



Supplementary Information for

Vocal learning and flexible rhythm pattern perception are linked:
evidence from songbirds

Andrew A. Rouse¹, Aniruddh D. Patel^{1,2}, and Mimi Kao^{3,4}

¹Department of Psychology, Tufts University; ²Program in Brain, Mind and Consciousness, Canadian Institute for Advanced Research; ³Department of Biology, Tufts University, ⁴Neuroscience Graduate Program, Tufts University.

*Corresponding author: Andrew A. Rouse

Email: andrew.rouse@tufts.edu

This PDF file includes:

Supplementary text
Figures S1 to S4
Table S1
Legends for Sound Files S2 to S5
SI References

Other supplementary materials for this manuscript include the following:

Sound Files S2 to S5

Supplementary Information Text

Subjects. Subjects were 16 experimentally naïve male zebra finches from the Tufts University breeding colony. Their care and treatment were approved by the Institutional Animal Care and Use Committee at Tufts University. Each male was housed individually in a custom-made sound attenuating chamber with a 12/12 h light/dark schedule. Birds had access to grit, cuttlebone, and seed (dehydrated at 44-50 °C) *ad libitum*. Prior to training birds had free access to water. At least once a week, birds had either vocal and/or physical interaction with other birds. Subjects that did not complete pre-training (“shaping”) within 30 days (see below) were removed from the study and excluded from further analysis (n=2).

Operant Chamber. Chambers were designed based on the setup described by Lim et al. (1) (see Fig. 1A). Each cage was equipped with two infrared beam-break sensors (Optek OPB815WZ) mounted on the wall that acted as peck switches: one for trial initiation (“trial switch”) and one for response (“response switch”). Each switch had an LED light that illuminated when the switch was active. A small 3W speaker (PUI Audio AS03208MS-3-R) was mounted above the center of the cage to deliver auditory stimuli (see below), and the output was calibrated for a received level of ~77.0 dB_A at 3 cm from the floor of the cage. Behavior was monitored by a closed-circuit camera.

Water was used as a reward for training as prior work has shown that zebra finches are highly motivated by water (1). A reservoir consisting of a 10-mL syringe attached to a solenoid (NResearch 161T012) was mounted outside each chamber ~20 cm above the floor, and was connected to an angled dispensing tip via Tygon tubing (.016 cm and .08 cm inner diameter in series) that slowed the flow, resulting in a 1–5 µL droplet. The water dispensing tip protruded from the center of the response switch, so that any attempt to drink would trigger the response switch. Water intake was monitored daily to ensure sufficient hydration, defined as ≥50% of their free-access consumption, or ~1 mL per day (2), averaged over three days. Birds that did not meet this intake were supplemented with free access to water to bring their consumption rate above this minimum level. Body condition was evaluated daily to check for signs of dehydration, and birds were weighed at least weekly. On average, each bird performed ~520 trials/day and consumed ≥1 mL of water per day.

Sensors, LED lights, auditory stimuli, water delivery, and house lights for each chamber were controlled by a microcontroller (Teensy 3.6) and sound card (Teensy Audio Board) attached to a custom-built printed circuit board. A Python package developed in the Gentner laboratory at UCSD (Pyoperant; original code available at <https://github.com/gentnerlab/pyoperant>) was used to control and monitor each operant box simultaneously. The Pyoperant code developed for this experiment automated stimulus presentation, reward delivery, and shaping and training phases (see below), and was controlled through a custom-made graphical interface (A.R., modified code available at <https://github.com/arouse01/pyoperant>).

Auditory operant training procedure. Using a modified go-no paradigm (go/interrupt), developed by Lim et al. (1), the bird only received a reward when it pecked the response key to one stimulus (S+ or ‘go’), but had the option to abort an ‘unrewarded’ (S- or ‘no-go’) trial by hitting the trial switch (the same switch used to initiate the trial). This paradigm attempted to mitigate a bias that can arise in the standard go/no-go task, when an incorrect response may be preferable to waiting for the opportunity to initiate another trial. In the go/interrupt paradigm a subject can demonstrate learning either by pecking the response key to the S+ stimulus or by interrupting playback of the S- stimulus.

Shaping. After acclimating to the operant chamber and prior to the discrimination tasks, each bird first learned the behavioral task of making a choice between auditory stimuli. Two unfamiliar conspecific songs were used as stimuli (~2.4 s long), one as the S+ and one as the S- stimulus. Initially, to encourage the association between pecking the trial switch and water delivery, water was dispensed on every S+ trial, either after pecking the response switch or at the end of the trial. After a bird performed at least 10 correct S+ trials, passive water delivery was discontinued. To encourage exploration of the apparatus, the lights-out punishment was not implemented until a bird successfully demonstrated reliable usage of the switches by performing at least 100 correct S+ trials.

