



Discrimination of High- and Low-Threat Vocalizations: An Examination of Referential Communication in Black-Capped Chickadee (*Poecile atricapillus*) Mobbing Calls

Jenna V. Congdon^{1,*}, Allison H. Hahn³, Kimberley A. Campbell¹, Erin N. Scully¹, Shannon K. Mischler¹, Carolina Montenegro¹, William D. Service¹, and Christopher B. Sturdy^{1,2}

¹Department of Psychology, University of Alberta

²Neuroscience and Mental Health Institute, University of Alberta

³Department of Psychology, St. Norbert College

*Corresponding author (Email: jcongdon@ualberta.ca)

Citation – Congdon, J. V., Hahn, A. H., Campbell, K. A., Scully, E. N., Mischler, S. K., Montenegro, C., Service, W. D., & Sturdy, C. B. (2022). Discrimination of high- and low-threat vocalizations: An examination of referential communication in black-capped chickadee (*Poecile atricapillus*) mobbing calls. *Animal Behavior and Cognition*, 9(3), 298-315. <https://doi.org/10.26451/abc.09.03.04.2022>

Abstract – Referential communication has been defined as the exchange of information regarding an object or event, but few studies have examined referential alarm calls in songbirds. In contrast, it has been well-supported that chickadees produce mobbing calls in response to predators that vary depending on the threat level posed, and the auditory brain areas of chickadees produce similar neural expression in response to predator calls and conspecific mobbing calls of the same threat level. This suggests that chickadees perceive these acoustically distinct vocalizations as similar, potentially as referent signals. In the current study, we trained 33 birds on an operant go/no-go discrimination task in which chickadees were presented with predator and mobbing calls of high- or low-threat. Following the first round of training, birds completed a second round with high- or low-threat calls and we predicted that birds would show transfer of training when contingencies (i.e., threat level regarding a species of predator) were the same between hetero- and conspecific vocalizations. However, high- and low-threat mobbing calls were not treated similarly to the corresponding predator's calls. Our asymmetrical results may be due to the acoustic distinction between calls produced by two owl predator species, but we cannot make definitive conclusions about referential communication from this study. Nonetheless, we believe that this experiment was an important exploration of the perception of referential signal.

Keywords – Referential communication, Predator perception, Black-capped chickadee, Operant conditioning

Referential communication has been defined as the exchange of information giving reference to objects or events in the environment (Seyfarth & Cheney, 1993). Animals that have the ability for referential communication must: 1) be able to produce signals, 2) produce these signals under the correct circumstances, and 3) have receivers respond appropriately (Maynard Smith & Harper, 2003). This type of communication is connected to semantics (Marler et al., 1992), and is known to be used by a wide range of species. For example, referential communication is evident in honeybees (*Apis mellifera carnica*) that dance to communicate the location of a food source (i.e., direction and distance, the language property of displacement; Riley et al., 2005), chickens (*Gallus gallus*) that produce 'food' calls that are distinct from their alarm calls (Evans & Evans, 1999; Evans et al., 1993), and arctic ground squirrels (*Citellus undulatus*)

that produce acoustically distinct *whistle* and *chatter* alarm calls in response to aerial danger and ground danger, respectively (Melchior, 1971).

The threat of predation influences individuals' behavior, which in turn impacts survival and fitness as an evolutionary force. Consequently, adaptive antipredator responses (such as informative signals) should be selected for over time (Lima & Dill, 1990). Vocal antipredator responses are a well-studied class of antipredator behaviors (Maynard Smith & Harper, 2003). In addition to signaling the presence of a predator, vocal antipredator responses can provide additional information about the predator type. Vervet monkeys (*Chlorocebus pygerythrus*) produce different alarm calls that signal either the presence of a leopard, eagle, or snake, and each of these calls results in an adaptive antipredator, species-specific behavioral response from nearby receiver vervet monkeys (Seyfarth et al., 1980). In addition, there also appear to be unspecific general alarm calls that are used more flexibly by vervet monkeys, using context and prior exposure to determine meaning, leaving much to be understood in the exploration of referential communication (Fichtel, 2020). Conversely, yellow warblers (*Setophaga petechia*) are known to produce specific referential calls (i.e., *seets*), rather than their more general alarm calls (i.e., *chips*), to alert conspecific mates to the presence of either brown-headed cowbirds (*Molothus ater*) that parasitize warbler nests, or in response to *seet* calls of other yellow warblers. However, warbler responses to each alarm call type depend on whether the birds were paired with their mates (Gill & Sealy, 2004; Lawson et al., 2021).

In addition to predator type, vocal antipredator responses can also communicate the threat level posed by a predator. Both Carolina chickadees (*Poecile carolinensis*) and black-capped chickadees (*P. atricapillus*) produce *chick-a-dee* calls containing four note types that can be included, omitted, or repeated, but occur in a fixed order (A→B→C→D; Bloomfield et al., 2005; Charrier et al., 2004). Both species produce the *chick-a-dee* call as a mobbing call in the presence of a predator, used to recruit and coordinate conspecifics and heterospecifics to attack (Smith, 1991). Typically, these mobbing calls are produced with a greater number of D notes per call in response to a small, high-threat predator (e.g., northern saw-whet owl) relative to a large, low-threat predator (e.g., great horned owl; Soard & Ritchison, 2009; Templeton et al., 2005). Thus, high-threat predators result in the production of high-threat *chick-a-dee* mobbing calls; this threat-specific modification in note composition allows for the communication of threat to other individuals and consequently adaptive responses from conspecifics and heterospecifics in the area (Soard & Ritchison, 2009; Templeton et al., 2005).

Furthermore, upon investigating neural expression in the auditory system of black-capped chickadees, Avey et al. (2011) found that more immediate early gene expression (IEG) occurs in response to high-threat predator vocalizations and high-threat mobbing calls compared to low-threat predator and mobbing calls. Typically, higher levels of IEG induction occur with more salient, biologically relevant signals (e.g., conspecific song > white noise). This similarity in IEG expression within threat level occurred despite acoustic differences between the calls by hetero- versus conspecifics (Avey et al., 2011), which supported and extended the findings of both Templeton et al. (2005) and Soard and Ritchison (2009). Based on these findings of concordance between IEG expression across stimuli of the same threat level, potentially referring to the specific predator species, we hypothesized that a similar pattern in perception of such calls would occur in a behavioral task. Specifically, if high-threat mobbing calls are produced in reference to high-threat predators, and the neural response to predator-produced calls and mobbing calls of the same threat level, connected to the same species are similar, is it possible that chickadee-produced mobbing calls provide referential information about which predator is nearby?

Referential communication has most commonly been studied in humans and nonhuman primates (Seyfarth & Cheney, 1993); however, little evidence has been provided to support these abilities in songbirds. Through playback, it has been found that incubating female Japanese tits (*Parus minor*) respond differentially to alarm calls made in response to snakes versus crows and martens (Suzuki, 2015). To further explore the perception of alarm calls, here we asked whether black-capped chickadees perceive predator calls produced by owls, and matched mobbing calls produced in response to those predators, as similar. Designed as an operant conditioning experiment, this study functions as a necessary and expanded exploration of how referential signals are perceived, directly comparing a transfer of responding based on the potentially referent information contained within mobbing calls.

Method

Ethics Statement

All 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 1937), 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 (#56066 and #56065), and City of Edmonton Parks permit. Throughout the experiment, birds remained in the testing apparatus to minimize the transport and handling of each bird. We monitored birds closely to ensure there were no ill effects as a result of the stimuli used in these experiments. Three male subjects died from natural causes during operant training. One male subject became ill during operant training and was consequently humanely euthanized. At the completion of the experiment, birds were returned to the colony room for use in future experiments.

