males had been previously trained to produce the same song type. This resulted in artificial dialects in the song of founders and their offspring. We investigated whether those birds would be able to discriminate between familiar and unfamiliar conspecifics based on song. Furthermore, we examined whether they would classify their own song as familiar or unfamiliar. We found that birds were able to discriminate between songs of familiar versus unfamiliar conspecifics, despite the fact that all songs were imitations of the same song type. This suggests that such discrimination is possible even based on songs with a high acoustic similarity. None of the subjects classified their own song as unfamiliar. Three out of eight males classified their own song as familiar. Thus zebra finches might recognise their own song. Further experiments are needed to confirm such self-recognition.

KEY WORDS: Conspecific discrimination, Self-perception, Bird's own song, Operant discrimination, GO-NOGO, *Taeniopygia guttata*

INTRODUCTION

The ability to discriminate among conspecifics plays an important role for communication between pair members, parents and offspring, members of a social group or rivals (Tibbetts and Dale, 2007). For example, territorial songbirds often respond less aggressively to playback of well-established territorial neighbours than to playback of strangers (the 'dear enemy phenomenon'; Temeless, 1994), demonstrating their ability for neighbour–stranger discrimination. In oscine songbirds, this ability to discriminate based on song is


closely intertwined with their ability for vocal production learning (*sensu* Janik and Slater, 1997; in the following referred to as vocal learning). Vocal learning is a rare feature in the animal kingdom (Janik and Slater, 1997; Knörmisch, 2014; Nowicki and Searcy, 2014). Next to learning of speech in humans, song learning in oscine songbirds is a particularly well studied case of vocal learning. Learning to modify one's own vocalisations according to auditory self-monitoring characterises the process of vocal learning as has been shown for humans and songbirds: both must be able to hear themselves in order to develop normal vocalisations (Doupe and Kuhl, 1999). Such auditory feedback allows vocal learners to compare their own vocal output with an internal model that they have acquired when listening to vocalisations of conspecifics in order to imitate them (Adret, 2004; Rossi and Derégnaucourt, 2020). In some species, auditory self-monitoring continues to play a role in maintaining stable vocal performance even in adulthood (reviewed in Mooney, 2018).

A bird's own song (BOS) is a salient signal when broadcast via a loudspeaker. Song-selective neurons, found throughout the song system, respond more to BOS playback than to equally complex conspecific songs and may play a role in the evaluation of auditory feedback during song learning (Brainard and Doupe, 2002; Ikeda et al., 2020). Likewise, behavioural responses to BOS playback suggest that an internal representation of BOS may serve as a reference aiding discrimination of individual conspecifics (McArthur, 1986; Cynx and Nottebohm, 1992).

Here, we studied auditory perception of self and conspecifics in zebra finches (*Taeniopygia guttata*). The zebra finch, an estrildid finch, is a highly gregarious and thus non-territorial songbird with male-only song production (Zann, 1996). This is one of the best studied songbird species for its ease of reproduction in the laboratory. Juvenile males learn their song during a sensitive period early in life often from their father (Immelmann, 1969; Zann, 1996). Song is used by females for mate choice decisions (Riebel, 2009; but see Wang et al., 2017) but might also serve for intra-group communication in this species. Under typical conditions, males of a given colony would not all produce the same song type, as they would have different male tutors – singing different song types – from which to learn their song or different parts of their song. For instance, Zann (1990) found that young males learned portions but not necessarily the entire song of a specific individual of the same colony, which was often but not always their father. The resulting diversity of song types within a colony might facilitate the ability to discriminate conspecifics. Discriminating conspecifics is probably particularly important in highly gregarious zebra finches assumed to live in fission–fusion societies. Vocal learning is thought to play a crucial role in the production of individualised vocal signatures (Elie and Theunissen, 2018). In fact, discrimination based on song has traditionally been demonstrated for zebra finches producing different song types (Clayton, 1988; Cynx and Nottebohm, 1992; Gess et al., 2011).

¹Laboratoire Ethologie Cognition Développement, UPL, Université Paris Nanterre, 92001 Nanterre Cedex, France. ²Neuro-PSI CNRS UMR 9197, Université Paris-Sud, 91405 Orsay Cedex, France. ³Centre for Ecological Research, Institute of Ecology and Botany, H-2163, Vácrátót, Hungary. ⁴Institut Universitaire de France, 75005 Paris, France.

*Author for correspondence (nicole.geberzahn@parisnanterre.fr)

 N.G., 0000-0001-9945-7067; S.Z., 0000-0003-4463-4753; S.D., 0000-0003-0160-1986

In another songbird species, the song sparrow (*Melospiza melodia*), males actually confused similar song types of different singers, suggesting that song sparrow songs of the same type do not easily allow identification of the singer (Beecher et al., 1994). In contrast, in zebra finches, individual identity is coded not only in learned song but also in unlearned calls (D'Amelio et al., 2017; Elie and Theunissen, 2018), suggesting that discrimination of individual conspecifics based on less individualised vocalisations is possible as well. Furthermore, we have recently shown that zebra finches can individually recognise conspecifics even if they all produce imitations of the same song type (Geberzahn and Derégnaucourt, 2020). Such individual recognition was based on individual variation in spectro-temporal details of syllables, i.e. unique variants of syllables in that same song type.

Here, we studied whether zebra finches can discriminate between songs produced by unfamiliar versus familiar conspecifics despite a lack of strongly individualised vocal signatures (i.e. when all conspecifics sing the same song type). Furthermore, we investigated whether zebra finches perceive BOS playback as song produced by conspecifics of their own group or as song produced by unknown individuals. To this end, we made use of two colonies of zebra finches housed in visually and acoustically separated aviaries throughout the experiment. Thus, even though those colonies were completely separated from each other, they were both founded by males having all vocally learned to imitate the same song type (L. Le Maguer, N.G., L. Nagle and S.D., unpublished results).

We predicted that male zebra finches would be able to discriminate songs of male individuals of their own colony from those of another colony despite the fact that all the songs were imitations of the same song type. We made this prediction because discriminating conspecifics should be important in this highly gregarious species. Given the prominent role of BOS as a salient signal, as revealed by neurobiological and behavioural studies (reviewed in Derégnaucourt and Bovet, 2016), and the fact that subjects should be very familiar

with their own song, we furthermore predicted that when having to classify BOS playback either into the category (1) songs from individuals of their own colony, i.e. familiar songs, or (2) songs from individuals of an unknown colony, i.e. unfamiliar songs, male zebra finches would classify BOS playback as song from a familiar subject of their own colony.

MATERIALS AND METHODS

Subjects and housing conditions

For the operant conditioning procedure, we used eight adult male zebra finches, *Taeniopygia guttata* (Vieillot 1817) aged 1778 ± 290 days (mean \pm s.d.) from one of our different communal aviaries at the University Paris Nanterre, France, that are all acoustically and visually separated from each other. Before and after the experiment, these males were housed in the same mixed-sex aviary. Each bird was identified using a numbered red-coloured ring. The colony in this aviary was originally founded by males that had been selected from amongst a pool of males according to the quality of their imitation. These males had been trained previously to produce the same song type by live tutoring or operant tutoring (Derégnaucourt et al., 2013). Founders as well as their offspring could reproduce freely over the course of about 400 days of a communal breeding experiment. Our eight subjects were male offspring of this colony and all sang an imitation of the same song type (Fig. 1; similarity to the song model based on comparisons of one representative song per subject with the song model: $82 \pm 16\%$, mean \pm s.d., percentage similarity score procedure in Sound Analysis Pro; Tchernichovski et al., 2000; <http://soundanalysispro.com>). These eight males had hatched in six different clutches (subject 1522 and 1523 were from the same clutch; subject 1567 and 1570 were from the same clutch). We have no information about the identity of the social and genetic parents of these male offspring, or about the genetic relationship between them. Subjects had *ad libitum* access to a commercial tropical seed mixture, egg food, cuttlebone, bird grit and water; the

