Black-capped chickadees (*Poecile atricapillus*) discriminate between naturally-ordered and scramble-ordered *chick-a-dee* calls and individual preference is related to rate of learning

Kimberley A. Campbell[ac], Marisa Hoeschele[de], Daniel Mann[c], Jenna V. Congdon[ac], Erin N. Scully[ac], Shannon K. Mischler[ac], Carolina Montenegro[ac], William D. Service[ac], & Christopher B. Sturdy[abc]

Department of Psychology[a], Neuroscience and Mental Health Institute[b], University of Alberta[c];
Acoustics Research Institute, Austrian Academy of Sciences[d]
Department of Cognitive Biology, Universität Wien[e].

Abstract

Though many forms of animal communication are not reliant on the order in which components of signals are combined to be effective, there is evidence that order does matter for some communication systems. In the light of differential responding to calls of varying note-order observed in black-capped chickadees in the field, we set out to determine whether chickadees recognize syntactically-ordered and incorrectly-ordered *chick-a-dee* calls as separate and distinct conceptual categories using both an auditory preference task and go/no-go operant conditioning paradigm. Results show that chickadees spent more time on the perch that did not produce sound (i.e., silent perch) than on either of the acoustic perches (i.e., natural and scrambled order *chick-a-dee* call playback) and visited the perch associated with naturally-ordered calls more often than the perch associated with scrambled-order calls. Birds in both the True natural- and scrambled-order call groups continued to respond according to the contingencies that they learned in Discrimination training, indicating that black-capped chickadees are capable of perceiving and acting upon the categories of natural- versus scrambled-ordered calls.

Introduction

Most forms of animal communication are thought to be non-syntactic, meaning that unique signals are employed to serve different purposes in various situations (Nowak, Plotkin, & Jansen, 2000). For example, vervet monkeys (*Cercopithecus aethiops*) and Japanese tits (*Parus minor*) produce distinct alarm vocalizations in response to avian versus terrestrial predators (Seyfarth, Cheney, & Marler, 1980; Suzuki & Ueda, 2013). In contrast, syntactic communication, as is seen in human language, involves individual components of a signal having their own meaning and the combination of those components producing different and varied meanings (Nowak, Plotkin, & Jansen, 2000). A major hurdle of understanding the origins of human language is determining how human language could have evolved from non-human animal communication through the process of natural selection (Suzuki, Wheatcroft, & Griesser, 2018). This endeavour is made more difficult by the fact that our closest evolutionary relatives, the great apes, do not possess syntactic communication systems (Fitch, 2010). Researchers have had to look more distantly to find species that share some of the traits involved in human language.

Many nonhuman animal species vocal communication putatively follow rules of syntax in their vocalizations. For example, non-human primates, humpback whales, and songbirds have all been shown to produce vocalizations that follow syntactical rules to different degrees (see Zuberbühler, 2019 for an extensive recent discussion of this topic). Specifically, Diana monkeys (*Cercopithecus diana*; Candiotti, Zuberbühler & Lemasson, 2012) alter the ordering of their vocalizations in a contextually-dependent manner, with different combinations being used during either positive or negative social interactions. Similarly, Campbell's monkey (*Cercopithecus campbelli*) alarm calls are delivered with different syllable order depending on the urgency of the situation (Ouattara, Lemasson, & Zuberbühler, 2009). In addition to nonhuman primates, several bird species are also sensitive to syntactical rules contained in their vocalizations. The Japanese great tit (*Parus minor*) have been shown to behave differently when they are exposed to different note types or different orders of call notes (Suzuki, Wheatcroft, & Griesser, 2016) and they are not the only members of the Parid family that follow syntactical rules in the perception or production of their species-specific vocalizations. Hailman and Ficken (1986) noted that the *chick-a-dee* call of black-capped chickadees (*Poecile atricapillus*), a small species of North American songbird, possesses a computable syntax not unlike that of human language. *Chick-a-dee* calls consist of four main note types that are amalgamated following a fixed order (A → B → C → D) though individual note types can be repeated or omitted, yielding a theoretically unlimited repertoire of producible call types (Hailman, Ficken, & Ficken, 1985). Chickadees can use their *chick-a-dee* call as a mobbing call, warning and recruiting both con- and heterospecifics of the presence of predators (Hurd, 1996). The composition of the call changes depending on how dangerous the predator is perceived to be: *chick-a-dee* calls with more D notes are produced to higher threat owls and hawks compared to low threat predators (Templeton, Green, & Davis, 2005). In a mobbing situation, the same vocalization type can thus convey different meanings depending on the number and type of notes produced. In addition to threat-level cues, the *chick-a-dee* call also contains information about species identity (Bloomfield & Sturdy, 2008; Bloomfield et al., 2005), flock membership (Mammen & Nowicki, 1981; Nowicki, 1989), and individual identity (Charrier, Bloomfield, & Sturdy, 2004) and birds in the field have been shown to be sensitive to the order of notes within calls, showing reduced responding to playback of calls with reversed syntax (Charrier & Sturdy, 2005).

In the light of differential responding to calls of varying composition observed in black-capped chickadees in the field, we set out to determine whether black-capped chickadees recognize syntactically-ordered and incorrectly-ordered *chick-a-dee* calls as separate and distinct conceptual categories. Specifically, we investigated the order of notes in the call in two stages. In the first stage of the current study, we designed a behavioural choice preference task to investigate if chickadees demonstrated a preference for correctly-ordered *chick-a-dee* calls over incorrectly-ordered (i.e., scrambled) *chick-a-dee* calls. In the second stage of the experiment, we designed an operant conditioning discrimination task to determine if black-capped chickadees perceive natural-ordered *chick-a-dee* calls as a separate perceptual category from scrambled-ordered *chick-a-dee* calls. We also tested whether there was a relationship between preference in stage one and performance in stage two.

The results of this study will strengthen our knowledge of the perception of structure and meaning of black-capped chickadee *chick-a-dee* calls and, more broadly, will add to our understanding of how syntactic communication systems may differ from non-syntactic systems in important ways. First, in study one, the choice-based preference task is designed to allow us to test which call order a bird chooses to hear, and not only whether birds respond with increased vocalizations or approach to particular vocalizations as in previous playback studies. Second, in experiment two, the operant discrimination study is designed to allow us to test for the particular perceptual and cognitive mechanisms that birds are tapping into to perceive natural versus altered-syntax vocalizations. Additionally, the interaction of both order preference and performance in the discrimination task will allow us to determine if birds are able to learn the discrimination better when reinforced for preferred versus not-preferred categories. Taken together, the results of these studies will expand our understanding of syntax usage in black-capped chickadees specifically and in nonhuman animals more generally.

Methods

Subjects

Twenty black-capped chickadees (9 males, 11 females) were tested between July 17, 2017 and March 9, 2018. Birds were captured in Edmonton, Alberta, Canada (North Saskatchewan River Valley, 53.53°N, 113.53°W; Mill Creek Ravine, 53.52°N, 113.47°W) between February 5, 2015 and February 6, 2017. Birds were determined to be at least one year of age at time of capture by examining the shape and colour of outer tail retrices (Pyle, 1997) and sex was determined by DNA analysis of blood samples (Griffiths, Double, Orr, & Dawson, 1998). No birds had experience with the experimental procedures or stimuli.

Housing

Birds were individually housed in 30 × 40 × 40 cm cages (Rolf C. Hagen, Inc. Montreal, QB) in colony rooms where they had visual and auditory, but not physical, contact with conspecifics when not in experimental apparatus. Colony rooms were maintained on a light:dark cycle that matched the natural light cycle of Edmonton, Alberta, Canada. Birds had *ad libitum* access to food (Mazuri Small Bird Maintenance Diet; Mazuri, St Louis, MO, U.S.A.), water, grit, and cuttlebone. Birds received additional supplementation in the form of a superworm (*Zophobas morio*) and vitamins added to water (Prime vitamin supplement; Hagen, Inc.) three times a week, a mixture of eggs and spinach or parsley twice a week, and three to five sunflower seeds daily.