Protocol

We designed a controlled laboratory experiment to train and test birds in their perception of natural, relevant stimuli in an operant go/no-go discrimination task. Predator-produced vocalizations (referred to as “predator calls” throughout) and mobbing calls have both been used in operant go/no-go discrimination tasks (Congdon et al., 2019; 2020a, b), but responding to both predator and mobbing calls have never been compared. We trained one group of chickadees to respond (‘go’) to high-threat predator calls and withhold responding (‘no-go’) to low-threat predator calls (Discrimination Training I) and then trained those subjects with conspecific mobbing calls (i.e., high- or low-threat; Discrimination Training II; Owl S+ subgroups). For purposes of counterbalancing subjects, another group of chickadees were first trained with conspecific mobbing calls then trained with predator calls (MOB S+ subgroups). These were the True Transfer category subgroups in which there were true categories (i.e., high- vs. low-threat categories compared to pseudorandomized stimuli) to learn and the rewarded contingency was consistent between training rounds. Considering that the contingencies (e.g., high threat) were the same between the first and second round of training (e.g., mobbing calls then predator calls), if mobbing calls contain information about predators, transfer of training would be possible between the two types of stimuli and the subsequent round of training should be completed in fewer trials. The design of this experiment was crucial to directly query if mobbing calls and predator calls are perceived similarly by birds (i.e., categorized using the same threat parameters) as examining the rate of learning in an experimental setting with high stimulus control allows us to equate the response across different stimuli types (i.e., conspecific and heterospecific).

For a control, we included a group in which birds were rewarded for pseudorandomized stimuli (selected using a random number generator) with no category of threat level (Pseudo category subgroups). In addition, we included a reversal group in which the threat-contingencies were reversed between Discrimination Training I and Discrimination Training II (e.g., first trained with high-threat owl calls then trained with low-threat conspecific mobbing calls, or vice versa; True Reversal category subgroups). Again, two subgroups were included for the purposes of counterbalancing. This resulted in three groups: (1) True Transfer, (2) Pseudo (i.e., control), and (3) True Reversal. The final, True Reversal category group provided a completely unique type of comparison in an attempt to ensure that birds in the True Transfer category group were completing the task based on consistent contingencies between rounds of training. Thus, the True Transfer category group should complete the second round of training in fewer trials than the first round (i.e., Discrimination Training II in fewer trials than Discrimination Training I) due to the use of this perceptual category (e.g., high threat owl = high threat conspecific mobbing call). Conversely, the True Reversal category group should take a similar number of trials to complete both rounds within training (i.e., Discrimination Training I and Discrimination Training II, described below), as the contingencies would

differ and transfer of training would not be possible (i.e., the rationale for naming this group “Reversal” rather than “Transfer”).

First, we predicted that True Transfer category groups and True Reversal category groups would complete Discrimination Training I in fewer trials compared to Pseudo category subgroups due to the ability to use categories of perceptually similar stimuli being learned in fewer trials than groups of pseudorandomized stimuli. We also predicted that True Transfer category groups and True Reversal category groups would complete Discrimination Training II in fewer trials compared to Pseudo category groups. Third, we predicted that within True Transfer, True Reversal, and Pseudo category groups, the subgroups would not differ between the first and second round of training as each of the subgroups had similar transfer, reversal, or no contingencies. Finally, we predicted that birds that received the same threat-contingencies in Discrimination Training I and Discrimination Training II would show transfer of training. For example, birds first trained to respond to high-threat predator calls were predicted to demonstrate transfer of training in that they would then discriminate high-threat mobbing calls (i.e., different stimulus type, but same contingency of high-threat) in fewer trials. If transfer of training was demonstrated by all True Transfer category subgroups completing Discrimination Training II in fewer trials than Discrimination Training I, it would suggest that chickadees perceive classes of mobbing calls and predator calls as similar. In summary, these results would suggest that mobbing calls provide referential information about threat levels, and thus, provide evidence of referential communication in a songbird.

Subjects

Thirty-seven black-capped chickadees (19 male, 18 female, identified by DNA analysis; Griffiths et al., 1998) were tested between June and September 2018; in total, thirty-three black-capped chickadees (15 males, 18 females) completed the experiment (see Ethics Statement). Chickadees at least one year of age (determined by examining the color and shape of their outer tail retrices; Pyle, 1997) were attracted to long-term feeder sites and captured using walk-in potter traps in Edmonton (North Saskatchewan River Valley, 53.53°N, 113.53°W, Mill Creek Ravine, 53.52°N, 113.47°W), Alberta, Canada between January 9-26, 2018.

Prior to the experiment, birds were individually housed in Jupiter Parakeet cages (30 × 40 × 40 cm; Rolf C. Hagen, Inc., Montreal, QB, Canada) in colony rooms that were maintained on a light:dark cycle that mimicked the natural light cycle for Edmonton, Alberta, Canada. In the colony rooms, birds had visual and auditory, but not physical, contact with one another. Birds had *ad libitum* access to food (Mazuri Small Bird Maintenance Diet; Mazuri, St Louis, MO, USA), water (vitamin supplemented on alternating days; Prime vitamin supplement; Hagen, Inc.), grit, and cuttlebones. Birds were given three to five sunflower seeds daily, one superworm (*Zophobas morio*) three times a week, and a mixture of greens (spinach or parsley) and eggs twice a week.

Throughout the experiment, birds were housed individually in operant chambers (see Apparatus), maintained on the natural light cycle for Edmonton, Alberta, and had *ad libitum* access to water (vitamin supplemented on alternate days), grit, and cuttlebone. Birds were given two superworms daily (one in the morning and one in the afternoon). Food (i.e., Mazuri diet) was only available as a reward for correctly responding during the operant discrimination task. Only three birds had previous experimental experience hearing black-capped chickadee-produced *fee-bee* songs during a playback experiment in March 2018 (Montenegro et al., unpublished), but were naïve to operant experiments and the stimuli used in the current experiment.

Apparatus

During the experiment, birds were housed individually in modified colony room cages (30 × 40 × 40 cm) placed inside a ventilated, sound-attenuating chamber. The chambers were illuminated by a 9-W, full spectrum fluorescent bulb. Each cage contained three perches, a water bottle, and a grit cup. An opening on the side of the cage (11 × 16 cm) provided each bird access to a motor-driven feeder (see Njegovan et

al., 1994). Infrared cells in the feeder and the request perch (perch closest to the feeder) monitored the position of the bird (i.e., perching and feeder entry). A personal computer connected to a single-board computer (Palya & Walter, 2001) scheduled trials and recorded responses to stimuli. Stimuli were played from the personal computer hard drive, through either a Cambridge A300 Integrated Amplifier, Cambridge Azur 640A Integrated Amplifier (Cambridge Audio, London, England), or an NAD310 Integrated Amplifier (NAD Electronics, London, England) and through a Fostex FE108 Σ or Fostex FE108E Σ full-range speaker (Fostex Corp., Japan; frequency response range 80-18,000 Hz) located beside the feeder. The operant system is closed, meaning that birds received life-sustaining food only from completing the task. See Sturdy and Weisman (2006) for a detailed description of the apparatus.

Acoustic Stimuli

A total of 68 vocalizations were used as stimuli in the current experiment: 17 predator calls produced by high-threat northern saw-whet owls (NSWO; *Aegolius acadicus*), 17 predator calls produced by low-threat great horned owls (GHOW; *Bubo virginianus*), 17 mobbing calls produced by black-capped chickadees made in response to mounts of NSWO, and 17 mobbing calls produced by black-capped chickadees made in response to mounts of GHOW. According to Templeton et al. (2005), NSWO and GHOW are on opposite ends of the spectrum in regards to threat level, and henceforth we refer to NSWO as ‘high-threat’ and GHOW as ‘low-threat’ for the procedures of our study.

All owl calls were obtained through the Borror Laboratory of Bioacoustics (The Ohio State University) and field recordings contributed by the Bayne Laboratory (Department of Biological Sciences, University of Alberta). All black-capped chickadee-produced mobbing calls were recorded in the laboratory by Avey et al. (2011) and used in Congdon et al. (2019). The average \pm standard deviation for the duration of NSWO call stimuli was 2037.6 ± 423.8 ms (range = 1558.2-2440.2 ms); the duration of GHOW call stimuli was 2095.6 ± 325.8 ms (range = 708.6-2618.2 ms); the duration of mobbing call stimuli produced by black-capped chickadees made in response to mounts of NSWO was 1536.1 ± 351.7 ms (range = 1050.9-2389.6 ms); and the duration of mobbing call stimuli produced by black-capped chickadees made in response to mounts of GHOW 765.3 ± 200.7 ms (range = 516.5-1293.1 ms). There was an average of 4.8 D notes per call for stimuli in response to NSWO and 2.2 D notes per call for stimuli in response to GHOW, and, thus, a significant difference in the duration of mobbing calls produced to NSWO versus GHOW ($t_{16} = -7.584$, $p < .001$, $d = 2.691$, 95% CIs = -986.283, -555.340).