Binned Performance Over Learning

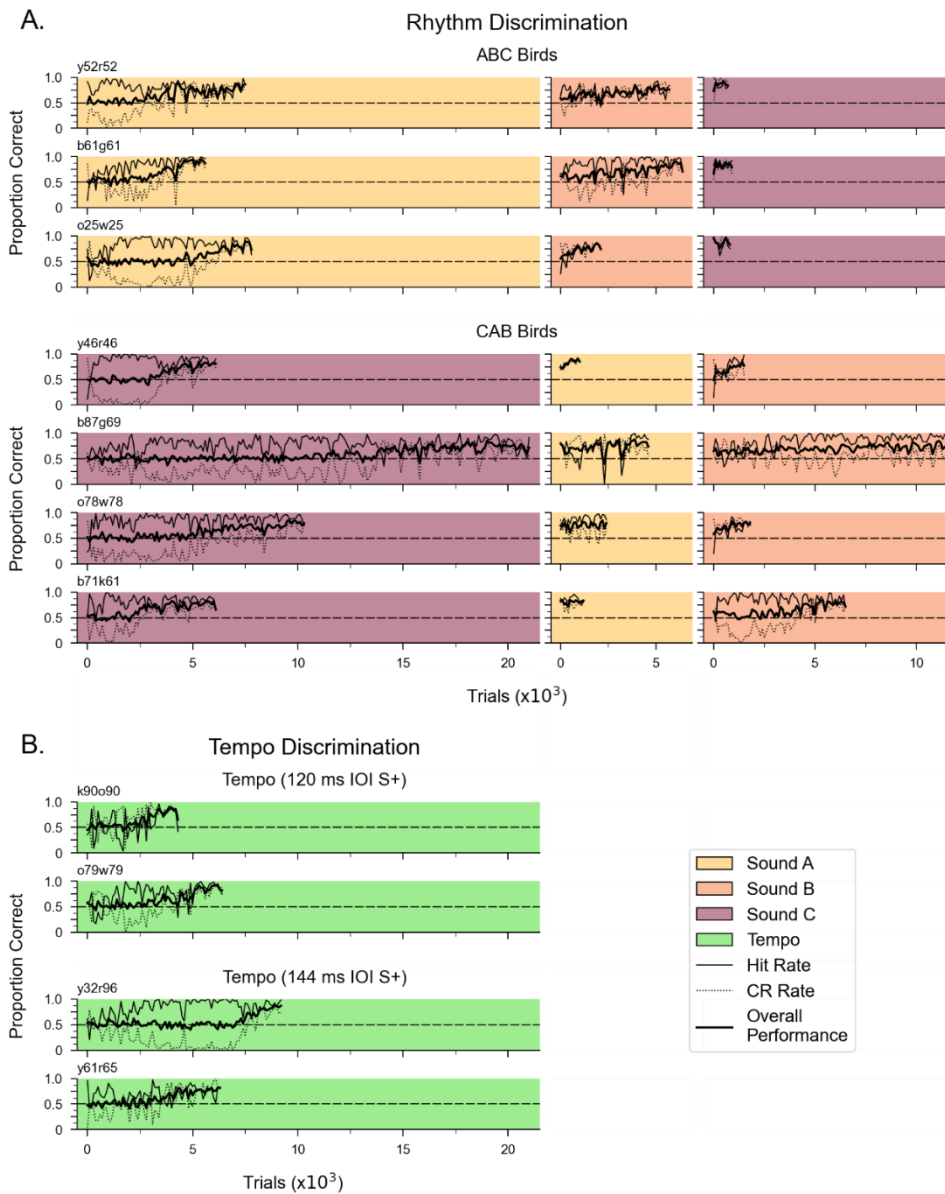


Fig. S1. Individual learning curves for discrimination based on temporal regularity or tempo. (A) Learning curves for 7 birds that successfully discriminated between isochronous and arrhythmic stimuli within 30 days per sound type. Thin black line shows proportion correct for isochronous stimuli (S+, “hit rate”); dotted gray line shows proportion correct for arrhythmic stimuli (S-, “correct rejection”); thick black line indicates overall proportion of correct responses. Dashed horizontal line indicates chance performance. Data are plotted until performance reached criterion (see *Materials and Methods*). Trials to criterion were 8563 ± 6211 trials (mean \pm SD) for the first rhythm discrimination training set. Graph color corresponds to sound type (see Fig 1D). (B) Learning curves for 4 birds trained to discriminate between sequences of sound A that differed in tempo (120ms vs. 144 ms IOI). Subjects were counterbalanced so that the rewarded (S+) stimulus was 120 ms IOI for 2 birds and 144 ms IOI for the other two. The mean number of trials to criterion was 6645 ± 1758 (SD) trials for tempo discrimination. Conventions as in (A).