Apparatus

Choice preference task. Birds were individually tested in a 67 × 116 × 116 cm testing space within a 117 × 120 × 200 cm sound-attenuating chamber (Industrial Acoustics Company, Bronx, NY). The testing space contained three 1.75 cm diameter, 10 cm long perches monitored by an infrared beam. The back and side walls each had a single perch and each perch was mounted in front of a Fostex FE108E S full-range speaker (Fostex Corp., Tokyo, Japan; frequency response range 80-18000 Hz) 100 cm above the floor of the testing space (see Figure 1 for diagram). Landing on one of the two acoustic perches would break the infrared beam, initiating stimulus playback and recording the visit. Landing on the silent perch would still break an infrared beam and record the visit, however no auditory playback occurred. Stimulus playback and response monitoring was controlled by a single-board computer (Palya & Walter, 2001) and a personal computer. Stimuli stored on a CD were played through a Cambridge Azur 640A Integrated Amplifier (Cambridge Audio, London, U.K.) to the speakers within the testing space. Birds had *ad libitum* access to food and water during testing sessions.

Operant conditioning task. Birds were individually housed in modified home cages (30 × 40 × 40 cm) located within ventilated sound-attenuating chambers during the instrumental learning task. Chambers were illuminated by 9W full-spectrum fluorescent bulbs on a light:dark cycle that matched the natural light cycle of Edmonton, Alberta, Canada. Cages contained three perches, a grit cup, a water bottle (water vitamin supplemented three days a week), and cuttlebone. Birds received one superworm twice daily. Mazuri food was accessible only after a correct response, as a reward. An 11 × 16 cm opening on one side of the cage allowed the birds to access a motorized feeder. Infrared beams in the perch closest to the feeder and in the feeder itself tracked the position of the bird during testing. A single-board computer tracked responses and set up trials in connection with a personal computer which stored and played the acoustic stimuli. Stimuli were amplified by an NAD310 Integrated Amplifier (NAD Electronics, London, U.K.) or a Cambridge A300 or 640A Integrated Amplifier (Cambridge Audio, London, U.K.) before playing through a Fostex FE108 Σ full range speaker (Fostex Corp., Tokyo, Japan; frequency response range: 80-18 000 Hz) located beside the feeder. See Sturdy and Weisman (2006) for a detailed description of the experimental setup.

Acoustic Stimuli

Four hundred and seventy-eight *chick-a-dee* calls were collected from four black-capped chickadees (two males: 290 calls, two females: 188 calls). Calls were recorded with an AKG C 1000S (AKG Acoustics, Vienna, Austria) microphone feeding into a Marantz PMD670 (Marantz America, Mahwah, NJ) digital sound recorder using a 16-bit, 44 100 Hz sampling rate in a 1.7 m × 0.84 m × 0.58 m sound-attenuating chamber (Industrial Acoustics Company, Bronx, NY). Birds remained in their home cages during recording and were permitted to acclimatize overnight in the chamber before recordings were obtained. Subjects had no experience with individuals that provided call stimuli. Note composition was determined by a single individual using visual inspection of spectrograms in SIGNAL (version 5.05.02, Engineering Design, 2013) and using Ficken, Ficken, and Witkin (1978) as a reference. Identified notes were extracted and saved individually using SIGNAL to create a pool of available notes.

One hundred seventy stimuli (85 natural order, 85 scrambled order) four to ten notes in length ($X \pm \mathrm{SD} = 6.99 \pm 2.04$ notes) were constructed by randomly selecting notes from the pool without replacement (i.e., a given note was only used once) such that each created stimulus contained at least one A, B, C, and D note. This length was chosen based on the composition of the recorded calls. After notes were selected for a stimulus, two paired stimuli were created: 1) a natural-ordered stimulus in which notes were ordered A → B → C → D, and 2) a scrambled-ordered stimulus where notes did not follow the natural order (e.g., D → B → A → C). Calls were assembled with 10 ms of silence between adjacent notes and were bandpass filtered using GoldWave 6.19 (GoldWave, Inc., St John's, NL, Canada) outside of the frequency range of *chick-a-dee* calls (2000-5000 Hz) to remove any noise. Using SIGNAL, 5 ms of silence was added to the beginning and end of each call. Each file was also tapered to remove transients and then amplitude was equalized using GoldWave.

Procedure

Choice preference task. The 170 call stimuli were pseudo-randomly divided into two stimulus sets (Set A: 84 stimuli, Set B: 86 stimuli) such that each set had the same number of calls of each note length (e.g., half of four-note-long calls assigned to Set A, half to Set B). Pairs of natural and scrambled calls that were constructed from the same set of notes were kept together such that if the natural call was assigned to Set A, so was the equivalent scrambled call. Within a test session, birds only heard stimuli from one set. The order in which stimulus sets were presented was randomized between birds such that half of subjects heard Set A first, and half heard Set B first. Once birds met criteria (see below) for the first stimulus set, they were then tested on the second stimulus set.

Within a test session, natural-ordered calls were assigned to one perch, scrambled-ordered calls to another, and no acoustic stimulus for the final perch. There were six unique ways that the three stimulus types could be assigned to the left, back, and right perches. We randomly assigned perch configurations for the first stimulus set presented, then randomly chose one of the two possible remaining configurations for the second set that ensured no stimulus category was associated with the same perch for both sets.

Test sessions lasted for two hours and occurred between 09:00 and 16:00 hours. Birds only received a single test session each day. For testing, birds were transported from the colony room in their home cage, removed from the cage, and released into the testing area. After the test session ended, the bird was caught, returned to their home cage, and transported back to the colony room.

If the perch was an acoustic perch (i.e., natural or scrambled order call), breaking the beam triggered playing a single call stimulus in its entirety. A second stimulus would not play unless the bird left and returned to the perch (i.e., calls did not play simultaneously). Stimuli were randomly selected without replacement from the pool of stimuli in the appropriate stimulus category for the set. Birds were tested on the same stimulus set until they had heard each acoustic stimulus at least five times or had completed a maximum of five test sessions (e.g., 10 h total). After meeting these criteria, birds moved on to testing with the second stimulus set after at least one day or rest. After completing the choice preference task for both stimulus sets, birds were tested on the operant conditioning task.

Operant conditioning task.

Pretraining. First, birds underwent basic training (i.e., shaping) to ensure that they were able to obtain food from the experimental apparatus, then they began Pretraining. During basic training, one male bird was removed from the experiment and returned to the colony room after failing to feed successfully from the apparatus. Pretraining was included to ensure that birds responded non-differentially to all stimuli that would be differentially rewarded during later stages and to ensure that they responded at a sufficiently high rate. To accomplish this, birds were presented with and non-differentially rewarded for responding to all of the stimuli that would be used throughout the experiment. Birds would initiate a trial by landing on the request perch, breaking the infrared beam and triggering the playback of a stimulus. During playback, a stimulus was randomly selected from the 170 stimuli in a bin without replacement until all stimuli had been heard. In order to ensure the entirety of a stimulus was heard, birds were required to remain on the perch for the duration of each stimulus (1470-2377 ms) after a trial was initiated. Trials were considered interrupted if the bird left the perch before playback was completed. This triggered a 30-s timeout in which the houselight was turned off and new trials could not be initiated. If the bird flew into the feeder within 1 s of the end of playback, they were rewarded with 1 s access to food followed by a 30-s intertrial interval in which the houselight remained on and new trials could not be initiated. If the bird remained on the perch for more than 1 s after the end of playback, a 60-s intertrial interval was initiated in which the houselight remained on and new trials could not be initiated. This interval ended if the bird subsequently left the perch. Birds remained on Pretraining until they responded to 60% or more of trials across six 170-trial bins, displayed a less than 3% difference in responding across four 170-trial bins to what would be rewarded and unrewarded stimuli in Discrimination training (see below), and displayed a less than 3% difference in responding across four 170-trial bins to categories of Transfer testing stimuli (i.e., transfer natural stimuli and transfer scrambled stimuli). During this stage, a female subject died due to human error in loading a program.

Discrimination training. After meeting criteria in pretraining, birds began Discrimination training. The method of stimulus presentation remained the same, however, only 80 of the total 170 stimuli were presented and half (40) of these calls were now unrewarded (i.e., responding to these stimuli now resulted in a 30-s intertrial interval in which the houselight was off and a new trial could not be initiated). Responding to the remaining 40 calls was rewarded as in Pretraining with 1 s access to food.