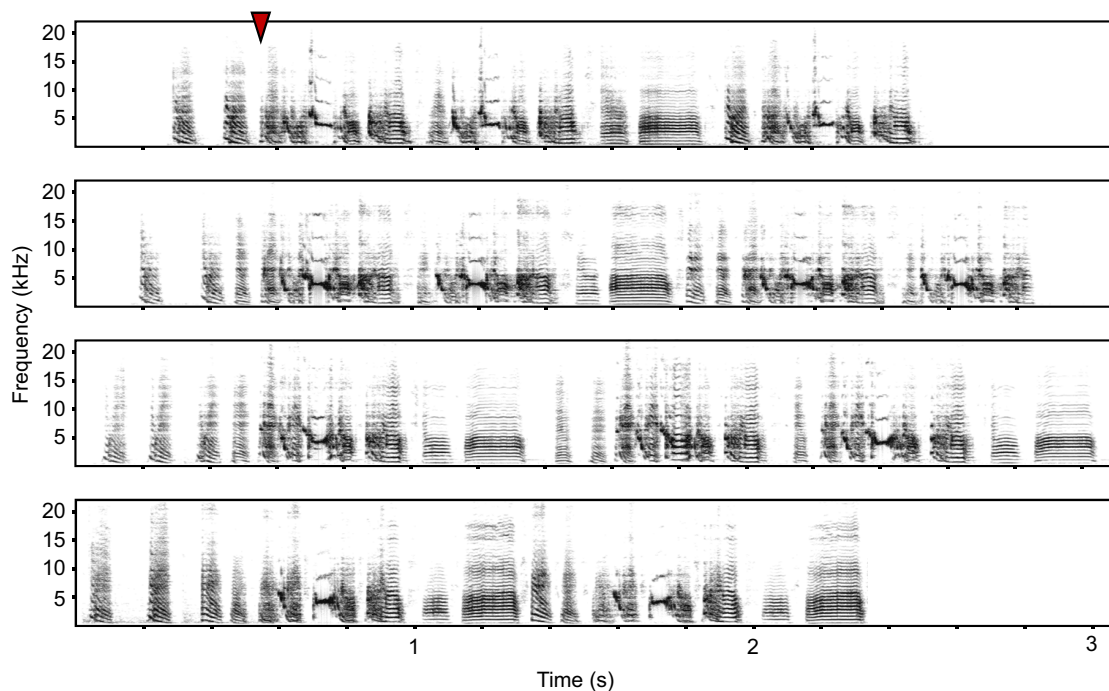


Fig. 1. Spectrograms of the song type produced by all males used in this study. Shown is one rendition produced by each of four subjects (from top to bottom: 1510, 1522, 1567 and 1581). For better comparability, all spectrograms are aligned to the onset of the same syllable indicated by the triangle on top of the first spectrogram.

food was supplemented with fresh lettuce once a week. Subjects were held on a 14 h:10 h light:dark cycle (lights on 08:00–22:00 h Central European Time). This study was conducted according to Association for the Study of Animal Behaviour guidelines on animal experimentation as well as the European law on animal experimentation. Experimental authorisation was provided by the French Ministry for National Education, Higher Education and Research (authorisation no. 02609.02).

Apparatus for operant conditioning

For the experiments, birds were transferred to custom-built sound-isolation chambers measuring 58 cm×51 cm and 57 cm high (internal measurements), lined with acoustic foam and containing a metal wire cage equipped with two perches. Next to one perch, subjects could peck on either of two keys (one situated to the right, the other to the left of the perch). At the other end of the same perch to the keys, a feeder (containing tropical seed mixture and egg food) was located to which access was blocked by a transparent plastic window. Access to food was possible only by mastering the task (see below), which led to the plastic window being pulled up by a string connected to the arm of a stepping motor (Modelcraft Premium RC-CarServo 4519DBB). A camera (HandyKam 420TVL COL CCD 12V PAL) placed inside the cage allowed monitoring of the birds' behaviour. A speaker (Yamaha Monitor Speaker MS101 III, frequency response: 30 Hz to 20 kHz) was placed behind the wire cage. All experimental events were controlled and recorded by a custom-written MATLAB program (see Gess et al., 2011) that we had slightly modified.

Experimental procedures

Social demonstration of the basic task

During this phase, subjects were continuously kept in the same experimental cage together with an expert conspecific (either a familiar male originating from the same colony or an unfamiliar female), allowing them to observe the expert bird mastering the basic task. Thus, both the subject and the expert conspecific had access to the same operant keys and feeder. The basic task consisted of pecking on the right key in order to elicit the playback of a song and to subsequently peck on the left key to open the feeder for 10 s. Subjects had access to cuttlebone, grit and water at all times. They had access to tropical seed mixture and egg food in the feeder only if the expert bird or the subject itself executed the basic task. The same stimulus was used for all demonstration sessions. This stimulus was a different song type from the stimuli used in the subsequent operant discrimination task. We monitored the behaviour and the mass of the birds in order to make sure that both the subject and the demonstrator gained regular access to the feeder (if not, we manually opened the feeder regularly for short periods of time).

Training phase – learning to discriminate the first set of stimuli

Once a subject reliably executed the basic task in the absence of the expert conspecific, he received training to discriminate two categories of song with a first set of stimuli ($N=16$ song stimuli). As before, pecking the right key elicited the playback of one of those stimuli. If a given stimulus belonged to the category GO ($N=8$ song stimuli, see 'Stimuli', below, for further details) the bird had to subsequently peck on the left key (i.e. to give a GO response) in order to obtain access to food for 10 s. However, if the stimulus belonged to the category NOGO ($N=8$ song stimuli), he had to withhold such a pecking response as pecking in this case would cause the light to go off for 10 s (constituting a mild punishment). Stimuli were presented in repeating sets of 16 songs (8 GO stimuli and 8 NOGO stimuli), but the order of stimulus presentation in each

set was random. Once a subject made correct responses to at least 75% of the stimuli for three or more consecutive blocks of 100 trials, he was transferred to the next experimental phase of the protocol in which he received a second set of stimuli. Two subjects failed to reach this criterion. We transferred them to the next experimental phase after their learning curves had reached an asymptotic plateau. Whether they had reached such a plateau was evaluated visually on figures plotting the percentage of correct responses as a function of trials, indicating that the maximum rate of correct responses was obtained. Both of them reached an average of 69% correct responses over the last 5 blocks of 100 trials.

Generalisation phase – transfer to the second set of stimuli

During the generalisation phase, birds were exposed to a second, novel set of stimuli ($N=16$ novel song stimuli). Eight of these stimuli belonged to the same category as the GO stimuli in the previous training phase (see 'Stimuli' below, for further details) and the other eight belonged to the same category as the previous NOGO stimuli. The aim was to test whether subjects had formed categories during the training phase, and thus were able to transfer knowledge about these categories from original stimuli to untrained stimuli or whether they had learned how to respond to each single original stimulus (rote memorisation of individual stimuli) without forming categories. The birds had to reach the same criteria of correct responses as in the training phase in order to be transferred to the next phase.

The two sets of stimuli combined – acclimation to reduced rate of reinforcement

In this phase, subjects were exposed to the two sets of stimuli combined ($N=32$, 16 GO, 16 NOGO). This time stimuli were presented in repeating sets of 32 songs, whereby the order of stimulus presentation in each set was random. As soon as they reached the criteria of correct responses (see 'Training phase – learning to discriminate the first set of stimuli', above), the rate of reinforcement for correct responses was lowered from 100% to 80% in order to them get used to a reduced rate of reinforcement. Subjects were transferred to the next experimental phase once they had reached the criteria of correct responses (see 'Training phase – learning to discriminate the first set of stimuli', above) with this lowered reinforcement rate.