All vocalizations were of high quality (i.e., no audible interference and low background noise when viewed on a spectrogram with amplitude cutoffs of -35 to 0 dB relative to vocalization peak amplitude) and were bandpass filtered (outside the frequency range of each vocalization type) using GoldWave version 5.58 (GoldWave, Inc., St. John’s, NL, Canada) to reduce any background noise. For each stimulus, 5 ms of silence was added to the leading and trailing portion of the vocalization and tapered to remove transients, and amplitude was equalized using SIGNAL 5.10.24 software (Engineering Design, Berkeley, CA, USA). During the experiment, stimuli were presented at approximately 75 dB as measured by a Brüel and Kjær Type 2239 (Brüel & Kjær Sound & Vibration Measurement A/S, Nærum, Denmark) decibel meter (A-weighting, slow response) at the approximate height and position of a bird’s head while sitting on the request perch directly in front of the feeder.

Detailed Procedure

Pretraining

Pretraining began once the bird learned to use the request perch and feeder to obtain food. During Pretraining, birds received food for responding to all stimuli (future S+, S-, and transfer stimuli). A trial began when the bird landed on the request perch and remained for between 900-1100 ms. The computer program played a randomly selected stimulus without replacement until all 68 stimuli had been heard. If

the bird left the request perch before a stimulus finished playing, the trial was considered interrupted, resulting in a 30-s time out with the houselight turned off. If the bird entered the feeder within 1 s after the entire stimulus played, it was given 1 s access to food, followed by a 30-s intertrial interval, during which the houselight remained on. If a bird remained on the request perch during the stimulus presentation and 1 s following the completion of the stimulus, it received a 60-s intertrial interval with the houselight on, but this intertrial interval was terminated if the bird left the request perch. This was to encourage a high level of responding on all trials. Birds continued on Pretraining until they completed six 340-trial bins with $\geq 60\%$ responding on average to all stimuli, at least four 340-trial bins with $\leq 3\%$ difference in responding to future S+ and S- stimuli, at least four 340-trial bins in which the bird had $\leq 3\%$ difference in responding to future high- and low-threat transfer stimuli, and at least four 340-trial bins in which the bird had $\leq 3\%$ difference in responding to short and long stimuli to ensure that birds did not display biases.

Once birds met the above criteria, they were given a day with unlimited access to food without auditory stimuli, then birds completed a second round in which they completed one 340-trial block with $\geq 60\%$ responding on average to all stimuli, completed one 340-trial block of $\leq 3\%$ difference in responding to future S+ and S- stimuli, completed one 340-trial block of $\leq 3\%$ difference in responding to future high- and low-threat transfer stimuli, and completed one 340-trial block of $\leq 3\%$ difference in responding to short and long stimuli, to confirm that each bird continued to not display biases following the break (this criterion has been used by similar operant go/no-go experiments; e.g., Hahn et al., 2015; 2016; McMillan et al., 2017). It is important to note that the operant system is a closed economy throughout all stages of the procedure, meaning that birds received life-sustaining food only from completing the task.

Discrimination Training I

The procedure was the same as during Pretraining; however, of the original 68 stimuli, only the 34 training stimuli were presented (with the remaining 34 withheld for use during Transfer testing) and responding to half of these stimuli were punished with a 30-s intertrial interval with the houselight off. As during Pretraining, responses to rewarded (S+) stimuli resulted in 1 s access to food and responses to unrewarded (S-) stimuli resulted in a 30-s time out with the houselight turned off. Discrimination Training I continued until birds completed six 340-trial bins with a discrimination ratio (DR) ≥ 0.80 with the last two bins being consecutive. For discrimination ratio (DR) calculations, see Response Measures, below.

The 33 birds (excluding the four birds that died during training/testing) were randomly assigned to either a True Transfer category discrimination group ($n = 15$), Pseudo category discrimination group ($n = 6$), or True Reversal category discrimination group ($n = 12$); to reduce the number of subjects included in this study, we assigned fewer birds to the Pseudo (i.e., control) groups than the True groups. Black-capped chickadees in the True Transfer category discrimination group were divided into four subgroups: 1) one subgroup discriminated 17 rewarded (S+) high-threat owl calls from 17 unrewarded (S-) low-threat owl calls (High Owl S+ Group: two male and two female subjects), 2) while the other subgroup discriminated 17 rewarded (S+) low-threat owl calls from 17 unrewarded (S-) high-threat owl calls (Low Owl S+ Group: two male and two female subjects); 3) another subgroup discriminated 17 rewarded (S+) high-threat mobbing calls from 17 unrewarded (S-) low-threat mobbing calls (High MOB S+ Group: one male and two female subjects), 4) while the other subgroup discriminated 17 rewarded (S+) low-threat mobbing calls from 17 unrewarded (S-) high-threat mobbing calls (Low MOB S+ Group: two male and two female subjects); see Figure 1.

The Pseudo category discrimination group was also divided into four subgroups. Two of the subgroups discriminated owl stimuli: eight randomly-selected rewarded (S+) high-threat owl and nine randomly-selected rewarded (S+) low-threat owl calls from nine unrewarded (S-) high-threat owl and eight unrewarded (S-) low-threat owl calls (Total of 34 stimuli; Pseudo 1 Owl Group: one female subject; Pseudo 2 Owl Group: one male and one female subject). The other two subgroups discriminated mobbing stimuli: eight randomly-selected rewarded (S+) high-threat mobbing and nine randomly-selected rewarded (S+) low-threat mobbing calls from nine unrewarded (S-) high-threat mobbing and eight unrewarded (S-) low-threat mobbing calls (Total of 34 stimuli; Pseudo 1 MOB Group: one female subject; Pseudo 2 MOB Group:

Figure 1

Schematic of the Stimulus Types and Reward Contingencies for Discrimination Training I and Discrimination Training II for the Four (4) Subgroups of the True Transfer Group

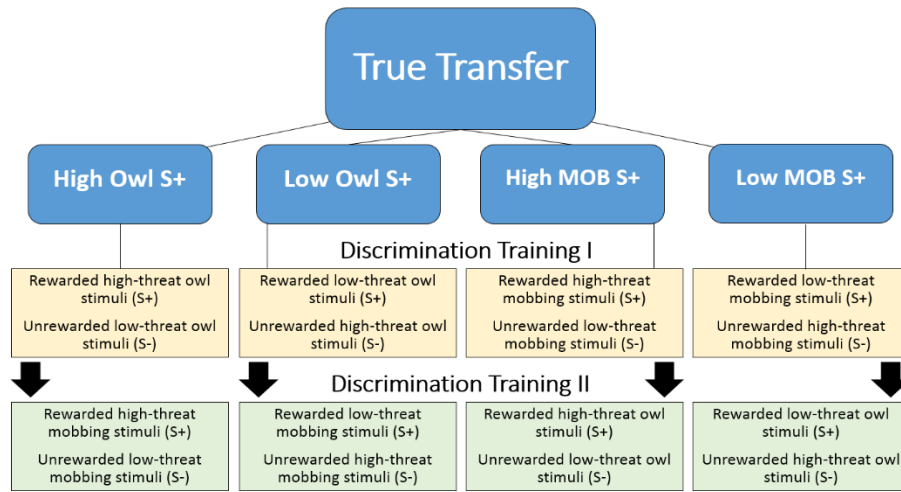
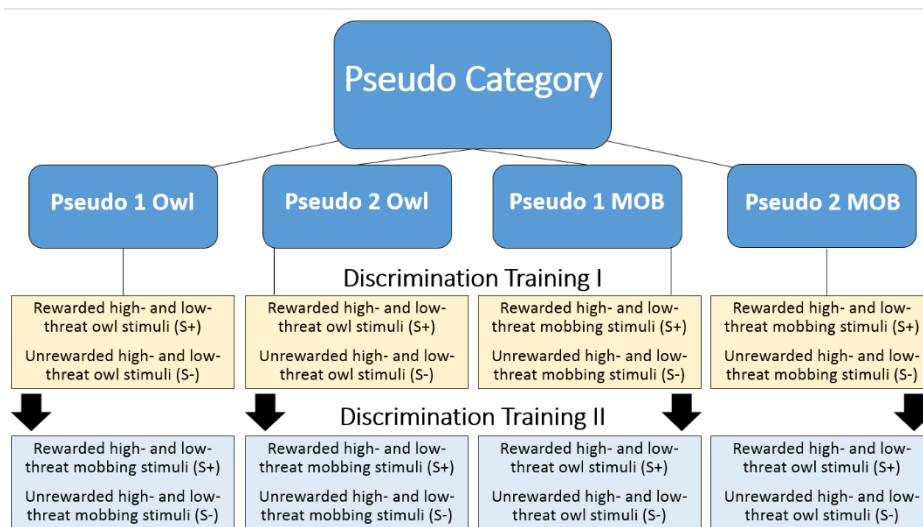