Birds were randomly assigned to either a True category discrimination group ($N = 12$) or a Pseudo category discrimination group ($N = 5$). The True category group consisted of two subgroups that were rewarded for either responding to natural-ordered calls ($N = 6$) or scrambled-ordered calls ($N = 6$). The Pseudo category discrimination group also consisted of two subgroups (Pseudo 1: $n = 2$; 1 male, 1 female; Pseudo 2: $n = 3$; 1 male, 2 females) that were rewarded for responding to 40 randomly-selected calls (20 natural-ordered, 20 scrambled-ordered) and unrewarded for responding to the remaining 40 calls (20 natural-ordered, 20 scrambled-ordered).

Birds remained on Discrimination training until they completed six 80-trial blocks with a discrimination ratio (DR) of at least 0.80 with the last two of these blocks being consecutive. DR was calculated by dividing the mean percentage of response to all rewarded stimuli by the mean percentage of response to rewarded stimuli plus the mean percentage of response to unrewarded stimuli, then multiplying by 100. With this calculation, a DR of 0.50 indicates equal responding to rewarded and unrewarded stimuli, while a DR of 1.00 indicates responding to only the rewarded stimuli (i.e., perfect discrimination). During this stage a male in the Pseudo 2 group died of natural causes.

Discrimination 85 training. Nearly identical to Discrimination training, Discrimination 85 training differed only in that the rewarded stimuli were reinforced only 85% of the time. This meant that when a stimulus from the rewarded category was played, on 15% of trials entering the feeder resulted in a 30-s intertrial interval in which the houselight remained on without access to food. This stage served to expose birds to trials in which responses to stimuli were neither rewarded nor punished, as would be encountered in Transfer trials. As with the Discrimination stage, birds continued on Discrimination training until they completed six 80-trial blocks with a DR of at least 0.80, where the last two of these blocks needed to be consecutive.

Transfer testing. In order to determine if birds respond to novel calls following the rules learned in Discrimination training, Transfer testing was conducted. After meeting criterion in Discrimination 85 training, birds began Transfer testing. The stimuli and reward-contingencies introduced in Discrimination 85 training remained the same. Birds were also introduced to an additional 90 calls (45 natural order, 45 scrambled order) that had previously been non-differentially reinforced (i.e., had been played during pretraining), but were not heard during Discrimination training. Responding to these novel transfer stimuli was neither rewarded nor punished, in that a response resulted in a 30-s intertrial interval in which the houselight remained on without access to food being provided. Within a 650-trial bin, the 80 discrimination stimuli were played seven times each while the new transfer stimuli were only played once. Upon completion of three bins, the experiment was completed and birds were returned to the colony room.

Response measures

Choice preference task. During this task, the absolute number of visits to each perch and the amount of time spent on each perch was recorded. However, individual birds took varying numbers of test sessions to meet criteria. To account for this, we calculated the average number of visits to each perch and the average amount of time spent on each perch by dividing the total measure by the number of 2-h sessions it took to meet criteria. We tested whether these values differed between set A and set B using paired-samples *t*-tests for each measure (e.g., average number of visits and average time) before combining the two data sets.

To quantify preference between the two types of acoustic stimuli, we calculated two preference scores, one using number of visits and the other using time on the perch, for each bird as follows:

$$preference\ score_{visits} = \frac{(visits\ to\ natural\ perch)}{(visits\ to\ natural\ perch) + (visits\ to\ scrambled\ perch)}$$

$$preference\ score_{time} = \frac{(time\ on\ natural\ perch)}{(time\ on\ natural\ perch) + (time\ on\ scrambled\ perch)}$$

Here, a preference score between 0.5 and 1 indicates a preference for natural ordered calls, a score of 0.5 indicates no preference, and a score between 0.5 and 0 indicates a preference for scrambled order calls.

Operant conditioning task. To analyze patterns of responding, we calculated the proportion of responding for each stimulus type by dividing the number of trials where the bird went into the feeder by the total number of trials in which the bird either went into the feeder or left the request perch after hearing the full call stimulus (i.e., all non-interrupted trials). A proportion of responding of 1 indicates responding to all stimuli in a category.

Statistical Analyses

For the choice preference task, we conducted a repeated measures ANOVA to evaluate differences in the average amount of time spent on each of the three perches. We conducted a similar repeated measures ANOVA on the average number of visits to each of the three perches.

To determine if the two True category groups differed from one another in speed of acquisition, we conducted an independent-samples *t*-test on the number of 80-trial blocks individuals took to reach criterion during Discrimination training. In the same fashion, we compared speed of acquisition between the two Pseudo category groups. In order to compare rate of learning between the True and Pseudo category groups, we conducted an independent-samples *t*-test on the number of 80-trial blocks individuals took to reach criterion during Discrimination training.

We then conducted a repeated measures ANOVA on the proportion of responding to each stimulus type (i.e., discrimination natural calls, discrimination scrambled calls, transfer natural calls, and transfer scrambled calls) during the three 650-trial blocks of Transfer testing.

Finally, we conducted two stepwise multiple regression analyses to determine if bins to criteria could be used to predict the strength of preference, as measured by preference score, for both time spent on and number of visits to the silent, natural, and scrambled perches.

All statistics were conducted in IBM SPSS Statistics v.22 (IBM Corp., Armonk, NY, U.S.A.)

Ethical Note

During the choice preference task, birds remained in the testing apparatus for only two hours a day and had free access to food and water. During the instrumental learning task, birds remained in the testing apparatus to minimize stress and discomfort caused by transport and handling. Following the experiments, birds were returned to the colony room for use in future experiments. Birds were closely monitored throughout the experiments. One bird died during pretraining due to a program error and one bird died during Discrimination training due to natural causes. All other birds remained healthy during the experiment. These procedures were conducted in accordance with the Canadian Council on Animal Care (CCAC) Guidelines and Policies with approval from the Animal Care and Use Committee for Biosciences for the University of Alberta (AUP 108), which is consistent with the Animal Care Committee Guidelines for the Use of Animals in Research. Birds were captured and research was conducted under an Environment Canada Canadian Wildlife Service Scientific permit (#13-AB-SC004), Alberta Fish and Wildlife Capture and Research permits (#56076 and #56077), and City of Edmonton Parks Permit.

Results

Response to Perches

Figure 2 shows the average amount of time spent on each of the three perches and Figure 3 shows the average number of visits to each of the three perches. We conducted a repeated measures ANOVA where the average amount of time on each of the three perches (e.g., natural order, scrambled order, silence) was the within-subjects factor and sex was the between-subjects factor. Mauchly's test indicated that the assumption of sphericity had not been violated, χ2(2) = 3.094, p = .213. There was a significant main effect of perch ($F_{2,30}$= 10.095, *p* < 0.001 , η2p = .402). A Bonferroni *post hoc* test revealed that birds spent significantly more time on the silent perch than on either acoustic perch (natural: *p* < .001, d = 0.698; scrambled: *p* < .001 , d = 1.649). There was no significant main effect of sex (p = .624, η2p = .016). A similar repeated measures ANOVA on average number of visits to each of the three perches (Mauchly's: χ2(2) = 9.276, p = .010; Greenhouse-Geiser correction applied) also revealed a significant main effect of perch ($F_{1.35,20.21}$= 4.158, *p* = .044 , η2p = .217), as well as a significant effect of sex (F1, 15 = 4.888, p = .043, η2p = .246). Here, birds were revealed to visit the natural order perch significantly more often than the scrambled order perch (*p* = .019 , d = 0.518). In both cases, there was no significant interaction (time: p = .528, η2p = .042; visits: *p* = .385, η2p = .062).

Trials to Criterion

The average ± SD number of trials required for each group to complete Discrimination training are as follows: Natural S+ group = 32.17 ± 14.05; Scrambled S+ group = 50.17 ± 35.43;

Pseudo group 1 = 23.67 ± 23.03; and Pseudo group 2 = 77.33 ± 23.97. Independent-samples *t*-tests on the number of 80-trial blocks individuals took to reach criterion during Discrimination training revealed no significant difference between either the two True category groups (i.e., Natural S+ and Scrambled S+ groups; Levene's test was significant (F = 7.112, p = .024), so equal variances not assumed: $t_{6.534} = -1.157$, $p = .288$, d = 0.668) or the two Pseudo category groups (equal variance assumed (F = 1.774, p = .275): $t_3 = -2.145$, $p = .121$, d = 2.096). There was also no difference in speed of acquisition between the True and Pseudo category groups (equal variance assumed (F < .001, p = .984): $t_{15} = -1.307$, $p = .211$, d = 0.684).