Test phase – exposure to BOS

Each subject continued to respond to the 32 song stimuli (two previous sets combined, random presentation within repeated sets of 32 stimuli, reinforcement rate at 90%, now considered as baseline stimuli) and on 9% of the trials one of eight different renditions of BOS was presented (BOS stimuli were never reinforced). The test phase went on until a subject was exposed 25 times to each of the eight renditions of BOS.

Additional generalisation phase

In an additional generalisation phase, birds continued to respond to the 32 baseline song stimuli (two previous sets combined, random presentation within repeated sets of 32 stimuli, reinforcement rate at 90%) and on 6% of the trials, one of 16 different novel stimuli (recorded from 16 other males) from the categories GO ($N=8$) and NOGO ($N=8$) was presented (they were never reinforced). This additional generalisation phase served to test once more whether subjects had formed categories during previous experimental phases or whether they had learned how to respond to each single stimulus without forming categories. This phase went on until a subject was exposed 10 times to each of the 16 stimuli.

Stimuli

All songs used as stimuli were undirected zebra finch songs recorded from males that were individually housed in sound-isolation chambers. A microphone (Behringer C-2) was placed above the cage and songs were digitally recorded into wav files (sampling rate: 44.1 kHz, accuracy: 16 bit) using PreSonus AudioBox 1818VSL and Sound Analysis Pro software run on a PC (Tchernichovski et al., 2004). Files were high-pass filtered at 0.45 kHz and amplitudes were normalised to 90% (i.e. the maximum amplitude of the sound sequence was set to 90% of the dynamic range using Goldwave Inc. software; www.goldwave.com/). All stimuli used from the training phase onwards were imitations of the same song type. Stimuli were of similar duration (2855 ± 437 ms, mean \pm s.d.). Within each set of stimuli (for the training and the generalisation phase), birds had to discriminate between eight versions of this song type produced by four familiar males (two song renditions of each male, ‘familiar songs’ in the following) versus eight versions of this same song type produced by four unfamiliar males (two song renditions of each male, ‘unfamiliar songs’ in the following). Familiar males were members of the same colony as the subject (four males chosen from a pool of $N=48$ males consisting of all the males of the colony except the founder males). If subjects had received social demonstration by a familiar male, songs of this familiar male were not chosen as stimuli for this particular subject. Likewise, we obviously did not chose song of the subject itself (BOS) from this pool of males as familiar song. Unfamiliar males were members of a second colony that had also been founded by males imitating the very same song type, but that the subject had never seen or heard before (four males chosen from a pool of $N=34$ males consisting of all the males of the colony except the founder males). Familiar and unfamiliar males all produced the same song type. Members of these two colonies had always been kept in different rooms and had thus never been in acoustic, visual or physical contact. Four subjects received familiar songs as GO stimuli and unfamiliar songs as NOGO stimuli and for the other four subjects this association of familiar/unfamiliar songs with the categories GO/NOGO was reversed. The sets of stimuli (for the training and the generalisation phase) were chosen individually for each subject in such a way that one familiar stimulus and one unfamiliar stimulus matched each other in their similarity to the subject’s BOS (percentage similarity score procedure in Sound Analysis Pro; Tchernichovski et al., 2000; familiar: $63.3 \pm 10.2\%$, unfamiliar: $63.4 \pm 10.3\%$, means \pm s.d.). This allowed us to rule out that a difference in similarity of BOS when compared with familiar songs as opposed to when compared with unfamiliar songs might bias the behavioural responses in the test phase (when subjects were exposed to BOS). Furthermore, detailed song analysis revealed high similarity between males of the two colonies (from which stimuli were chosen) in other aspects of the song such as overall song bout structure, composition of syllable and element repertoires as well as inter-syllabic gap distributions (L. Le Maguer, N.G., L. Nagle and S.D., unpublished results).

Analyses

Analyses were conducted on individual and group levels (see Spierings and ten Cate, 2016; Geberzahn and Derégnaucourt, 2020). Analysis on individual levels was carried out with a binominal test and a subsequent Benjamini–Hochberg correction for eight comparisons (Benjamini and Hochberg, 1995). To test for group level differences between responses to stimuli from the categories GO and NOGO in the generalisation phases, we used general linear mixed-effect models. The proportion of GO responses was the response variable. To control for a possible impact of clutch identity, we included this variable as a predictor. Subject identity was a random factor. To compare the proportion of

GO responses to BOS stimuli between those birds that had been previously trained with familiar songs as GO stimuli and those that had been previously trained with unfamiliar songs as GO stimuli, we used a two sample *t*-test. We used R (version 3.5.0; <http://www.cran.r-project.org>), except for the Benjamini–Hochberg corrections which were calculated using spreadsheets.

In addition, we report d' , a measure of the ability to discriminate stimuli taking into account potential response bias (Macmillan and Creelman, 2005), using the formula:

$$d' = z(\text{hit rate}) - z(\text{false alarm rate}), \quad (1)$$

where hit rate is the rate of correct responses to stimuli of the category GO and false alarm rate is the rate of incorrect responses to stimuli of the category NOGO (hit rate and false alarm rate being terms established in signal detection theory; see Macmillan and Creelman, 2005). One occasion of a proportion of 0 for hit rate was converted to $1/(2N)$ in order to avoid infinite values, where N is the number of trials on which this proportion was based (Macmillan and Creelman, 2005).

First generalisation phase

The proportion of hits (correct responses to stimuli of the category GO) averaged over the first four exposures for each of eight stimuli was measured against a binominal distribution with a probability equivalent to chance level. The same was done for false alarms (incorrect responses to stimuli of the category NOGO). Chance level was calculated for each subject based on his individual performance just before the transfer to the novel set of stimuli. To this end, we first calculated for each subject the proportion of hits over the last four exposures for each of eight stimuli in the preceding training phase. Then, we likewise calculated for each subject the proportion of false alarms over the last four exposures for each of eight stimuli in the preceding training phase. Chance level was then calculated as the mean of these two proportions following the formula: chance level = (proportion of hits in preceding phase + proportion of false alarms in preceding phase) / 2.

We only considered the last four exposures of the training phase in order to capture the performance right at the end of the training phase. Likewise, we selected only the first four exposures of each stimulus in the generalisation phase to reduce the risk that birds had already learned to associate this new set of stimuli with the categories GO and NOGO (note that in this first generalisation phase all stimuli were reinforced). In addition to the individual-based analysis, we conducted a comparison of the proportion of correct responses to each stimulus of the category GO (first four exposures for each of eight stimuli) and the incorrect responses to each stimuli of the category NOGO (first four exposures for each of eight stimuli) of the generalisation phase at the group level. We also report the mean and s.e.m. for d' values over all subjects.