Figure 2

Schematic of the Stimulus Types and Reward Contingencies for Discrimination Training I and Discrimination Training II for the Four (4) Subgroups of the Pseudo Category Group

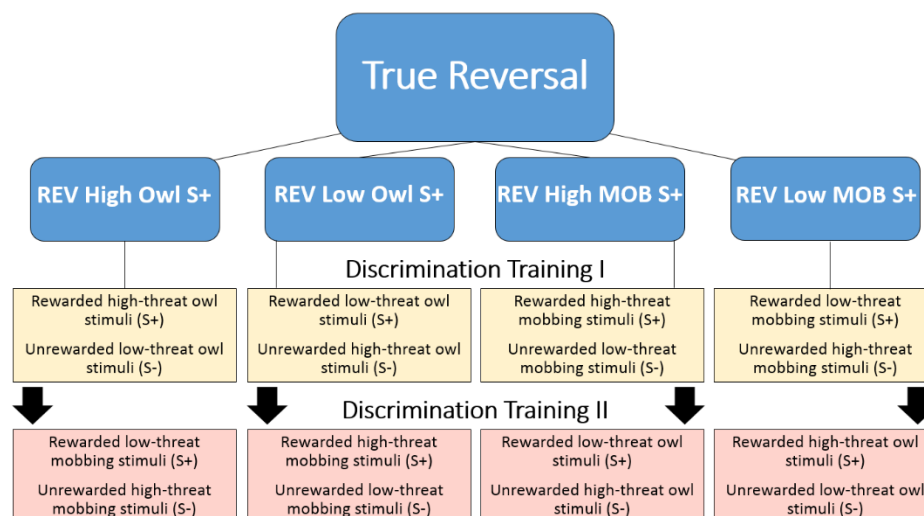


one male and one female subject); see Figure 2. The purpose of the Pseudo category group was to include a control in which subjects were not trained to categorize according to threat level.

Black-capped chickadees in the True Reversal category discrimination group were divided into four subgroups (REV High Owl S+: one male and two female subjects, REV Low Owl S+: two males and one female subject, REV High MOB S+: two males and one female subject, REV Low MOB S+: one male and two female subjects) in which Discrimination Training I was the same as the True Transfer subgroups, but Discrimination Training II differed (see Discrimination Training II for this differentiation); Figure 3.

Figure 3

Schematic of the Stimulus Types and Reward Contingencies for Discrimination Training I and Discrimination Training II for the Four (4) Subgroups of the True Reversal Group



Discrimination Training II

The procedure was the same as during Discrimination Training I; however, the 34 stimuli from Pretraining that were withheld from Discrimination Training I were presented. As during Pretraining and Discrimination Training I, responses to rewarded (S+) stimuli resulted in 1 s access to food and responses to unrewarded (S-) stimuli resulted in a 30-s time out with the houselight turned off.

True Transfer Owl S+ (High Owl S+, Low Owl S+) and Pseudo Owl S+ groups (Pseudo 1 Owl, Pseudo 2 Owl) were presented with mobbing stimuli during Discrimination Training II, whereas True Transfer MOB S+ (High MOB S+, Low MOB S+) and Pseudo MOB S+ groups (Pseudo 1 MOB, Pseudo 2 MOB groups) were presented with owl stimuli during Discrimination Training II (i.e., the opposite type of stimuli as presented with during Discrimination I). For example, High Owl S+ Group birds that were rewarded for responding to high-threat NSWO stimuli in Discrimination Training I were then rewarded for responding to high-threat mobbing stimuli in Discrimination Training II (i.e., the same contingency of 'high-threat'; see Figure 1).

True Reversal groups also received the opposite type of stimulus during Discrimination Training II (owl stimuli during Discrimination Training I then mobbing stimuli during Discrimination Training II, or vice versa), but received the opposite contingencies during Discrimination Training II (high-threat stimuli were rewarded during Discrimination Training I then low-threat stimuli were rewarded during Discrimination Training II, or vice versa). For example, REV High Owl S+ Group birds that were rewarded for responding to high-threat NSWO stimuli in Discrimination Training I were then rewarded for responding to low-threat mobbing stimuli in Discrimination Training II; see Figure 3. The purpose of the True Reversal group was to determine if birds in this group would take longer in Discrimination Training II to complete a reversal of training compared to the True Transfer groups transferring training. This was expected as learning a reversal of contingencies (e.g., rewarded for responding to high-threat owl stimuli then low-threat mobbing stimuli) should take more trials than learning a transfer of contingencies (e.g., rewarded for responding to high-threat owl stimuli then high-threat mobbing stimuli; referential information of high-threat/NSWO).

Discrimination Training II continued until birds completed six 340-trial bins with a discrimination ratio ≥ 0.80 with the last two bins being consecutive. For DR calculations see Response Measures, below.

Response Measures

For each stimulus exemplar, a percent response was calculated by the following formula: $R^+/(N-I)$, where R^+ is the number of trials in which the bird went to the feeder, N is the total number of trials, and I is the number of interrupted trials in which the bird left the perch before the entire stimulus played. For Discrimination Training I, we calculated a discrimination ratio (DR) by dividing the mean percent response to all S+ stimuli by the mean percent response of S+ stimuli plus the mean percent response of S- stimuli. A DR of 0.50 indicates equal responding to rewarded (S+) and unrewarded (S-) stimuli, whereas a DR of 1.00 indicates perfect discrimination.

To determine whether groups differed in speeds of acquisition during Discrimination Training II compared to Discrimination Training I, the number of trial bins to criterion during Discrimination Training I was subtracted from the number of trial bins to criterion during Discrimination Training II (DIS2-DIS1).

Statistical Analyses

We conducted one-way analyses of variance (ANOVAs) on the number of bins to discrimination criterion to compare between the subgroups. We conducted a separate ANOVA for the True Transfer, Pseudo, and True Reversal category groups. We also conducted between-groups independent-samples t -tests on the number of bins to discrimination criterion to compare True Transfer v. Pseudo, True Reversal v. Pseudo, and True Transfer v. True Reversal. Note: Our sample size remained small, despite preferring the inclusion of more birds for greater statistical power, but it was critical to consider the implications of removing more birds from the wild.

We conducted one-way ANOVAs on the number of bins to Discrimination Training II criterion between the subgroups of the True Transfer, Pseudo, and True Reversal category groups. We also conducted between-groups independent samples t -tests on the number of bins to Discrimination Training II criterion to compare True Transfer v. Pseudo, True Reversal v. Pseudo, and True Transfer v. True Reversal.

We conducted one-way ANOVAs and independent samples t -tests, with Bonferroni corrections ($p = .008$), on the difference between the number of bins in Discrimination Training II and Discrimination Training I (Discrimination Training II-Discrimination Training I) to reach criterion for the True Transfer, Pseudo, and True Reversal category groups.

We conducted independent samples t -tests on the difference between the number of bins in Discrimination Training II and Discrimination Training I (Discrimination Training II-Discrimination Training I = DIS2-DIS1) to reach criterion for the True Transfer, Pseudo, and True Reversal category groups according to MOB S+ compared to Owl S+.