Transfer testing

True category groups. We conducted a repeated measures ANOVA on the proportion of responding during the three 650-trial bins of Transfer testing to evaluate if individuals in the True category groups continued to respond to the reward-contingencies learned in Discrimination training when presented with novel stimuli in Transfer testing. Stimulus type (e.g., discrimination natural order calls, discrimination scrambled order calls, transfer natural order calls, transfer scrambled order calls) was the within-subjects factor and both category group (e.g., natural S+ group, scrambled S+ group) and sex as between-subjects factors. Mauchly's test indicated that the assumption of sphericity had not been violated, $\chi2(5) = 7.117$, p = .216. There was a significant stimulus type × group interaction ($F_{3,24}$= 129.70, $p < .001$, η2p = .942). All other main effects and interactions were not significant (stimulus type: F3,24 = 0.053, p = .984, η2p = .007; stimulus type × sex: F3,24 = 0.411, p = .739, η2p = .050; stimulus type × group × sex: F3,24 = 1.687, p = .196, η2p = .174).

We conducted *post hoc* pairwise comparisons using Bonferroni corrections on the proportion of responding to each stimulus type for each True group separately. Birds in the natural order S+ group responded significantly more to rewarded natural order calls than to unrewarded scrambled order calls (adjusted $p < .001$, 95% Confidence Interval (CI) = -0.743 – -0.546) and to transfer scrambled order calls (adjusted $p = .001$, CI = -0.857 – -0.352). They also responded significantly more to transfer natural order calls than to both unrewarded scrambled order calls (adjusted $p = .001$, CI = 0.330 – 0.818) and transfer scrambled order calls (adjusted $p = .001$, CI = -0.756 – -0.313). Similarly, birds in the scrambled order S+ group responded significantly more to rewarded scrambled order calls than to both unrewarded natural order calls (adjusted $p = .004$, CI = 0.268 – 0.990) and transfer natural order calls (adjusted $p = .002$, CI = -0.871 – -0.321). They also responded significantly more to transfer scrambled order calls than to both unrewarded natural order calls (adjusted $p = .007$, CI = 0.206 – 0.909) and transfer natural order calls (adjusted $p = .001$, CI = 0.296 – 0.752). All other pairwise comparisons were not significant (adjusted $p > .472$).

Four independent-samples *t*-tests were conducted to compare the proportion of responding to each stimulus type (e.g., discrimination natural order calls, discrimination scrambled order calls, transfer natural order calls, transfer scrambled order calls) between the two True groups (e.g., natural order S+ group, scrambled order S+ group). The adjusted *p*-value for significance became p = .05/4 = .013 with Bonferroni corrections for four comparisons. Levene's test was not significant for any comparison (p > .135), so equal variances were assumed. There was a significant difference in responding between groups such that birds rewarded for responding to natural-ordered stimuli (e.g., natural order S+ group) responded more to natural-ordered stimuli (both from discrimination and from transfer) than did birds that were rewarded for responding to scrambled order stimuli (discrimination: $t_{10} = -5.003$, $p = .001$, d = 2.888, CI = -0.779 – -0.299; transfer: $t_{10} = -4.489$, $p = .001$, d = 2.592, CI = -0.652 – -0.219) and birds rewarded for responding to scrambled order stimuli (e.g., scrambled order S+ group) responded more to scrambled order stimuli (both from discrimination and from transfer) than did birds that were rewarded for responding to natural-ordered stimuli (discrimination: t_{10} = 24.005, *p* < .001, d = 13.861, CI = 0.666 – 0.802; transfer: t_{10} = 10.835, *p* < .001, d = 6.255, CI = 0.495 – 0.751; see Figure 4).

Pseudo category groups. In the same manner as was done for the True category groups, we also conducted a repeated measures ANOVA on the proportion of responding during the three 650-trial bins of Transfer testing to evaluate how individuals in the Pseudo category groups responded to the novel stimuli in Transfer testing. Mauchly's test could not be conducted due to insufficient residual degrees of freedom. Again, there was a significant stimulus type × group interaction ($F_{3,3}$= 95.486, *p* = .002, η2p = .990). There was also a significant main effect of both group and sex (group: $F_{1,1}$= 697.173, *p* = .024, η2p = .999; sex: $F_{1,1}$= 244.052, *p* = .041, η2p = .996). All other main effects and interactions were not significant (stimulus type: F3,3 = 1.588, p = .357, η2p = .614; stimulus type × sex: F3,3 = 1.637, p = .348, η2p = .621; stimulus type × group × sex: F3,3 = 2.094, p = .280, η2p = .677).

To examine the main effects, we conducted *post hoc* Bonferroni comparisons. Birds in the Pseudo 1 group responded more overall than did birds in the Pseudo 2 group (*p* = .024). Similarly, males responded more overall than females (*p* = .041). We conducted *post hoc* pairwise comparisons using Bonferroni corrections on the proportion of responding to each stimulus type for each Pseudo category group separately. Levene's test was not significant for any comparison (p > .103), so equal variances were assumed. Birds in the Pseudo 1 group responded significantly more to rewarded discrimination calls than to transfer scrambled order calls (adjusted *p* = .036, CI = 0.037 – 0.233) and more to transfer natural order calls than unrewarded discrimination calls (adjusted *p* = .026, CI = -0.908 – -0.289). All other pairwise comparisons were not significant (adjusted *p* > .053). See Figure 5 for summary.

Four independent-samples *t*-tests were conducted to compare the proportion of responding to each stimulus type between the two Pseudo category groups. The adjusted *p*-value for significance became *p* = .05/4 = .013 with Bonferroni corrections for four comparisons. There was a significant difference in responding between groups such that birds in the Pseudo 1 group responded more to Pseudo 1 S+ stimuli (i.e., Pseudo 1 rewarded discrimination stimuli) than did birds in the Pseudo 2 group (t_3 = 10.254, *p* = .002, d = 10.800, CI = 0.461 – 0.877). Similarly, birds in the Pseudo 2 group responded more Pseudo 2 S+ stimuli (i.e., Pseudo 2 rewarded discrimination stimuli) than did birds in the Pseudo 1 group (t_3 = -18.409, *p* < .001, d = 15.423, CI = -0.812 – -0.573). There were no significant differences in responding between groups to transfer natural order stimuli (t_3 = 3.137, *p* = .052, d = 3.224) or to transfer scrambled order stimuli (t_3 = 3.392, *p* = .043, d = 3.575).

Acoustic preference and performance. We conducted stepwise multiple regression analyses to determine if a measure of learning speed, here bins to criteria, could be used to predict the strength of preference, as measured by preference score, for both time spent on and number of visits to each of the three perches. In both instances, the relationship was linear and the data did not violate assumptions of homoscedasticity or independence of errors. For the preference score calculated using the amount of time spent on the three perches, the regression equation was not significant (R^2 = 0.075, R^2_{adj} = -0.131, $F_{2,9}$ = 0.364, *p* = .704). However, for the preference score calculated using the number of visits to each of the three perches, the regression was significant (R^2 = 0.486, R^2_{adj} = 0.372, $F_{2,9}$ = 4.258, *p* = .050). The number of bins required to reach criterion significantly predicted the number of visits to the perch (β = 0.711, *p* = .017) such that birds that learned the discrimination in more trials demonstrated a stronger preference toward what would be the rewarded stimulus type (see Figure 6).

Discussion

In a series of two experiments, both a playback experiment where birds could request the playback stimuli, and an operant conditioning experiment, we examined the preference for and perceptual mechanisms underlying, naturally-ordered and scrambled chick-a-dee calls in black-capped chickadees. In this manner, we could evaluate the relationship between individual preference and discrimination performance as it pertains to natural- and unnaturally-ordered conspecific calls. We first conducted a choice preference task to evaluate individual black-capped chickadees' responses to our two types of calls without the influence of differential appetitive (i.e., food) reward. We then conducted an instrumental discrimination task in which responding to calls was differentially-rewarded (i.e., food-rewarded). This allowed us to examine how individual variation in stimulus preference could influence discrimination performance in natural and unnatural signals.