Test phase – exposure to BOS

The proportion of GO responses to BOS stimuli (first four exposures for each of eight stimuli) was measured against a binominal distribution with a probability equivalent to chance level. Here, chance level was individually calculated as the grand mean of the average of GO responses to the last six presentations of GO (i.e. hits) and the average of GO responses to the last six presentations of NOGO stimuli (i.e. false alarms) prior to a given BOS stimulus presentation, thus reflecting the performance to already well known stimuli (baseline stimuli) during the test phase. We selected only the first four exposures of each BOS stimulus because cumulative plots

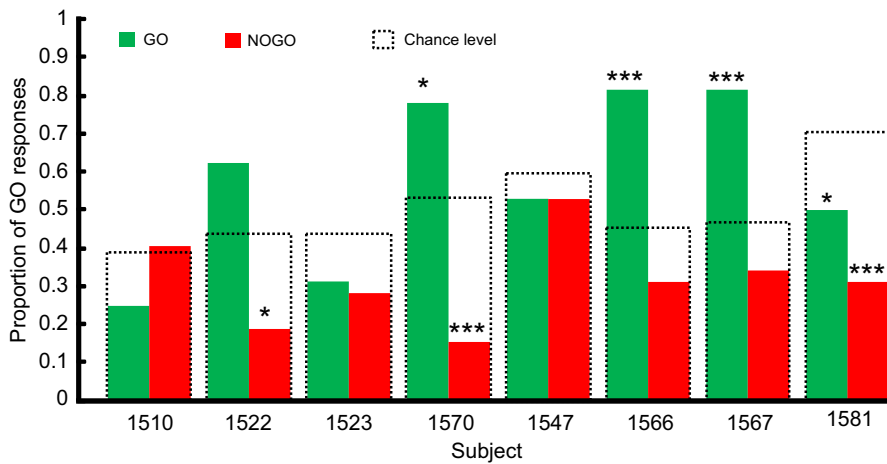


Fig. 2. Proportion of GO responses to presentation of a novel set of stimuli at the beginning of the first generalisation phase.

Responses to GO stimuli (green) and NOGO stimuli (red) are shown. Birds 1510, 1522, 1523 and 1570 were trained with familiar songs as GO stimuli and birds 1547, 1566, 1567 and 1581 were trained with unfamiliar songs as GO stimuli. These novel stimuli were reinforced (GO: reward, NOGO: punishment). Chance levels were calculated for each subject as the mean of the proportion of hits and false alarms at the end of the preceding training phase.

* $P_{\text{adjusted}} \leq 0.05$, *** $P_{\text{adjusted}} \leq 0.001$. For details of statistics, see Table 1. Bird ID is based on individual ring numbers.

of pecking responses to BOS stimuli showed asymptotic curve shapes suggesting a gradual decline (Fig. S1), probably due to the fact that the birds learned over the course of this experimental phase that BOS stimuli were never reinforced. Such a decline was not yet visible during the first 32 stimulus presentations (corresponding to four presentations of each of eight stimuli).

Additional generalisation phase

The proportion of correct responses to each stimulus of the category GO (first four exposures) and of incorrect responses to each stimulus of the category NOGO (first four exposures) of this second generalisation phase was measured against a binominal distribution with a probability equivalent to chance level. Chance level was calculated for each subject by taking into account the average of the last six presentations of already known GO and NOGO stimuli prior to the presentation of a given novel stimulus (either of the GO or NOGO category). In addition to this individual-based analysis, we conducted a comparison at the group level. To this end, we compared the proportion of correct responses to each stimulus of the category GO (first four exposures) and the incorrect responses to each stimulus of the category NOGO (first four exposures). We also report the mean and s.e.m. for d' values over all subjects.

RESULTS

Discrimination between songs produced by familiar and unfamiliar males

All eight subjects learned to discriminate between stimuli of the category GO versus NOGO of the first set of stimuli. That means they learned to discriminate between songs produced by familiar

males and songs produced by unfamiliar males. The number of training blocks (each block counting 100 trials) necessary to pass to the first generalisation phase ranged from 9 to 74 (mean 41.5).

Performance after transfer to the second (novel) set of stimuli

Just after having been transferred to the second and novel set of GO and NOGO stimuli, three out of eight subjects gave a higher proportion of correct responses to the GO stimuli (i.e. a higher proportion of hits) than expected by chance (Fig. 2, Table 1). Only one subject gave significantly fewer correct responses to the GO stimuli than expected by chance (subject 1581; see Fig. 2, Table 1). Three out of eight subjects responded incorrectly significantly less often to NOGO stimuli than expected by chance, i.e. they gave a false alarm less often than expected by chance (Fig. 2, Table 1). At the group level, birds responded with a significantly higher proportion of GO responses to GO stimuli than to NOGO stimuli (Fig. 3, GLMM: $F_{1,7} = 8.06$, $P = 0.025$). The mean \pm s.e.m. d' over all subjects was 0.72 ± 0.28 , which is above chance level ($d' = 0$).

Responses to BOS

For three out of eight birds the proportion of GO responses differed significantly from chance level (Fig. 4, Table 2). One of them (1523) had previously been trained to associate familiar songs with the category GO and unfamiliar songs with the category NOGO. This subject responded significantly more often to BOS with a GO response, suggesting that he associated his own song with the category GO (i.e. familiar songs for him). The other two subjects whose responses differed from chance level (1567, 1581) had

Table 1. Proportion of GO responses to the new set of GO stimuli (i.e. hits) and to the new set of NOGO stimuli (i.e. false alarms)

Bird ID	GO stimuli			NOGO stimuli		
	Proportion of responses	P	P_{adjusted}	Proportion of responses	P	P_{adjusted}
1510	0.25 (8/32)	0.146	0.194	0.41 (13/32)	0.858	0.980
1522	0.63 (20/32)	0.048	0.077	0.19 (6/32)	0.004	0.011
1523	0.31 (10/32)	0.212	0.242	0.28 (9/32)	0.078	0.155
1570	0.78 (25/32)	0.004	0.012	0.16 (5/32)	<0.001	<0.001
1547	0.53 (17/32)	0.477	0.477	0.53 (17/32)	1.000	1.000
1566	0.81 (26/32)	<0.001	<0.001	0.31 (10/32)	0.154	0.247
1567	0.81 (26/32)	<0.001	<0.001	0.34 (11/32)	0.214	0.286
1581	0.50 (16/32)	0.019	0.037	0.31 (10/32)	<0.001	<0.001

Proportions are shown for individual birds in the first generalisation phase, measured against a binominal distribution with a probability equivalent to chance level. Significant P -values are in bold. Adjusted P -values were assessed by Benjamini–Hochberg correction for multiple testing. Bird ID is based on individual ring numbers.

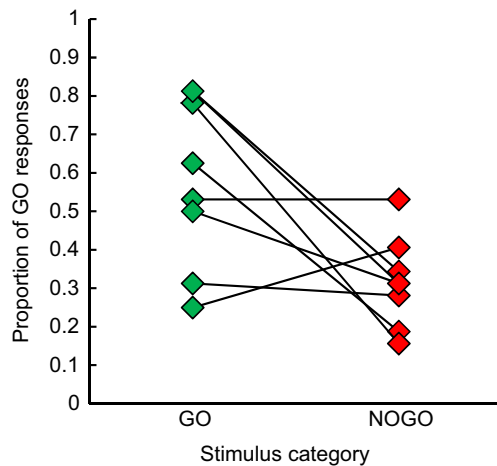


Fig. 3. Proportion of GO responses to presentation of a novel set of stimuli at the beginning of the first generalisation phase. GO responses are hits and NOGO responses are false alarms. Comparison was made at the group level: GLMM: $F_{1,7}=8.06$, $P=0.025$.

previously been trained to associate unfamiliar songs with the category GO and familiar songs with the category NOGO. These two subjects exhibited a significantly lower proportion of GO responses to BOS than expected by chance. This suggests that they associated their own song with the category NOGO (i.e. familiar songs for them).

On a group level, the proportion of GO responses to BOS stimuli was 0.48 ± 0.12 (mean \pm s.d.) for those birds that had previously been trained with familiar songs as GO stimuli. This proportion was 0.30 ± 0.19 (mean \pm s.d.) for those birds that had previously been trained with unfamiliar songs as GO stimuli. These two groups of birds did not significantly differ in their proportion of GO responses to BOS (two-sample t -test, $t=1.675$, d.f.=6, $P=0.145$).