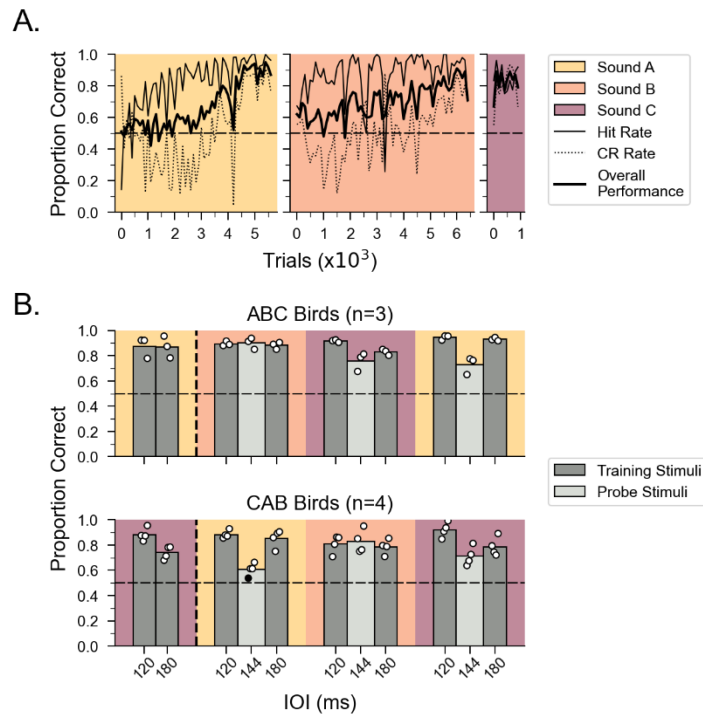


Fig. S2. Performance on rhythmic discrimination plotted as Hit Rate – False Alarm rate. Filled circles denote performance for each bird; bars represent average performance across birds in each group.

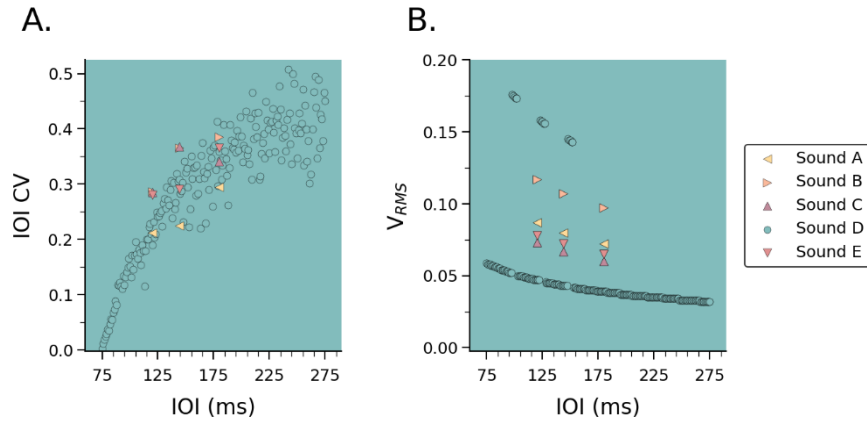


Fig. S3. Variability in the timing and amplitude of auditory stimuli. (A) Coefficient of variation (CV) of inter-onset-intervals in each arrhythmic stimulus plotted against stimulus mean IOI. As the average IOI of an arrhythmic stimulus approaches the length of the sound element used to make the sequence, variability in IOI duration approaches 0. (B) Stimulus amplitude plotted against IOI. In general, most stimuli from a single sound type have a similar amplitude across the range of rates. The exception is a subset of stimuli from sound D: 98-102 ms, 123-127 ms, and 148-152 ms, which were generated separately and were inadvertently louder than the other sound D stimuli. Circles denote sequences with sound D; other sound types are indicated by colored triangles (training stimuli at 120 and 180 ms IOI; probe stimuli at 144ms IOI). The gradual RMS decrease with increasing IOI reflects an increased proportion of silent intervals in the stimuli as IOI increases.