Results

Discrimination Training I

First, the differences between subgroups in speed of acquisition was investigated by comparing each subgroups' bins to criterion (i.e., Transfer category subgroups, Pseudo category subgroups, and Reversal category subgroups). The data was subsequently pooled into groups as there were no significant differences between subgroups during Discrimination Training I, as expected ($ps \geq .157$). The differences between groups in speed of acquisition was then investigated by comparing each groups' bins to criterion. Thus, to compare the acquisition performance during Discrimination Training I of the True Transfer and Pseudo category groups and to determine if the True Transfer group learned to discriminate in fewer bins than the Pseudo category group, we conducted an independent-samples t -test on the number of 340-trial bins to reach criterion for the True Transfer category and Pseudo category groups. As predicted, there was a significant difference between the groups ($t_{5,079} = -3.422, p = .018, d = -3.037, 95\% \text{ CIs} = -38.010, -5.500$)

in that chickadees in the True Transfer group learned to discriminate significantly faster than chickadees in the Pseudo category group.

To compare the acquisition performance during Discrimination Training I of the True Reversal and Pseudo category groups and to determine if the True Reversal group learned to categorize in fewer bins than the Pseudo category group, we conducted an independent-samples *t*-test on the number of 340-trial bins to reach criterion for the True Reversal category and Pseudo category groups. As predicted, there was a significant difference between the groups ($t_{5,267} = -3.253$, $p = .021$, $d = -2.835$, 95% CIs = -37.649, -4.639) in that chickadees in the True Reversal group learned to discriminate significantly faster than chickadees in the Pseudo category group.

Last, to compare the acquisition performance during Discrimination Training of the True Transfer and True Reversal category groups, we conducted independent-samples *t*-test on the number of 340-trial bins to reach criterion for the True Transfer category and True Reversal category groups. As predicted, there was no significant difference between the groups ($t_{25} = -.792$, $p = .436$, $d = -.319$, 95% CIs = -30.080, -13.520). See Table 1.

Discrimination Training II

Subgroup data was again pooled as there were no significant differences ($ps \geq .069$). To compare the acquisition performance during Discrimination Training II of the True Transfer and Pseudo category groups and to determine if the True Transfer group learned to transfer training in fewer bins than the Pseudo category group discriminated, we conducted an independent-samples *t*-test on the number of 340-trial bins to reach criterion for the True Transfer category and Pseudo category groups. There was no significant difference between the groups ($t_{5,067} = -1.482$, $p = .198$, $d = -1.099$, 95% CIs = -42.637, 11.371) in that True Transfer birds did not learn to transfer discrimination significantly faster than Pseudo birds.

To compare the acquisition performance during Discrimination Training II of the True Reversal and Pseudo category groups, and to determine if the True Reversal group learned to reverse discrimination in fewer bins than the Pseudo category group discriminated, we conducted an independent-samples *t*-test on the number of 340-trial bins to reach criterion for the True Reversal category and Pseudo category groups. There was no significant difference between the groups ($t_{5,072} = -1.453$, $p = .205$, $d = -1.290$, 95% CIs = -42.336, 11.669) in that True Reversal birds did not learn to discriminate faster than Pseudo birds.

Last, to compare the acquisition performance during Discrimination Training II of the True Transfer and True Reversal category groups, and to determine if the True Transfer group learned to transfer training in fewer bins than the True Reversal group learned the reverse contingencies, we conducted an independent-samples *t*-test on the number of 340-trial bins to reach criterion for the True Transfer category and True Reversal category groups. There was no significant difference between the groups ($t_{25} = -.240$, $p = .812$, $d = -.096$, 95% CIs = -2.869, 2.269) in that True Transfer birds did not learn to transfer discrimination significantly faster than True Reversal birds learned to reverse discrimination. See Table 1.

Discrimination Training II vs. Discrimination Training I

To determine whether birds in the four True Transfer category groups differed in their speed of acquisition between Discrimination Training I and Discrimination Training II, we conducted a one-way ANOVA on the difference between the number of 340-trial bins in Discrimination Training II and Discrimination Training I to reach criterion (Discrimination Training II-Discrimination Training I; DIS2-DIS1) for the True Transfer category. There was a significant difference, $F_{3,11} = 4.187$, $p = .033$, $\eta^2 = .533$, 95% CIs = -1.964, 1.964. We conducted independent sample *t*-tests and applied Bonferroni corrections ($p = .05/6$ tests = .008). There were no significant differences between subgroups ($ps \geq .032$); see Table 1 and Figure 4.

Table 1*Training and Transfer/Reversal Results*

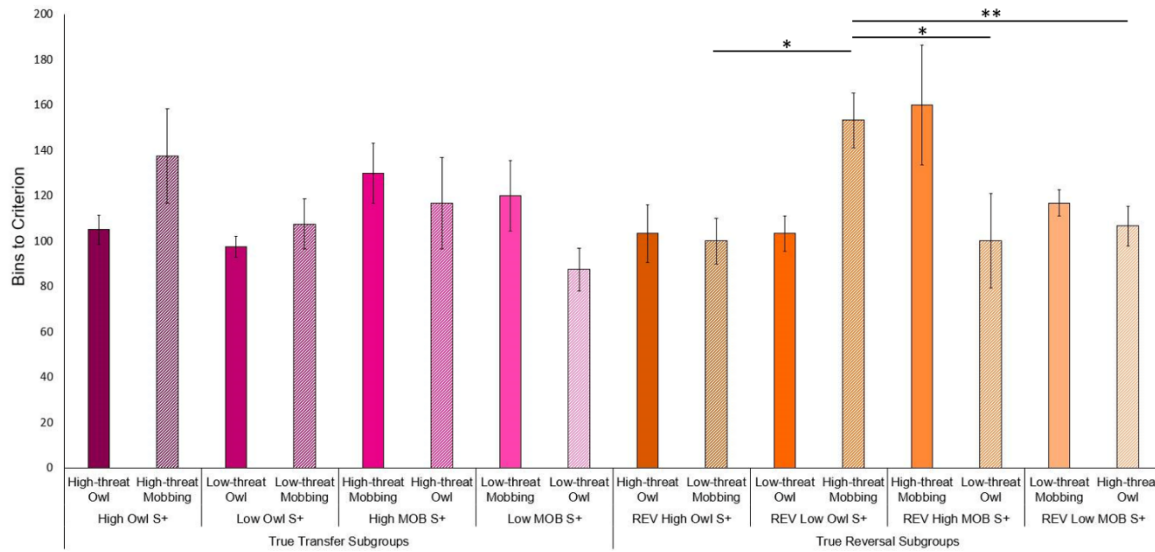
Subgroup Names	Average Bins to Criterion in Discrimination Training I	Average Bins to Criterion in Discrimination Training II	Difference Score in Average Bins to Criterion (DIS2-DIS1)
True Transfer Group ($n = 15$)			
High Owl S+	105 ± 6.5 (High-threat owl)	137.5 ± 20.9 (High-threat mobbing)	32.5
Low Owl S+	97.5 ± 4.8 (Low-threat owl)	107.5 ± 11.1 (Low-threat mobbing)	10
High MOB S+	130 ± 15.3 (High-threat mobbing)	116.7 ± 20.3 (High-threat owl)	-13.3
Low MOB S+	120 ± 13.5 (Low-threat mobbing)	87.5 ± 9.5 (Low-threat owl)	-32.5
True Reversal Group ($n = 12$)			
REV High Owl S+	103.3 ± 14.5 (High-threat owl)	100 ± 10 (Low-threat mobbing)	-3.3
REV Low Owl S+	103.3 ± 8.8 (Low-threat owl)	153.3 ± 12 (High-threat mobbing)	50
REV High MOB S+	160 ± 30.6 (High-threat mobbing)	100 ± 20.8 (Low-threat owl)	-60
REV Low MOB S+	116.7 ± 6.7 (Low-threat mobbing)	106.7 ± 8.8 (High-threat owl)	-10
Pseudo Category Group ($n = 6$)			
Pseudo 1 Owl	220 (High- and low-threat owl)	210 (High- and low-threat mobbing)	-10
Pseudo 2 Owl	425 ± 175 (High- and low-threat owl)	335 ± 12.5 (High- and low-threat mobbing)	-90
Pseudo 1 MOB	430 (High- and low-threat mobbing)	230 (High- and low-threat owl)	-200
Pseudo 2 MOB	240 ± 40 (High- and low-threat mobbing)	575 ± 135 (High- and low-threat owl)	335

Note. The average number of bins to criterion ± SEM during Discrimination Training I and Discrimination Training II, and the difference scores, for each subgroup of the True Transfer, Pseudo, and True Reversal groups. The stimuli discriminated for each subgroup in Discrimination Training I and Discrimination Training II is indicated below values in parentheses.