Individual Preference

We found that black-capped chickadees spent more time on the perch that did not produce sound (i.e., silent perch) than they spent on either of the perches that would produce sound (i.e., natural and scrambled order *chick-a-dee* call playback). This finding is consistent with observations by Hahn and colleagues (2017) who were investigating preferences between dominant and subordinate *fee-bee* songs. Hahn et al. (2017) observed that birds spent more time on the silent perch compared to either of the acoustic perches. However, in addition to preferring the silent perch to the acoustic perches, there was also a preference among the acoustic perches, with chickadees in the current study visiting the perch associated with naturally-ordered calls more often than the perch associated with scrambled-order calls. Our results further indicate that when birds selected the silent perch, they remained on the silent perch, whereas when birds selected the natural-order perch, the birds would visit the perch, then leave and return again , actively choosing the perch, and thus opting for more song playback of the preferred acoustic perch.

Perceptual Categorization

In this experiment, birds in the True category groups were differentially-reinforced for responding to one category of stimuli but not the other (e.g., rewarded for responding to natural-ordered stimuli but not scrambled-order stimuli). Whereas birds in the Pseudo category groups were differentially-reinforced for responding to randomly-selected groups of vocalizations that did not form categories (i.e., rewarded for responding to randomly-selected natural- and randomly-selected scrambled-order calls). This distinction allows us to evaluate whether black-capped chickadees preferentially make use of open-ended categorization when learning this discrimination. Black-capped chickadees have been shown to use open-ended categorization to discriminate between male- and female-produced *fee-bee* songs (Hahn et al., 2015), and con- and heterospecific *chick-a-dee* calls (Bloomfield, Farrell, & Sturdy, 2008; Bloomfield, Sturdy, Phillmore, & Weisman, 2003). By the nature of the training contingencies, only True category groups would be capable of learning the two categories (i.e., natural and scrambled) from Discrimination training. Evidence in support of True category groups using categorization can come from two sources. First, if we assume that learning a category (possible mechanism for the True groups) is easier than memorizing each individual call (as required with Pseudo category groups), we would expect True category groups to learn the discrimination at a faster rate than Pseudo category groups. By comparing the number of bins required to meet Discrimination training criteria between True and Pseudo category groups, we found no such difference in rates of acquisition. This could imply that both True and Pseudo category groups were relying on rote memorization to learn the discrimination. Second, only birds that learned a category could transfer that knowledge, and hence their pattern of responding, to new stimuli. If birds had simply memorized the Discrimination training stimuli, they would respond non-differentially to the testing stimuli. Our results indicate that birds in both the True natural- and scrambled-order call groups continued to respond according to the contingencies that they learned in Discrimination training during Transfer testing. Additionally, birds in the Pseudo category groups showed non-differential responding to the Transfer testing stimuli, as would be expected since they did not learn True categories in Discrimination training. These results taken together indicate that black-capped chickadees are capable both of learning and using the categories of natural- versus scrambled-ordered calls, while also being able to use rote memorization of calls to solve our operant discrimination task. The ability to use both open-ended categorization and rote memorization has significant implications beyond our study here that examined the preference and perception of naturally-ordered and scrambled-ordered calls. For instance, these seemingly complimentary cognitive mechanisms could be at play during critical tasks such as seed caching and recovery as well as during social interactions among birds. Seed caching could be aided by open-ended categorization to initially guide a bird to the general cache location with similar physical characteristics, while memorization could then hone the search for the particular cache location. In a similar way, open-ended categorization could help a bird determine if a call was produced by a flockmate or non-flockmate, and hence guide behavioural decisions at this level (e.g., defend territory against a nonflockmate or not), while rote memorization could then aid a bird in determining which particular bird from their flock emitted the call and thus guide a more fine-grained behavioural response (e.g., informing a bird how to behave to a particular bird depending on the dominance status of the sender and receiver).

Acoustic Preference and Performance

In this study, we presented black-capped chickadees with two types of *chick-a-dee* calls that had been constructed by artificially assembling individual notes into either natural- or scrambled-order calls. In nature, *chick-a-dee* calls are almost exclusively produced with notes in a fixed order of A through D notes (A → B → C → D; Hailman, Ficken, & Ficken, 1985). Calls that deviate from this set syntax are rare (personal observation), suggesting that there may be information conveyed by the order itself, rather than subtleties in the structure of individual notes (Hailman & Ficken, 1986). This quality can be seen in that some notes produced earlier in a call are thought to contain redundant information on notes that come later in the call (Freeberg, Lucas, & Clucas, 2003). For example, the structure of an A note produced early in the call will change depending on how many other A notes will follow it. In our experiment, we observed a preference for natural-ordered calls over scrambled-ordered calls in that birds visited the perch associated with natural-ordered calls more compared to the perches associated with scrambled-order calls. This preference may be driven by an avoidance of the “unnatural” scrambled-order calls. Since both the natural-order and scrambled-order calls used in this study were both constructed from manually assembled notes that may have been used in different positions as compared to the position they were produced in (e.g., an A note may have been the first A note produced in the original call, but it could be used as the second or third A note in a manufactured call), this may create discrepancies within the information passed along in the manufactured calls. The notes in any given manufactured call contained notes from multiple calls of differing compositions and from multiple birds, so future research should investigate if our manufactured natural-ordered calls are perceived by black-capped chickadees as perceptually-similar to naturally-produced *chick-a-dee* calls. If both individual notes and the overall note structure both contain information, birds may perceive these manufactured calls as lacking information in comparison to naturally-produced calls, even though the note order is grammatically-correct. Alternatively, birds may simply prefer naturally-ordered calls due to the very high frequency with which these calls are encountered in the wild (nearly 100% of the time) compared to infrequently-encountered scrambled-ordered calls.

Individual differences in order (natural vs. scrambled) and strength of preference may be related to how individuals perceive and respond to different vocalizations. For example, Riters and colleagues (2013) showed that some European starlings (*Sturnus vulgaris*) find hearing conspecific vocalizations to be rewarding whereas other individuals do not. A similar difference in perception or motivation may be at work here, and may explain differences in motivation and, subsequently, learning. In this experiment, we found that birds with a strong preference for a stimulus type took longer to learn the discrimination task if the preferred stimulus type was the rewarded category. For example, birds in the Natural Order S+ group took longer to learn to respond to natural order stimuli if they had a strong preference for natural stimuli compared to if their preference was for scrambled order stimuli.

Conclusions

In the present study, we used two behavioural tasks to evaluate the relationship between preference for and ability to discriminate between natural- and scrambled-order *chick-a-dee* calls. Our results indicate that, not only do chickadees display individual preference for one type of call over the other, and interestingly a preference for silence over any playback, the magnitude and direction of their preference influences the rate at which they learn to discriminate between the two types of calls, natural- and scrambled-order chick-a-dee calls. Chickadees not only learned to discriminate between natural- and scrambled-order *chick-a-dee* calls, but they did so in two different, complimentary ways: open-ended categorization and rote memorization. The former mechanism was then demonstrated when the birds applied the learned contingencies to testing stimuli, suggesting that natural- and scrambled-order calls are distinct perceptual open-ended categories. The fact that the birds also discriminated randomly selected groups of calls suggests that they performed this task using rote memorization. These two, complimentary cognitive mechanisms can have significant impacts beyond the acoustic discriminations discussed here, and might be widely-used for other activities critical to survival. To extend these findings, future research should evaluate the validity of using manufactured natural-ordered call stimuli as a substitute for naturally-produced calls, and also determine which other natural tasks employ open-ended categorization and rote memorization and the particular manner in which their use differs depending upon the type and stage of the task they are employed.