Performance in the second generalisation phase

During the first four exposures to novel stimuli of this second generalisation phase (half of which were from the same category as the baseline GO stimuli and the other half from the same category as the baseline NOGO stimuli), none of the subjects gave a higher proportion of correct responses to the GO stimuli than expected by chance (Fig. 5, Table 3). One subject gave significantly fewer correct responses to the GO stimuli than expected by chance (subject 1510; see Fig. 5, Table 3). Seven out of eight subjects responded incorrectly significantly less often to NOGO stimuli than expected by chance (Fig. 5, Table 3).

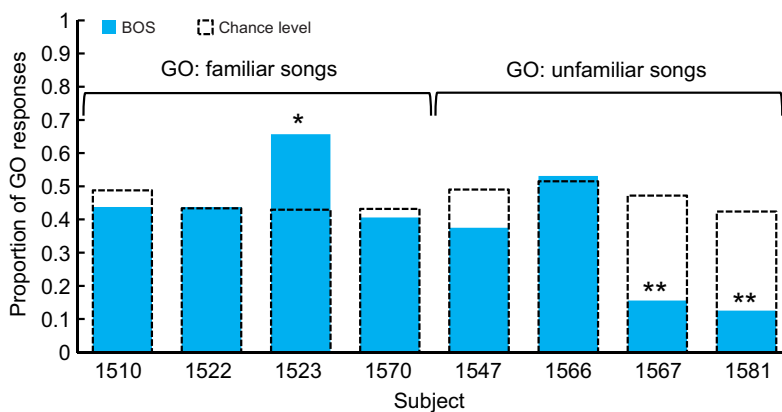


Fig. 4. Proportion of GO responses to presentation of bird's own song (BOS) of individual subjects during the test phase. Chance levels were calculated for each subject as the grand mean of the average proportion of responses to the last six presentations of GO and the average proportion of responses to the last six presentations of NOGO stimuli prior to a given BOS stimulus presentation. * $P_{\text{adjusted}}\leq 0.05$, ** $P_{\text{adjusted}}\leq 0.01$. For details of statistics, see Table 2. Bird ID is based on individual ring numbers.

Compared with the first generalisation test, subjects refrained from giving a GO response much more often in this second generalisation test. This might be due to having been exposed to stimuli that were never reinforced (BOS stimuli) during the previous test phase. Thus, subjects might have learned that certain stimuli are never reinforced. However, at a group level, birds responded with a significantly higher proportion of GO responses to novel GO stimuli than to novel NOGO stimuli (Fig. 6; GLMM: $F_{1,7}=10.34$, $P=0.015$). Furthermore, the means \pm s.e.m. d' over all subjects was 0.53 ± 0.24 , which is above chance level ($d'=0$).

DISCUSSION

Male zebra finches were able to discriminate between songs produced by several familiar versus several unfamiliar conspecifics, even though all of these songs were imitations of the same song type. In addition, when having to classify BOS playback into the category familiar or unfamiliar songs, three out of eight male zebra finches treated their own song as being produced by a familiar rather than as being produced by an unfamiliar conspecific.

Discrimination based on hardly individualised song

Zebra finches correctly classified novel stimuli in two generalisation phases, suggesting that they were able to discriminate between songs from familiar and unfamiliar conspecifics, even if the songs that those conspecifics sang were very similar. Our set of stimuli was built in such a way that familiar stimuli were as similar to a given subject's own song as were the unfamiliar stimuli. As a consequence, there should be no clear acoustic differences in the songs of familiar and unfamiliar conspecifics, or, in other words, no clear vocal group signatures, that would allow the subjects to carry out this discrimination. Hence, a likely mechanism by which such discrimination could come about is individual recognition. In this scenario, a subject would recognise songs as being either produced by a particular familiar individual or not. Based on the presence or absence of such recognition, he would then classify those songs as either familiar or unfamiliar. In fact, we have recently demonstrated that individual recognition is possible in this species even if a finch has to distinguish songs that are imitations of the same song type (Geberzahn and Derégnaucourt, 2020). The results of the current study are clear when analysed at the group level. Interestingly, when analysed at the individual level, subjects of the current study performed less well in the generalisation tasks compared with subjects in Geberzahn and Derégnaucourt (2020), which supports our idea that individual recognition rather than familiar–unfamiliar discrimination might be the prevailing mechanism here. Zebra finches are social birds living in flocks, with a high level of

Table 2. Proportion of GO responses to baseline and bird's own song (BOS) stimuli of individual birds in the test phase

Bird ID	GO stimuli	Baseline		Test		
		GO	NOGO	BOS	<i>P</i>	<i>P</i> _{adjusted}
1510	Familiar songs	0.81	0.17	0.44 (14/32)	0.600	0.960
1522	Familiar songs	0.72	0.15	0.44 (14/32)	1.000	1.000
1523	Familiar songs	0.67	0.19	0.66 (21/32)	0.012	0.031
1570	Familiar songs	0.77	0.10	0.41 (13/32)	0.859	0.985
1547	Unfamiliar songs	0.63	0.35	0.38 (12/32)	0.218	0.437
1566	Unfamiliar songs	0.92	0.11	0.53 (17/32)	0.862	0.985
1567	Unfamiliar songs	0.88	0.07	0.16 (5/32)	<0.001	0.002
1581	Unfamiliar songs	0.68	0.16	0.13 (4/32)	<0.001	0.002

Responses to BOS were measured against a binomial distribution with a probability equivalent to chance level. Significant *P*-values are in bold. Adjusted *P*-values were assessed by Benjamini–Hochberg correction for multiple testing. Bird ID is based on individual ring numbers.

dispersion (Zann, 1996). It should be noted that in the current study we were pushing the capacity to discriminate between familiar and unfamiliar conspecifics to an extreme by testing subjects with non-individualised song types: during social interactions in flocks, wild zebra finches would hardly ever be exposed to several different conspecifics that all sing the same song type. Furthermore, groups of wild zebra finches probably change in size by means of the fission and fusion of subgroups, resulting in frequent turnover in flock composition. Such dynamics would result in large and frequently changing communication networks such as in parrots (Balsby et al., 2012). However, whether zebra finches have group structures characterised by fission/fusion dynamics remains speculative because the data in the wild are relatively scarce. Anyway, it is likely that in each of our colonies, each individual experienced some social interactions, positive or negative, with other birds. It has already been evidenced that male song could be used as a vocal label in such interactions (Miller, 1979). In ravens (*Corvus corax*), it has been experimentally shown that subjects associate individual calls with social status by using playback of vocal interactions that violate or comply with the current rank hierarchy (Massen et al., 2014).

In oscine songbirds, vocal learning is thought to play a crucial role in the production of individualised vocal signatures (Elie and Theunissen, 2018). Vocal learning allows for a larger inter-individual diversity of song types and thus more individualised vocalisations because different juveniles learn different songs from different male tutors, mainly their fathers (Derégnaucourt, 2011). In our case, birds were able to discriminate conspecifics based on song types shared between all colony members and thus must have used other acoustic features such as individual variation in spectro-temporal details of syllables (Geberzahn and Derégnaucourt, 2020).

Such individual variation might reflect individually unique voice characteristics related to individual properties of the vocal organ and the vocal tract (Lambrechts and Dhondt, 1995). As such, voice characteristics should be inherent to an individual's vocalisations; the ultimate proof of this would be to show that subjects tested in an experimental recognition task are able to generalise between different types of songs of one and the same individual (e.g. Beecher et al., 1994). Here, we did not test for such generalisation. Therefore, whether vocal discrimination in our study was due to individual voice cues remains speculative.