To determine whether birds in the four Pseudo category subgroups differed in their speed of acquisition, we conducted a one-way ANOVA on the difference between the number of 340-trial bins in Discrimination Training II and Discrimination Training I to reach criterion for the Pseudo category conditions. There was no significant difference ($F_{3,2} = 2.694$, $p = .282$, $\eta^2 = .802$, 95% CIs = -22.454, 31.787); see Table 1 and Figure 5.

Figure 4

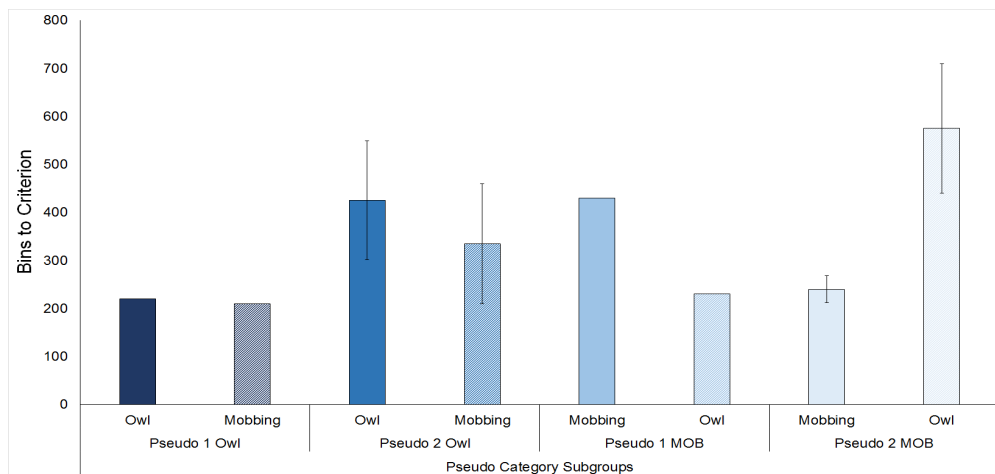
Discrimination Training II vs. Discrimination Training I: True Groups



Note. The average number of 340-trial bins to criterion \pm SEM during Discrimination Training I and Discrimination Training II (left and right bars of each subgroup, respectively) for each subgroup in True Transfer and True Reversal. The stimuli discriminated for each subgroup in Discrimination Training I and Discrimination Training II is indicated below each bar. There were no significant differences between True Transfer category subgroups (p s $>$.05/6 tests $>$.008; $n = 15$). There were significant differences, however, between True Reversal category subgroups as indicated by the * ($p = .004$) and ** ($p = .002$) on the histogram (all other p s $>$.008; $n = 12$).

Figure 5

Discrimination Training II vs. Discrimination Training I: Pseudo Category Subgroups



Note. The average number of 340-trial bins to criterion \pm SEM during Discrimination Training I and Discrimination Training II for each subgroup in the Pseudo Category (left and right bars of each subgroup, respectively). The stimuli discriminated for each subgroup in Discrimination Training I and Discrimination Training II is indicated below each bar. There were no significant differences ($p = .282$; $n = 6$). Missing error bars indicate no calculated

To determine whether birds in the four True Reversal category subgroups differed in their speed of acquisition, we conducted a one-way ANOVA on the difference between the number of 340-trial bins in

Discrimination Training II and Discrimination Training I to reach criterion for the True Reversal category conditions. There was a significant difference ($F_{3,8} = 19.703, p < .001, \eta^2 = .881, 95\% \text{ Cis} = -3.339, 2.172$). To further examine the significant ANOVA, we conducted independent samples t -tests and applied Bonferroni corrections ($p = .05/6 \text{ tests} = .008$). REV Low Owl S+ took significantly more bins to reverse their discrimination learning compared to all three of the other subgroups: REV Low MOB S+ ($t_4 = 7.348, p = .002, d = 7.378, 95\% \text{ Cis} = 3.733, 8.267$); REV High Owl S+ ($t_4 = -6.047, p = .004, d = -6.047, 95\% \text{ Cis} = -7.782, -2.885$); and REV High MOB S+ ($t_4 = -6.025, p = .004, d = -6.025, 95\% \text{ Cis} = -16.070, -5.931$); all other p s $\geq .038$. See Table 1 and Figure 4.

MOB S+ vs. Owl S+

Due to finding significant differences between the True Reversal subgroups that indicated owl stimuli may have been easier to learn to discriminate, we conducted an independent samples t -test on the difference between the number of 340-trial bins in Discrimination Training II and Discrimination Training I (DIS2-DIS1) for the True Transfer groups (MOB S+, Owl S+). There was a significant difference between DIS2-DIS1 for MOB S+ subgroups compared to Owl S+ subgroups ($t_{13} = -3.195, p = .007, d = -1.772, 95\% \text{ Cis} = -7.633, -1.475$). Specifically, both the MOB S+ and Owl S+ subgroups discriminated the owl stimuli in fewer trials than the mobbing stimuli. This result is not in line with our prediction as we predicted that Discrimination Training II would be completed in fewer trials than Discrimination Training I, regardless of whether stimuli were produced by heterospecifics or conspecifics.

To further investigate if this difference occurred between the Pseudo category subgroups, we conducted an independent samples t -test on the difference between the number of 340-trial bins in Discrimination Training II and Discrimination Training I for the Pseudo category groups (Pseudo MOB, Pseudo Owl). We found no significant difference between Pseudo MOB subgroups compared to Pseudo Owl subgroups ($t_4 = 1.054, p = .351, d = 1.054, 95\% \text{ Cis} = -35.944, 79.944$) in that both pseudorandomized owl and mobbing stimuli were discriminated in approximately the same number of trials.

To further confirm this difference between the True Reversal subgroups, we conducted an independent samples t -test on the difference between the number of 340-trial bins in Discrimination Training II and Discrimination Training I for the True Reversal groups (REV MOB S+, REV Owl S+). There was a significant difference between REV MOB S+ subgroups compared to REV Owl S+ subgroups ($t_{10} = -3.121, p = .011, d = -1.974, 95\% \text{ Cis} = -9.999, -1.668$) in that both the REV MOB S+ subgroups and the REV Owl S+ subgroup discriminated the owl stimuli in fewer trials than the mobbing stimuli.

Discussion

In the current study, we tested whether black-capped chickadees perceived two types of predator calls (high-threat NSW0 and low-threat GH0W) as similar to conspecific mobbing calls produced in response to each of these predators (i.e., MOB NSW0 and MOB GH0W, respectively) in order to investigate the referential information contained within chickadee mobbing calls. Despite the differences in threat posed by the animals that produced the acoustic stimuli, High S+ and Low S+ groups required approximately the same number of trials to successfully acquire the task. During Discrimination Training I and II, chickadees in the True Transfer and True Reversal groups learned to discriminate between rewarded and unrewarded stimuli significantly faster than chickadees in the Pseudo category group (Prediction 1 & 2), suggesting that true categories were easier to learn than pseudo categories. This result was expected as there is ample evidence that songbirds learn acoustic categories faster than they memorize similar acoustic stimuli not arranged into categories (Sturdy et al., 2000). During Discrimination Training I, the discrimination task was the same for chickadees in the True Transfer and True Reversal groups and, as expected, there was no significant difference in the number of trials to reach criteria between the two groups (Prediction 3). If mobbing calls contain referential information about the threat level posed by predators, chickadees in the True Transfer group would learn to transfer responding in fewer trials compared to chickadees in the True Reversal group during Discrimination Training II (Prediction 4). However, during

Discrimination Training II, the True Transfer group did not learn to transfer significantly faster than the True Reversal group. Therefore, in this study, we did not find any evidence of referential communication in black-capped chickadees.