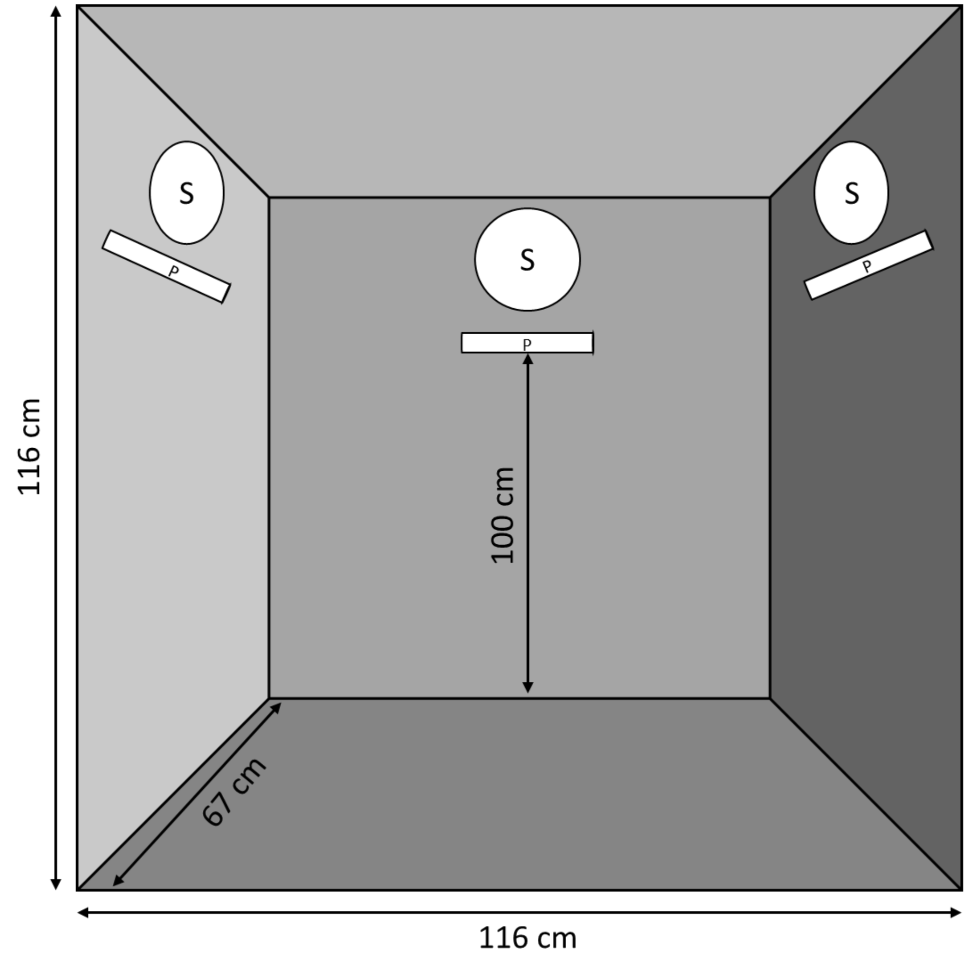

Figure 1. Diagram depicting the layout of the choice preference task testing chamber when viewed from the front. “S” indicates a speaker. “P” indicates a perch with infrared sensors.

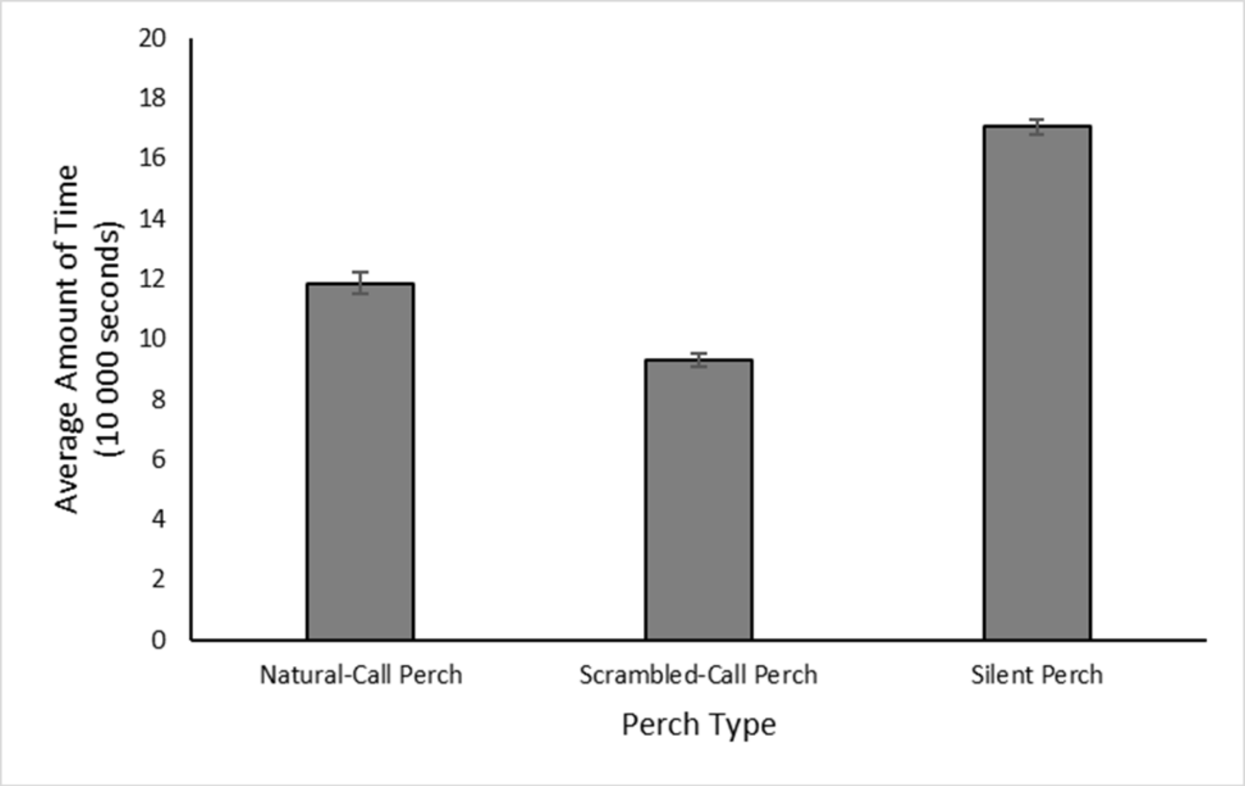

Figure 2. Average ± SE amount of time spent on natural-call, scrambled-call, and silent perches during the choice preference task.

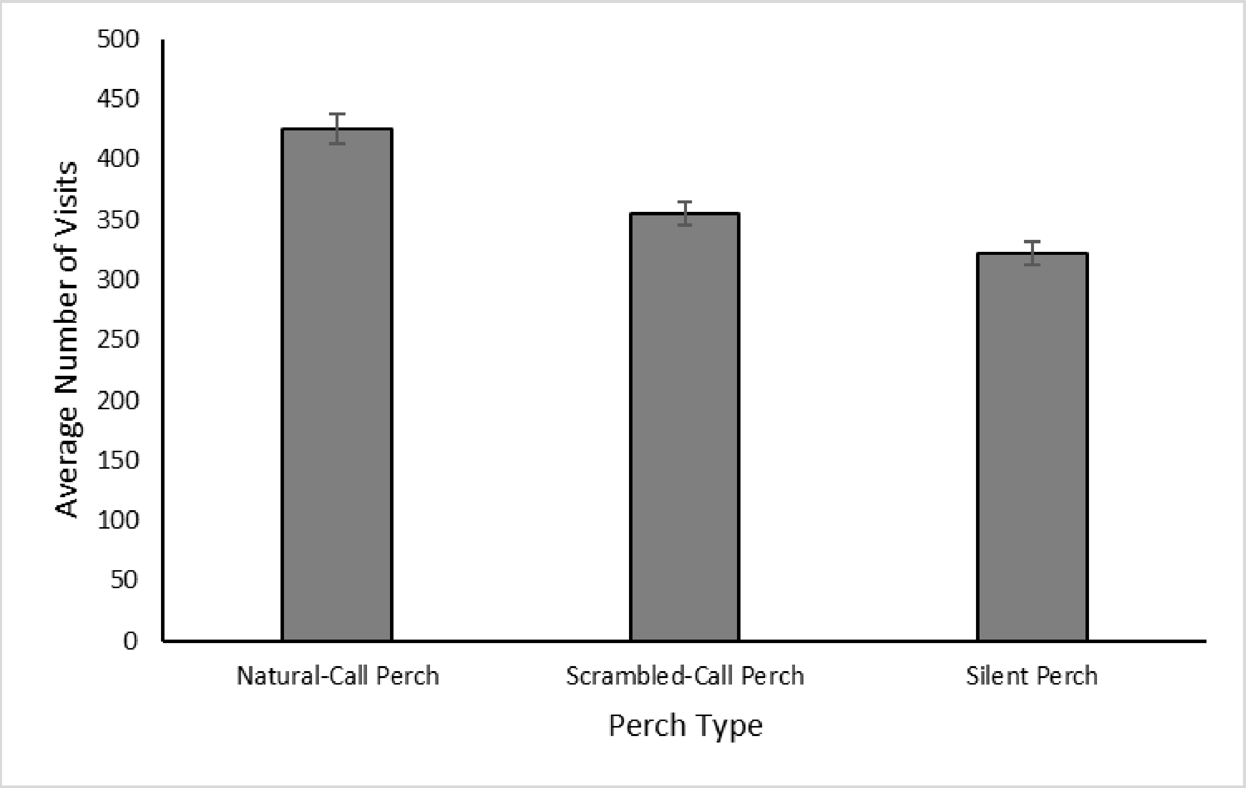

Figure 3. Average ± SE number of visits to natural-call, scrambled-call, and silent perches during the choice preference task.