Perception of BOS

Five out of eight zebra finches classified BOS playback as neither familiar nor unfamiliar: their behavioural responses did not differ significantly from chance level. It is worth mentioning that none of the subjects classified BOS playback as unfamiliar song. Classification of BOS as unfamiliar song would be expected if subjects completely failed to recognise BOS playback as their own song. The inability to classify BOS playback as either familiar or unfamiliar in those five subjects might thus indicate an ambiguity in recognising their own song. Such ambiguity might be due to differences between passive exposure to BOS via playback and exposure to a bird's own song whilst singing naturally. There are at least three types of potential differences. First, perception of song produced by the bird itself might be characterised not only by air conduction but also by bone conduction, such as in humans (Maurer and Landis, 1990), even though the effect might be smaller in birds, given that their bones are much thinner than in humans. Second, the precise acoustic properties of the song broadcast by a loudspeaker might slightly differ from properties of song produced by the bird. Third, when singing, birds might produce an efference copy of the

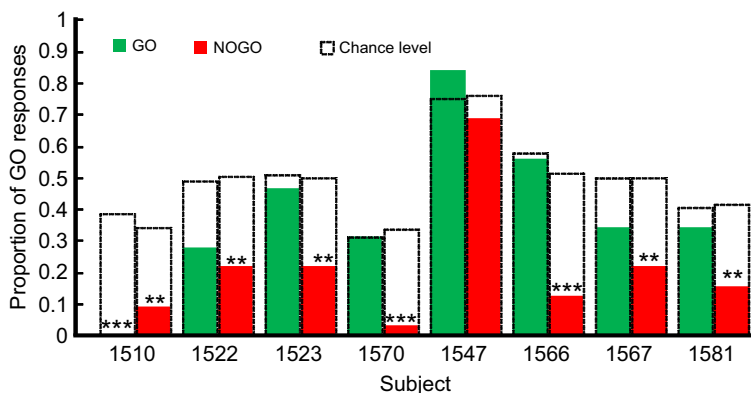


Fig. 5. Proportion of GO responses to presentation of a novel set of stimuli at the beginning of the second generalisation phase.

Responses to stimuli from the same category as GO stimuli (green) and from the same category as NOGO stimuli (red) are shown. These novel stimuli were never reinforced. Birds 1510, 1522, 1523 and 1570 were initially trained with familiar songs as GO stimuli and birds 1547, 1566, 1567 and 1581 were trained with unfamiliar songs as GO stimuli. Chance levels were calculated for each subject as the grand mean of the average proportion of responses to the last six presentations of baseline GO stimuli and the average proportion of responses to the last six presentations of baseline NOGO stimuli prior to a given novel stimulus presentation. **P*_{adjusted} ≤ 0.05, ***P*_{adjusted} ≤ 0.01, ****P*_{adjusted} ≤ 0.001. For details of statistics, see Table 3. Bird ID is based on individual ring numbers.

Table 3. Proportion of GO responses to novel stimuli from the category GO (i.e. hits) and NOGO (i.e. false alarms)

Bird ID	GO stimuli			NOGO stimuli		
	Proportion of responses	<i>P</i>	<i>P</i> _{adjusted}	Proportion of responses	<i>P</i>	<i>P</i> _{adjusted}
1510	0.00 (0/32)	<0.001	<0.001	0.09 (3/32)	0.002	0.003
1522	0.28 (9/32)	0.021	0.084	0.22 (7/32)	0.001	0.003
1523	0.47 (15/32)	0.725	0.967	0.22 (7/32)	0.002	0.003
1570	0.31 (10/32)	1.000	1.000	0.03 (1/32)	<0.001	<0.001
1547	0.84 (27/32)	0.306	0.613	0.69 (22/32)	0.306	0.306
1566	0.56 (18/32)	1.000	1.000	0.13 (4/32)	<0.001	<0.001
1567	0.34 (11/32)	0.110	0.294	0.22 (7/32)	0.002	0.003
1581	0.34 (11/32)	0.590	0.944	0.16 (5/32)	0.002	0.003

Proportions are shown for individual birds in the second generalisation phase measured against a binomial distribution with a probability equivalent to chance level. Significant *P*-values are in bold. Adjusted *P*-values were assessed by Benjamini–Hochberg correction for multiple testing. Bird ID is based on individual ring numbers.

corresponding motor command (Troyer and Doupe, 2000) and should receive proprioceptive information from the vocal organ. Such an efference copy as well as proprioceptive information should theoretically be absent in the nervous system when BOS is broadcast from a loudspeaker. To conclude, we speculate that birds responding at chance level classified BOS as neither familiar nor unfamiliar because perception of such BOS playback slightly differed from perception of their own song when they would really sing themselves.

It is also worth mentioning that neuronal responses to BOS playback are modulated by the behavioural state in the zebra finch: neurons with responses to BOS playback in anaesthetised or sleeping animals do not always exhibit activity when finches are awake, indicating that auditory responses to sounds are ‘gated’ by the behavioural state of the bird. Thus, little or no auditory-evoked activity is detectable in the so-called song control system during periods of wakefulness (Cardin and Schmidt, 2003). But the neural processing of BOS is a multistage process and the inter-individual variability in individual birds’ arousal or attentional state might explain why some finches classified BOS as familiar (cases further discussed in the following paragraph), whereas others responded at chance level.

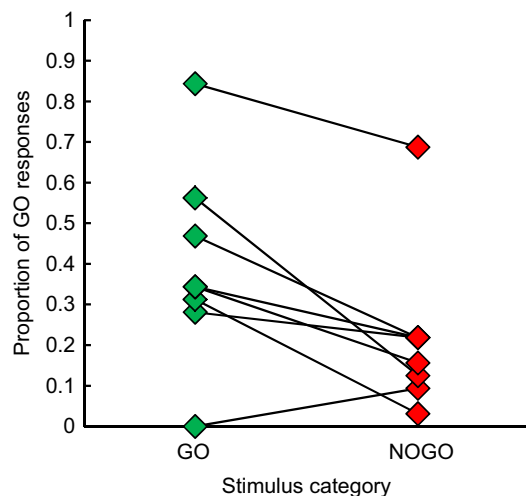


Fig. 6. Proportion of GO responses to presentation of generalisation stimuli of individual subjects at the beginning of the second generalisation phase. Generalisation stimuli were novel stimuli of the categories GO and NOGO. Comparison at the group level: GLMM: $F_{1,7}=10.34$, $P=0.015$.

Three of our zebra finches did provide behavioural responses to BOS that differed from chance and those males classified BOS as a familiar song: they responded significantly more often with the behavioural response associated with the category of familiar songs. This indicates that they perceived BOS as familiar. This is in line with the findings of a phonotaxis experiment, in which adult male zebra finches expressed a robust behavioural preference for the playback of their BOS compared with conspecific male song (Remage-Healey et al., 2010). The role of familiarity in discriminating conspecific songs has been studied by Cynx and Nottebohm (1992) using operant conditioning. They showed that males trained to discriminate between their own song and another song from their aviary reached criterion within the fewest number of trials. Males that had to discriminate between two conspecific songs from their aviary needed more trials. The most training was required by males discriminating between songs they had not heard before (Cynx and Nottebohm, 1992). Whilst differences in such earlier studies were suggestive of BOS being more familiar to subjects than other conspecific stimuli, the responses of those three males in our study clearly indicate that BOS stimuli were perceived as familiar by them.

During the process of song acquisition, auditory neurons in the song control system are shaped to respond best to the BOS (Margoliash, 1983; Doupe and Konishi, 1991; for a recent review, see Ikeda et al., 2020). This shaping could contribute to a bird’s ability to discriminate among conspecific songs by acting as an ‘autogenous reference’ (Margoliash, 1986). This pattern of self-responsiveness is found in adult birds even if they were raised without a tutor, indicating that self-experience is a critical factor in shaping BOS selectivity (Kojima and Doupe, 2007). Furthermore, BOS-selective auditory responses in the song control system emerge as sensorimotor learning progresses (Volman, 1993; Doupe, 1997; Solis and Doupe, 1999). And in the zebra finch, song development occurs during puberty, an important period for individual construction in humans and other animals. Such BOS-selective auditory responses might thus constitute a neuronal signature of the bird recognising its own song (Brecht and Nieder, 2020).