Seyfarth et al. (1980) initiated the discussion of functionally referential alarm calls by suggesting that vervet monkeys produce specific alarm calls in response to multiple predators and respond distinctly to playback of each of those conspecific-produced alarm call types. Two and a half decades later, Templeton et al. (2005) demonstrated that chickadees' *chick-a-dee* call D-note production is negatively correlated with predator size (i.e., more D notes produced in response to small, high-threat hawks and owls). These results inspired Avey et al.'s (2011) study that found similar neural expression in chickadee auditory areas in response to hearing conspecific mobbing calls compared to hearing the predator's calls that the mobbing calls were produced in response to. This suggested that these two acoustically distinct vocalizations are perceptually similar, referring to the predator species and due to shared threat level. The notion that information regarding predator threat could be contained in the varied mobbing calls produced by chickadees would be evidence of referentiality in a songbird species. Exploring how referential signals are perceived by addressing this specific question required a novel expansion from Templeton et al.'s (2005) and Avey et al.'s (2011) studies, using an entirely different but well-established behavioral go/no-go operant paradigm. However, the current study did not provide evidence supporting that owl calls and mobbing calls produced in response to the same species of owls were perceptually similar. These findings do not provide support for referential communication as there are a number of factors that could contribute to these results. Additional perspectives to consider are discussed below.

Additional tests compared the difference in bins to criteria between Discrimination Training I and Discrimination Training II and indicated that there were significant differences in that both REV MOB S+ groups reversed to owl stimuli of opposite reward contingencies in fewer trials than REV Low Owl S+ reversed discrimination to high-threat mobbing stimuli. Thus, it appears as though categorizing the owl species' calls according to threat level may be easier than categorizing mobbing calls. This suggests that, despite the difference between high- and low-threat mobbing calls, chickadees were unlikely to be attending to the duration of the call for appropriate discrimination and categorization. Learning about the variations in production of mobbing calls produced by conspecifics is critical to survival as chickadees live in flocks and need to recognize a nearby threat and assist in antipredator mobbing behavior (Charrier et al., 2004; Templeton et al., 2005); however, chickadees generally took longer to properly discriminate mobbing calls, compared to owls' calls, according to threat level. Perhaps, due to the biological relevance of conspecific mobbing calls it would be more appropriate to initially respond with mobbing behavior rather than feeding; however, the required method of responding was approaching the feeder in the current task, which is consistent with mobbing behavior. In addition to subjects in this experiment passing Pretraining (i.e., responding at high rates to all stimuli), a previous study has shown that chickadees are capable of learning to approach operant feeders in response to high-arousal stimuli, including *chick-a-dee* mobbing calls produced in response to predators (Congdon et al., 2019). Congdon et al. investigated the perception of arousal by training chickadees to respond to high- or low-arousal stimuli, then testing with novel high- and low-arousal vocalizations and found strong transfer of training to black-capped chickadee stimuli. Conversely, the design of the current experiment was quite different; our intention was to focus on referential signals (rather than vocalizations of arousal), and we directly compared the rates of acquisition in discrimination between calls produced by conspecifics and calls produced by two heterospecific species. We determined that this type of operant conditioning design was critical to addressing our question of referentiality, as this design allowed us to standardize responses across a wide range of natural stimuli (i.e., conspecific vocalizations compared to heterospecific vocalizations) that otherwise might evoke vastly different responses, via examining transfer of training. However, we recommend that future experiments explore this concept using more than two species' calls (i.e., one high-threat, one low-threat predator) to test referentiality of individual species, and further separating generalization based on the concept of arousal level. In addition, we note the restriction of our subject size (i.e., 33 birds); when running a laboratory study it is critical to reduce the number of subjects included, but due to the number of subgroups, our null effects could be a direct result of a lack of power regarding effect size. Considering our recommendation of adding

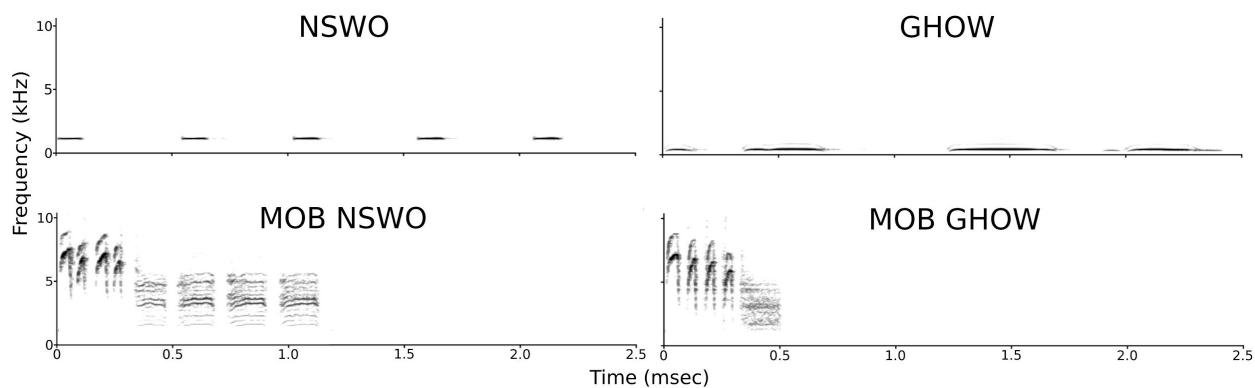
more stimuli, future experiments should also attempt to include more subjects and/or employ a within-subjects design regarding the pseudo groups.

It is evident that not all animals, solitary or social, require the use of referential communication as long as their communication system can allow the individuals to forage, reproduce, and survive. Originally it was thought that North American red squirrels (*Tamiasciurus hudsonicus*) a relatively solitary species, produced functionally referential alarm calls (i.e., *seet* and *bark* vocalizations in response to aerial and ground predators, respectively; Greene & Meagher, 1998), but more recent research has indicated that squirrels produce a *seet-bark* vocalizations to all predators, regardless of type (Digweed & Rendall, 2009). In contrast to Templeton et al.'s (2005) conclusions that more D notes per call are produced to high-threat predators compared to low-threat predators, Baker and Becker (2002) and Wilson and Mennill (2011) argued that the rate of calling and duty cycle (i.e., proportion of time filled by vocalizations) of the *chick-a-dee* mobbing calls is the element that indicates urgency tied to the level of posed threat. The mobbing calls produced to NSWO and GHOW (i.e., MOB NSWO and MOB GHOW), used as stimuli in the current study, were individual calls with varying D notes; due to the constraints of the design (i.e., chickadees remaining on the request perch for the entirety of the call), we were able to use only individual calls rather than strings of calls. It is possible that if we were able to train birds to mobbing call stimuli that varied in both note repetition and calling rate, there would have been transfer of contingencies from the True Transfer MOB S+ subgroups in Discrimination Training II. Alternatively, it is also possible that because the owl call stimuli used in this experiment were produced by two different species (high-threat NSWO vs. low-threat GHOW) that produce acoustically distinct calls, birds found it perceptually easier to categorize acoustically distinct vocalizations rather than *chick-a-dee* mobbing calls produced by the same species,

but which vary in D-note composition and average duration (see Figure 6). Additionally, predator-produced calls are likely more salient stimuli as owls prey upon chickadees; thus, an incorrect response is far more costly in face-to-face situations with a confirmed predator than whether to engage in a potential mobbing scenario with conspecifics.

Figure 6

Spectrograms of Experimental Stimuli



Note. Sample sound spectrograms of the vocalizations produced by northern saw-whet owls (NSWO), great horned owls (GHOW), and black-capped chickadees' mobbing in response to both owl predators (MOB NSWO and MOB GHOW), used as experimental stimuli with time (msec) on the *x*-axis and frequency (kHz) on the *y*-axis.