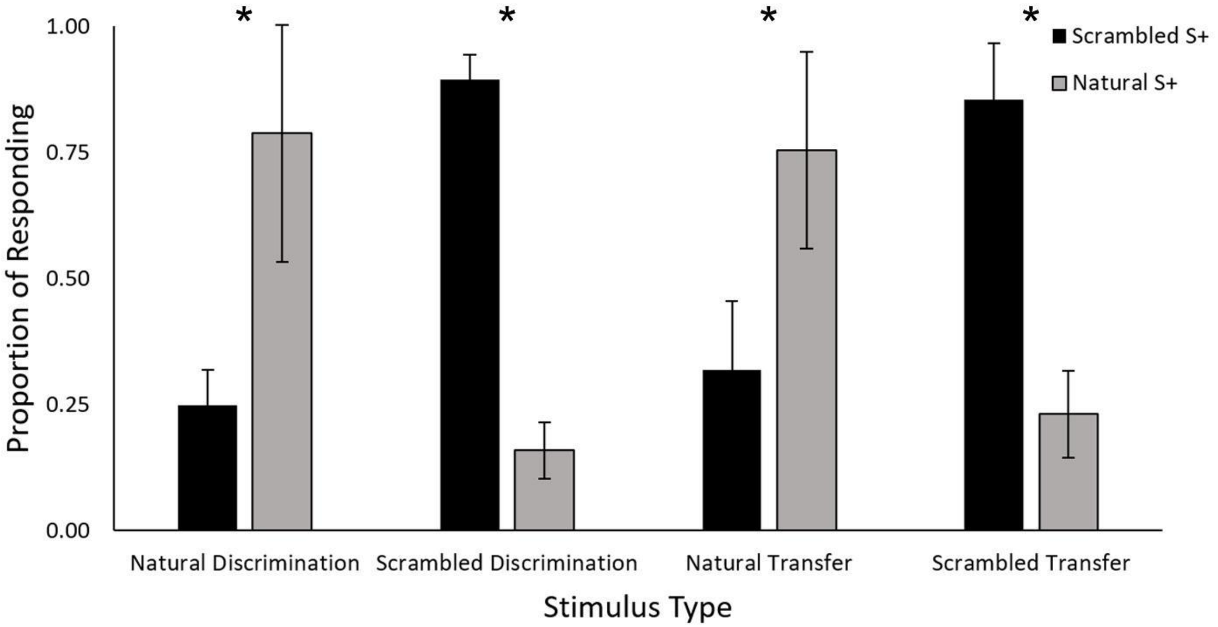

Figure 4. Average ± SE proportion of responding by birds in the True category groups (e.g., Natural S+, Scrambled S+) to four types of stimuli: natural order stimuli from Discrimination training, scrambled order stimuli from Discrimination training, novel natural-ordered stimuli in Transfer testing, and scrambled order stimuli in Transfer testing. * indicates a significant difference ($p \leq .05$) between group means.

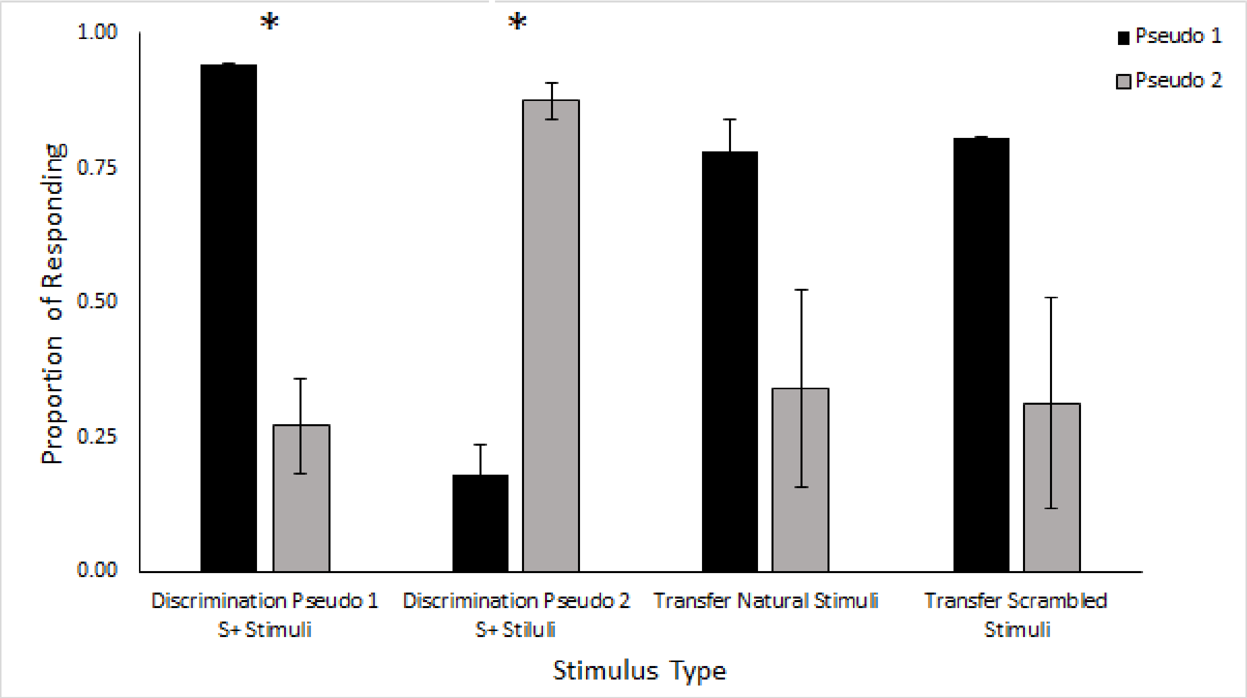

Figure 5. Average ± SE proportion of responding by birds in the Pseudo category groups (e.g., Pseudo 1, Pseudo 2) to four types of stimuli: Pseudo 1 reinforced stimuli from Discrimination training, Pseudo 2 reinforced stimuli from Discrimination training, novel natural-ordered stimuli in Transfer testing, and scrambled order stimuli in Transfer testing. * indicates a significant difference ($p \leq 0..05$) between group means.