Individual Rule Training Performance

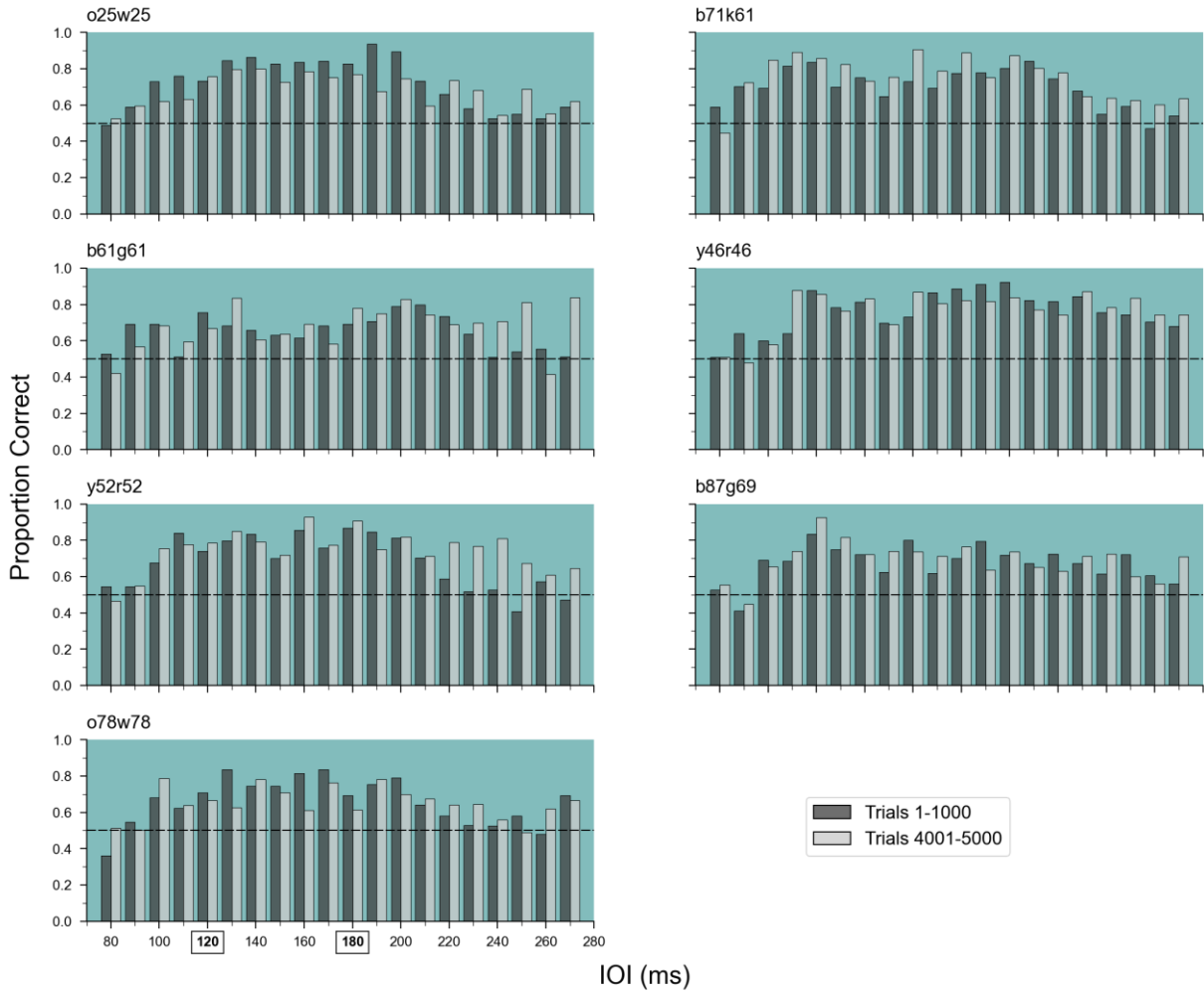


Fig. S4. Performance during rule training for each bird. Average performance for each bird for trials 1-1000 (*dark gray*) and trials 4001-5000 (*light gray*), in 10 ms bins. In the first 1000 trials all subjects showed a reduction in performance IOIs beyond 200 ms, but several birds showed improvement at these slower tempi over the course of rule training.

Table S1. Classification of possible responses during training and probe testing, and associated consequences. S+ is the rewarded stimulus; S- is the unrewarded stimulus.

Response	S+ (Regular)	S- (Irregular)
Response Switch	Water reward (Hit)	Lights off (False Alarm)
Trial Switch	Playback stops (Miss)	Playback stops (Correct Rejection)
No Response	None (Miss)	None (Correct Rejection)

Sound File S2. Corresponds to spectrogram in Fig 1C, top left (120 ms IOI, rhythmic).

Sound File S3. Corresponds to spectrogram in Fig 1C, bottom left (120 ms IOI, arrhythmic).

Sound File S4. Corresponds to spectrogram in Fig 1C, top right (180 ms IOI, rhythmic).

Sound File S5. Corresponds to spectrogram in Fig 1C, bottom right (180 ms IOI, arrhythmic).

All sound files in the study are available in the external data repository (see Data Availability).

SI References

1. Y. Lim, R. Lagoy, B. G. Shinn-Cunningham, T. J. Gardner, Transformation of temporal sequences in the zebra finch auditory system. *Elife* **5**, 1–18 (2016).
2. T. J. Cade, C. A. Tobin, A. Gold, Water economy and metabolism of two estrildine finches. *Physiol. Zool.* **38**, 9–33 (1965).