Such an autogenous reference could also be used for neighbour–stranger recognition in the wild by territorial birds (reviewed in Derégnaucourt and Bovet, 2016). Territorial songbirds often respond less aggressively to playback of well-established territorial neighbours than to playback of strangers (the ‘dear enemy phenomenon’; Temeless, 1994). Furthermore, some of those species have been shown to respond differently towards neighbour song broadcast from opposite versus regular boundaries (e.g. Godard and Wiley, 1995) allowing interpretation of such reduced aggression as indicative of individual recognition rather than

habituation. Thus, should territorial birds consider BOS playback as song produced by themselves, they should react less aggressively towards BOS playback than towards playback of a stranger's song. This has indeed been found for several species (e.g. song sparrows: McArthur, 1986; swamp sparrows, *Melospiza georgiana*: Searcy et al., 1981; white-throated sparrows, *Zonotrichia albicollis*: Brooks and Falls, 1975; but see Akçay and Beecher, 2020). However, playback experiments in the wild cannot rule out the possibility that a reduced territorial response of the focal bird towards BOS playback might be a consequence of a reduced territorial response of neighbouring birds eavesdropping on that playback. Those neighbours should recognise the BOS playback as the song of the neighbour and respond less vigorously. Such reduced responses of neighbours to BOS might stimulate the focal birds less than responses of neighbours to playback of strangers (Brooks and Falls, 1975).

By applying an operant discrimination task under laboratory conditions in the current study, we were able to rule out such potential influences from other birds eavesdropping on the playback of stimuli. The only way for zebra finches in the current study to get access to food was to correctly discriminate stimuli, resulting in a high motivation of the birds to provide the appropriate behavioural response. Using this approach revealed a clear result for three out of eight subjects: they perceived BOS as familiar and might have recognised themselves in the BOS playback. This could provide support for the idea that an internal representation of BOS may serve for self-recognition and as a reference aiding vocal recognition of conspecifics.

However, we cannot rule out an alternative explanation: there could have been a conspecific in the subject's aviary that produced a song virtually identical or very similar to the subject's BOS. In this case, subjects might have classified BOS as familiar not because they recognised their own song but because they misinterpreted the BOS stimuli as the song of this familiar colony member. Indeed, a comparison between all colony members revealed that 6 out of 8 birds (1510, 1523, 1571, 1547, 1566 and 1581) – including two out of the three finches that classified their BOS as familiar (1523, 1581) – had at least one colony mate with a song revealing the same percentage similarity score when compared with the subject's song as for comparison of the subject's song with itself.

Further experiments are needed to confirm that zebra finches are able to recognise their own song. If such confirmation were provided, this could then be interpreted as a certain level of self-other distinction. Furthermore, additional experiments investigating the perception of self across different sensorial modalities (e.g. combining a mirror test with playback of BOS) would be welcomed to address these issues in the future (Derégnaucourt and Bovet, 2016).

Acknowledgements

We thank Sarah Woolley for providing the MATLAB script for the GO/NOGO procedure and Chloé Huetz for helping to adapt it. Many thanks go to Henrik Brumm for discussions. We thank Thierry Aubin, Héléne Courvoisier and Laurent Nagle for their help with building the sound proof chambers. We thank Philippe Groué, Emmanuelle Martin and Ophélie Bouillet for taking care of the birds. We are grateful to two anonymous reviewers for their valuable comments on a previous version of the manuscript.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: N.G., S.D.; Methodology: N.G., S.Z., S.D.; Software: S.Z.; Formal analysis: N.G.; Investigation: N.G., S.D.; Writing - original draft: N.G.; Writing - review & editing: N.G., S.Z., S.D.; Visualization: N.G.; Project administration: S.D.; Funding acquisition: S.Z., S.D.

Funding

Funding was provided by the Agence Nationale de la Recherche (ANR-12-BSH2-0009) and the Institut Universitaire de France (IUF) to S.D., and by the Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal (grant numbers K-129215 and PD-115730) to S.Z.

Supplementary information

Supplementary information available online at <https://jeb.biologists.org/lookup/doi/10.1242/jeb.233817.supplemental>

References

- Akçay, Ç. and Beecher, M. D. (2020). Song sparrows do not discriminate between their own song and stranger song. *Behav. Process.* **178**, 104184. doi:10.1016/j.beproc.2020.104184
- Adret, P. (2004). In search of the song template. *Ann. N. Y. Acad. Sci.* **1016**, 303-324. doi:10.1196/annals.1298.005
- Balsby, T. J. S., Momberg, J. V. and Dabelsteen, T. (2012). Vocal imitation in parrots allows addressing of specific individuals in a dynamic communication network. *PLoS ONE* **7**, e49747. doi:10.1371/journal.pone.0049747
- Beecher, M., Campbell, S. E. and Burt, J. (1994). Song perception in the song sparrow: birds classify by song type but not by singer. *Anim. Behav.* **47**, 1343-1351. doi:10.1006/anbe.1994.1182
- Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Methodol.* **57**, 289-300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Brainard, M. S. and Doupe, A. J. (2002). What songbirds teach us about learning. *Nature* **417**, 351-358. doi:10.1038/417351a
- Brecht, K. F. and Nieder, A. (2020). Parting self from others: Individual and self-recognition in birds. *Neurosci. Biobehav. Rev.* **116**, 99-108. doi:10.1016/j.neubiorev.2020.06.012
- Brooks, R. J. and Falls, J. B. (1975). Individual recognition by song in white-throated sparrows. I. Discrimination of songs of neighbors and strangers. *Can. J. Zool.* **53**, 879-888. doi:10.1139/z75-101
- Cardin, J. A. and Schmidt, M. F. (2003). Song system auditory responses are stable and highly tuned during sedation, rapidly modulated and unselective during wakefulness, and suppressed by arousal. *J. Neurophysiol.* **90**, 2884-2899. doi:10.1152/jn.00391.2003
- Clayton, N. S. (1988). Song discrimination learning in zebra finches. *Anim. Behav.* **36**, 1016-1024. doi:10.1016/S0003-3472(88)80061-7
- Cynx, J. and Nottebohm, F. (1992). Role of gender, season, and familiarity in discrimination of conspecific song by zebra finches (*Taeniopygia guttata*). *Proc. Natl. Acad. Sci. USA* **89**, 1368-1371. doi:10.1073/pnas.89.4.1368
- D'Amelio, P. B., Klumb, M., Adreani, M. N., Gahr, M. L. and ter Maat, A. (2017). Individual recognition of opposite sex vocalizations in the zebra finch. *Sci. Rep.* **7**, 5579. doi:10.1038/s41598-017-05982-x
- Derégnaucourt, S. (2011). Birdsong learning in the laboratory, with especial reference to the song of the zebra finch (*Taeniopygia guttata*). *Interaction Studies* **12**, 323-349. doi:10.1075/is.12.2.07der
- Derégnaucourt, S. and Bovet, D. (2016). The perception of self in birds. *Neurosci. Biobehav. Rev.* **69**, 1-14. doi:10.1016/j.neubiorev.2016.06.039
- Derégnaucourt, S., Poirier, C., Kant, A. V., Linden, A. V. and Gahr, M. (2013). Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J. Physiol. Paris* **107**, 210-218. doi:10.1016/j.jphysparis.2012.08.003
- Doupe, A. J. (1997). Song-and order-selective neurons in the songbird anterior forebrain and their emergence during vocal development. *J. Neurosci.* **17**, 1147-1167. doi:10.1523/JNEUROSCI.17-03-01147.1997
- Doupe, A. J. and Konishi, M. (1991). Song-selective auditory circuits in the vocal control system of the zebra finch. *Proc. Natl. Acad. Sci. USA* **88**, 11339-11343. doi:10.1073/pnas.88.24.11339
- Doupe, A. J. and Kuhl, P. K. (1999). Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* **22**, 567-631. doi:10.1146/annurev.neuro.22.1.567
- Elie, J. E. and Theunissen, F. E. (2018). Zebra finches identify individuals using vocal signatures unique to each call type. *Nat. Commun.* **9**, 4026. doi:10.1038/s41467-018-06394-9
- Geberzahn, N. and Derégnaucourt, S. (2020). Individual vocal recognition in zebra finches relies on the song syllable structure rather than on song syllable order. *J. Exp. Biol.* **223**, jeb220087. doi:10.1242/jeb.220087
- Gess, A., Schneider, D. M., Vyas, A. and Woolley, S. M. (2011). Automated auditory recognition training and testing. *Anim. Behav.* **82**, 285-293. doi:10.1016/j.anbehav.2011.05.003
- Godard, R. and Wiley, R. H. (1995). Individual recognition of song repertoires in two wood warblers. *Behav. Ecol. Sociobiol.* **37**, 19-23. doi:10.1007/BF00164157
- Ikeda, M. Z., Trusel, M. and Roberts, T. F. (2020). Memory circuits for vocal imitation. *Curr. Opin. Neurobiol.* **60**, 37-46. doi:10.1016/j.conb.2019.11.002
- Immelmann, K. (1969). Song development in the Zebra Finch and other estrildid finches. In *Bird Vocalizations* (ed. R. A. Hinde). pp. 61-74. London: Cambridge University Press.