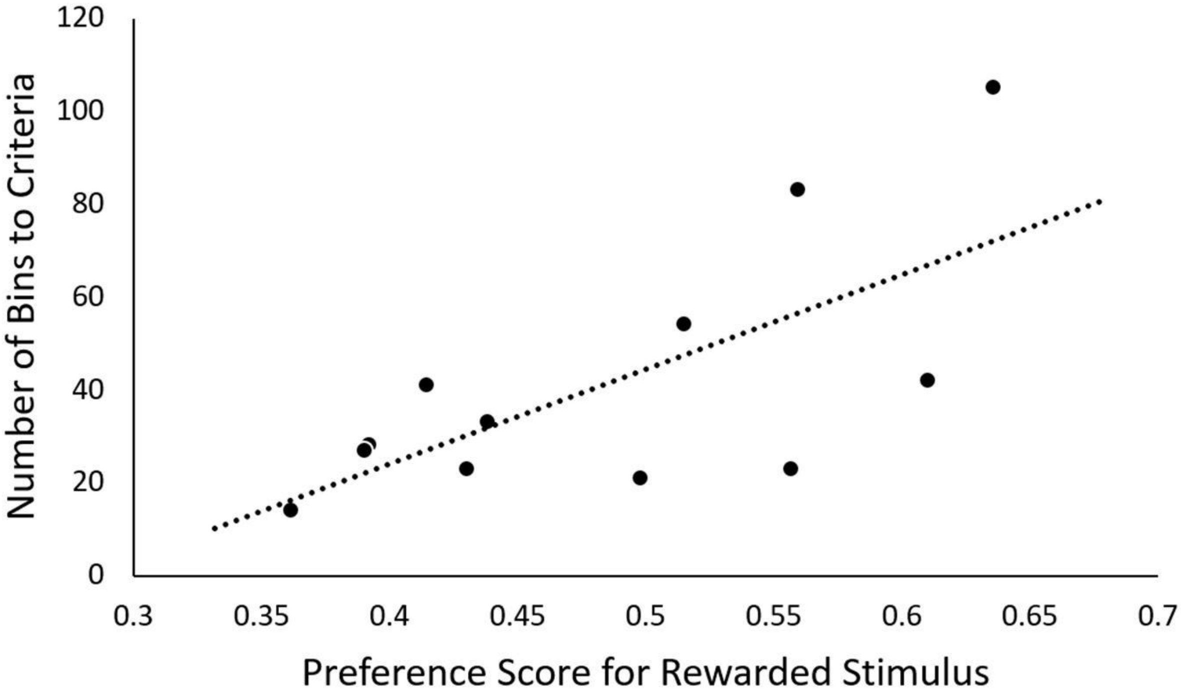

Figure 6. The relationship between the number of bins required to meet criteria in Discrimination training and the preference score calculated for the rewarded stimulus during Discrimination training using the number of visits to each perch.

Acknowledgements

All animal studies were conducted in accordance with the Canadian Council on Animal Care guidelines and policies and with approval from the Animal Care and Use Committee for Biosciences for the University of Alberta. Chickadees were captured and research was carried out under an Environment Canada Canadian Wildlife Service Scientific permit, Alberta Fish and Wildlife Capture and Research permits, and a City of Edmonton Parks Permit. This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant RGPIN-2016-04591) and Discovery Accelerator Supplement (NSERC RGPAS-2016-412311) funding to CBS. JVC and KAC were supported by Alexander Graham Bell Canada Graduate Scholarships-Doctoral (NSERC CGS D). Dan Mann was supported by funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 101028815. We thank Isaac Lank, Phil May, and Al Denington for their technical assistance.

References

Bloomfield, L. L., Phillmore, L. S., Weisman, R. G., & Sturdy, C. B. (2005). Note types and coding in Parid vocalizations. III: The *chick-a-dee* call of the Carolina chickadee (*Poecile carolinensis*). *Canadian Journal of Zoology*, *833*, 820-833. https://doi.org/10.1139/Z05-067

Bloomfield, L. L., Farrell, T. M., & Sturdy, C. B. (2008). All "*chick-a-dee*" calls are not created equally. Part II: Mechanisms for discrimination by sympatric and allopatric chickadees. *Behavioural Processes*, *77*, 87-99. https://doi.org/10.1016/j.beproc.2007.06.008

Bloomfield, L. L., & Sturdy, C. B. (2008). All "*chick-a-dee*" calls are not created equally. Part I: Open-ended categorization of chick-a-dee calls by sympatric and allopatric chickadees. *Behavioural Processes*, *77*, 73-86. https://doi.org/10.1016/j.beproc.2007.06.011

Bloomfield, L. L., Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (2003). Open-ended categorization of chick-a-dee calls by black-capped chickadees (*Poecile atricapilla*). *Journal of Comparative Psychology*, *117*, 290-301. https://doi.org/10.1037/0735-7036.117.3.290

Candiotti, A., Zuberbühler, K. & Lemasson, A. (2012). Context-related call combinations in female Diana monkeys. *Animal Cognition, 15*, 327-339. https://doi.org/10.1007/s10071-011-0456-8

Charrier, I., Bloomfield, L. L., & Sturdy, C. B. (2004). Note types and coding in parid vocalizations. I: The chick-a-dee call of the black-capped chickadee (*Poecile atricapillus*). *Canadian Journal of Zoology*, *779*, 769–779. https://doi.org/10.1139/Z04-045

Fitch, W. T. (2010). *The Evolution of Language*. New York, NY: Cambridge University Press.

Freeberg, T. M., Lucas, J. R., & Clucas, B. (2003). Variation in *chick-a-dee* calls of a Carolina chickadee population, *Poecile carolinensis*: Identity and redundancy within note types. *The Journal of the Acoustical Society of America*, *113*, 2127-2136. https://doi.org/10.1121/1.1559175

Griffiths, R., Double, M. C., Orr, K., & Dawson, R. J. (1998). A DNA test to sex most birds. *Molecular Ecology*, *7*, 1071-1075.

Hahn, A. H., Guillette, L. M., Hoeschele, M., Otter, K. A., Ratcliffe, L. M., & Sturdy, C. B. (2017). Discrimination of male black-capped chickadee songs: Relationship between acoustic preference and performance accuracy. *Animal Behaviour*, *126*, 107–121. https://doi.org/10.1016/j.anbehav.2017.02.001

Hahn, A. H., Hoang, J., McMillan, N., Campbell, K. A, Congdon, J. V., & Sturdy, C. B. (2015). Biological salience influences performance and acoustic mechanisms for the discrimination of male and female songs. *Animal Behaviour*, *104*, 213-228. https://doi.org/10.1016/j.anbehav.2015.03.023

Hailman, J. P. & Ficken, M. S. (1986). Combinatorial animal communication with computable syntax: *Chick-a-dee* calling qualifies as ‘language’ by structural linguistics. *Animal Behaviour*, *34*, 1899–1901.

Hailman, J. P., Ficken, M. S., & Ficken, R. W. (1985). The *'chick-a-dee'* calls of *Parus atricapillus*: a recombinant system of animal communication compared with written English. *Semiotica*, *56*, 191-224.

Hurd, C. R. (1996). Interspecific Attraction to the Mobbing Calls of Black-Capped Chickadees (*Parus atricapillus*). *Behavioral Ecology and Sociobiology*, *38*, 287-292. https://doi.org/10.1007/s002650050244

Mammen, D. L., & Nowicki, S. (1981). Individual differences and within-flock convergence in chickadee calls. *Behavioral Ecology and Sociobiology*, *9*, 179-186.

Nowak, M. A., Plotkin, J. B., & Jansen, V. A. (2000). The evolution of syntactic communication. *Nature*, *404*, 495-498.

Nowicki, S. (1989). Vocal plasticity in captive black-capped chickadees: the acoustic basis and rate of call convergence. *Animal Behaviour*, *37*, 64-73. https://doi.org/10.1016/0003-3472(89)90007-9

Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's Monkeys Use Affixation to Alter Call Meaning. *PLOS ONE, 4*(11), e7808. https://doi.org/10.1371/journal.pone.0007808

Pyle, P. (1997). *Identification Guide to North American Birds*. Bolinas, Slate Creek Press, Bolinas, CA.

Riters, L. V., Ellis, J. M., Angyal, C. S., Borkowski, V. J., Cordes, M. A., & Stevenson, S. A. (2013). Links between breeding readiness, opioid immunolabeling, and the affective state induced by hearing male courtship song in female European starlings (*Sturnus vulgaris*). *Behavioural Brain Research*, *2*47, 117-124.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, *210*, 801-803. https://doi.org/10.1126/science.7433999

Suzuki, T. N. & Ueda, K. (2013). Mobbing Calls of Japanese Tits Signal Predator Type: Field Observations of Natural Predator Encounters. *The Wilson Journal of Ornithology*, *125*, 412-415.

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2018). Call combinations in birds and the evolution of compositional syntax. *PLoS Biology*, *16*, e2006532. https://doi.org/10.1371/journal.pbio.2006532

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications, 7*, 10986.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of Alarm Calls: Black-Capped Chickadees Encode Information about Predator Size. *Science*, *308*, 1934-1937.