- Janik, V. M. and Slater, P. B.** (1997). Vocal learning in mammals. *Adv. Study Behav.* **26**, 59–99. doi:10.1016/S0065-3454(08)60377-0
- Knörnschild, M.** (2014). Vocal production learning in bats. *Curr. Opin. Neurobiol.* **28**, 80–85. doi:10.1016/j.conb.2014.06.014
- Kojima, S. and Doupe, A. J.** (2007). Song selectivity in the pallial-basal ganglia song circuit of zebra finches raised without tutor song exposure. *J. Neurophysiol.* **98**, 2099–2109. doi:10.1152/jn.00916.2006
- Lambrechts, M. M. and Dhondt, A. A.** (1995). Individual voice discrimination in birds. In *Current Ornithology* (ed. D. M. Power), pp. 115–139. New York: Plenum.
- Margoliash, D.** (1983). Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *J. Neurosci.* **3**, 1039–1057. doi:10.1523/JNEUROSCI.03-05-01039.1983
- Margoliash, D.** (1986). Preference for autogenous song by auditory neurons in a song system nucleus of the white-crowned sparrow. *J. Neurosci.* **6**, 1643–1661. doi:10.1523/JNEUROSCI.06-06-01643.1986
- Massen, J. J., Pašukonis, A., Schmidt, J. and Bugnyar, T.** (2014). Ravens notice dominance reversals among conspecifics within and outside their social group. *Nat. Commun.* **5**, 1–7. doi:10.1038/ncomms4679
- Maurer, D. and Landis, T.** (1990). Role of bone conduction in the self-perception of speech. *Folia Phoniatrica* **42**, 226–229. doi:10.1159/000266070
- McArthur, P. D.** (1986). Similarity of playback songs to self song as a determinant of response strength in song sparrows (*Melospiza melodia*). *Anim. Behav.* **34**, 199–207. doi:10.1016/0003-3472(86)90024-2
- Macmillan, N. A. and Creelman, C. D.** (2005). *Detection Theory: a User's Guide*, 2nd edn. Mahwah, NJ: Lawrence Erlbaum Associates.
- Miller, D. B.** (1979). Long-term recognition of father's song by female zebra finches. *Nature* **280**, 389. doi:10.1038/280389a0
- Mooney, R.** (2018). The song remains the same. *Trends Neurosci.* **41**, 167–170. doi:10.1016/j.tins.2018.02.006
- Nowicki, S. and Searcy, W. A.** (2014). The evolution of vocal learning. *Curr. Opin. Neurobiol.* **28**, 48–53. doi:10.1016/j.conb.2014.06.007
- Remage-Healey, L., Coleman, M. J., Oyama, R. K. and Schlinger, B.** (2010). Brain estrogens rapidly strengthen auditory encoding and guide song preference in a songbird. *Proc. Natl. Acad. Sci. USA* **107**, 3852–3857. doi:10.1073/pnas.0906572107
- Riebel, K.** (2009). Song and female mate choice in zebra finches: a review. *Adv. Study Behav.* **40**, 197–238. doi:10.1016/S0065-3454(09)40006-8
- Rossi, N. and Derégnaucourt, S.** (2020). Mechanisms of recognition in birds and social Hymenoptera: from detection to information processing. *Phil. Trans. R. Soc. B* **375**, 20190483. doi:10.1098/rstb.2019.0483
- Searcy, W. A., McArthur, P. D., Peters, S. S. and Marler, P.** (1981). Response of male song and swamp sparrows to neighbour, stranger, and self songs. *Behaviour* **77**, 152–163. doi:10.1163/156853981X00202
- Solis, M. M. and Doupe, A. J.** (1999). Contributions of tutor and bird's own song experience to neural selectivity in the songbird anterior forebrain. *J. Neurosci.* **19**, 4559–4584. doi:10.1523/JNEUROSCI.19-11-04559.1999
- Spierings, M. J. and ten Cate, C.** (2016). Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proc. Natl. Acad. Sci. USA* **113**, E3977–E3984. doi:10.1073/pnas.1600483113
- Tchernichovski, O., Nottebohm, F., Ho, C. E., Pesaran, B. and Mitra, P. P.** (2000). A procedure for an automated measurement of song similarity. *Anim. Behav.* **59**, 1167–1176. doi:10.1006/anbe.1999.1416
- Tchernichovski, O., Lints, T. J., Derégnaucourt, S., Cimenser, A. and Mitra, P. P.** (2004). Studying the song development process rationale and methods. In *Behavioral Neurobiology of Birdsong* Zeigler (ed. H. P. and P. Marler), pp. 348–363. New York, NY: New York Academy of Sciences.
- Temeless, E.** (1994). The role of neighbours in territorial systems: When are they 'dear enemies'? *Anim. Behav.* **52**, 856–859.
- Tibbetts, E. A. and Dale, J.** (2007). Individual recognition: it is good to be different. *Trends Ecol. Evol.* **22**, 529–537. doi:10.1016/j.tree.2007.09.001
- Troyer, T. and Doupe, A. J.** (2000). An associational model of birdsong sensorimotor learning. I. Efference copy and the learning of song syllables. *J. Neurophysiol.* **84**, 1204–1223. doi:10.1152/jn.2000.84.3.1204
- Volman, S. F.** (1993). Development of neural selectivity for birdsong during vocal learning. *J. Neurosci.* **13**, 4737–4747. doi:10.1523/JNEUROSCI.13-11-04737.1993
- Wang, D. W., Forstmeier, W. and Kempenaers, B.** (2017). No mutual mate choice for quality in zebra finches: time to question a widely-held assumption. *Evolution* **71**, 2661–2676. doi:10.1111/evo.13341
- Zann, R. A.** (1990). Song and call learning in wild zebra finches in south-east Australia. *Anim. Behav.* **40**, 811–828. doi:10.1016/S0003-3472(05)80982-0
- Zann, R.** (1996). *The zebra finch: a synthesis of field and laboratory studies*. Oxford: Oxford University Press.