The purpose of this study was to explore how referential signals are perceived, focused on the investigation of referential communication in a species of songbird, as semantics are primarily studied in human and nonhuman primates (Seyfarth & Cheney, 1993). It has been suggested that a signal is referential if that vocalization contains variations that inform the receivers about environmental events, such as nearby predators (Evans, 1997). Maynard Smith and Harper (2003) suggested that animals that have the ability for

referential communication must: 1) be able to produce signals, 2) produce these signals under the correct circumstances, and 3) have receivers respond correctly. The findings provided by Baker and Becker (2002; call rate variation), Templeton and colleagues (2005; D note variation), and Wilson and Mennill (2011; duty cycle variation) indicate that *chick-a-dee* mobbing calls contain variations that inform the receiver about nearby predators and that receivers respond appropriately. The lack of findings in the current study suggest that chickadee mobbing calls may not be perceived to be signaling about a specific owl species, and thus, not parallel with the owl call stimuli, in the way that we anticipated (i.e., indicating specific predator or threat level). Theoretically, chickadee mobbing calls may still be referential according to the criteria suggested by Maynard Smith and Harper (2003) in that black-capped chickadees are (1) able to produce *chick-a-dee* mobbing calls, (2) produce these calls under the correct circumstances (i.e., in the presence of predators), and (3) have receivers respond correctly (i.e., becoming alert; participating in calling; actively mobbing the nearby threat). However, this experiment cannot make any definitive conclusions about referential communication.

Conclusion

To our knowledge, no other studies have examined referential signals and communication using a go/no-go procedure, as in the current experiment. We found that chickadees were able to discriminate the true categories of threat in fewer trials than pseudo categories, and True groups completed discriminations of acoustically distinct owl stimuli in the fewest number of trials. However, chickadees in the True Transfer group did not learn the second discrimination faster in Discrimination Training II compared to the True Reversal group in the way that we predicted. Thus, we propose that the current task (go/no-go discrimination) may not be an optimal experimental technique to detect whether mobbing calls are referential. For example, perhaps both True Transfer and True Reversal groups' 'true' categories are easy to discriminate, whether a transfer or reversal of threat contingencies is necessary (e.g., similar to mid-session reversal paradigms; Rayburn-Reeves & Cook, 2016). Future studies should continue this exploration of referential signals by further investigating the potential referential elements of chickadees' mobbing calls by considering additional stimuli and with varying duty cycles and/or alternative experimental designs, including a larger number of subjects; this may provide evidence to support, or fail to support, chickadees' use of referential communication, using an experimental design that may better address this question. Although this experiment was unsuccessful in providing substantial evidence, further studies will build a wealth of knowledge regarding referential communication in songbird species and the perception of referential signals generally.

Acknowledgements

We thank Isaac Lank, Phil May, and Al Denington for their technical assistance.

Author Contributions: **Jenna V. Congdon:** Conceptualization, Software, Investigation, Formal analysis, Writing – Original draft preparation **Allison H. Hahn:** Conceptualization, Investigation **Kimberley A. Campbell:** Investigation, Formal analysis, Writing – Reviewing & Editing **Erin N. Scully:** Investigation, Writing – Reviewing & Editing **Shannon K. Mischler:** Investigation **Carolina Montenegro:** Investigation **William D. Service:** Investigation **Christopher B. Sturdy:** Writing – Reviewing & Editing, Supervision

Funding: This research was supported by a Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant and Discovery Accelerator Supplement (RGPIN-2016-04591) funding at the University of Alberta (UofA) to CBS. JVC and KAC were each supported by an Alexander Graham Bell Canada Graduate Scholarship-Doctoral (NSERC CGS D).

Conflict of Interest: The authors declare no conflict of interest.

References

- Avey, M. T., Hoeschele, M., Moscicki, M. K., Bloomfield, L. L., & Sturdy, C. B. (2011). Neural correlates of threat perception: Neural equivalence of conspecific and heterospecific mobbing calls in learned. *PLoS ONE*, *6*, 1–7.
- Baker, M. C., & Becker, A. M. (2002). Mobbing calls of black-capped chickadees: Effects of urgency on call production. *The Wilson Journal of Ornithology*, *114*, 510–517.
- 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*, *83*, 820–833.
- 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*, *82*, 769–779.
- Congdon, J. V., Hahn, A. H., Filippi, P., Campbell, K. A., Hoang, J., Scully, E. N., Bowling, D. L., Reber, S. A., & Sturdy, C. B. (2019). Hear them roar: A comparison of black-capped chickadee (*Poecile atricapillus*) and human (*Homo sapiens*) perception of arousal in vocalizations across all classes of terrestrial vertebrates. *Journal of Comparative Psychology*, *133*, 1–22.
- Congdon, J. V., Hahn, A. H., Campbell, K. A., Scully, E. N., Yip, D. A., Bayne, E. M., & Sturdy, C. B. (2020a). Acoustic discrimination of predators by black-capped chickadees (*Poecile atricapillus*). *Animal Cognition*, *23*, 595–611.
- Congdon, J. V., Hahn, A. H., Campbell, K. A., Scully, E. N., Yip, D. A., Bayne, E. M., & Sturdy, C. B. (2020b). Can you hear me now? The effect of signal degradation on perceived predator threat in black-capped chickadees. *Animal Cognition*, *24*, 1–12.
- Digweed, S. M., & Rendall, D. (2009). Predator-associated vocalizations in North American red squirrels, *Tamiasciurus hudsonicus*: Are alarm calls predator specific? *Animal Behaviour*, *78*, 1135–1144.
- Evans, C. S., (1997). Referential signals. In D. H. Owings, M. D. Beecher, & N. S. Thompson, *Communication* (pp. 99–143). Springer.
- Evans, C. S., & Evans, L. (1999). Chicken food calls are functionally referential. *Animal Behaviour*, *58*, 307–319.
- Evans, C. S., Evans, L., & Marler, P. (1993). On the meaning of alarm calls: Functional reference in an avian vocal system. *Animal Behaviour*, *46*, 23–38.
- Fichtel, C. (2020). Monkey alarm calling: It ain't all referential, or is it. *Animal Behavior and Cognition*, *7*, 101–107.
- Gill, S. A., & Sealy, S. G. (2004). Functional reference in an alarm signal given during nest defence: Seet calls of yellow warblers denote brood-parasitic brown-headed cowbirds. *Behavioral Ecology and Sociobiology*, *56*, 71–80.
- Greene, E., & Meagher, T. (1998). Red squirrels, *Tamiasciurus hudsonicus*, produce predator-class specific alarm calls. *Animal Behaviour*, *55*, 511–518.
- 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., Hoang, J., McMillan, N., Campbell, K., Congdon, J., & Sturdy, C. B. (2015). Biological salience influences performance and acoustic mechanisms for the discrimination of male and female songs. *Animal Behaviour*, *104*, 213–228.
- Hahn, A. H., Hoeschele, M., Guillette, L. M., Hoang, J., McMillan, N., Congdon, J. V., ...Sturdy, C. B. (2016). Black-capped chickadees categorize songs based on features that vary geographically. *Animal Behaviour*, *112*, 93–104.
- Lawson, S. L., Enos, J. K., Mendes, N. C., Gill, S. A., & Hauber, M. E. (2021). Pairing status moderates both the production of and responses to anti-parasitic referential alarm calls in male yellow warblers. *Ethology*, 1–10.
- Lima, S. L., & Dill, L. M. (1990). Behavioural decisions made under the risk of predation: A review and prospectus. *Canadian Journal of Zoology*, *68*, 619–640.
- Maynard Smith, J. & Harper, D. (2003). *Animal signals*. Oxford University Press.
- McMillan, N., Hahn, A. H., Congdon, J. V., Campbell, K. A., Hoang, J., Scully, E. N., Spetch, M. L., & Sturdy, C. B. (2017). Chickadees discriminate contingency reversals presented consistently, but not frequently. *Animal Cognition*, *20*, 655–663.
- Melchior, H. R. (1971). Characteristics of arctic ground squirrel alarm calls. *Oecologia*, *7*, 184–190.
- Montenegro, C., Service, W. D., Campbell, K. A., Scully, E. N., Mischler, S. K., Congdon, J. V., & Sturdy, C. B. (2018). Behavioural response to female song playback in black-capped chickadees (*Poecile atricapillus*) [unpublished manuscript].

- Njegovan, M., Hilhorst, B., Ferguson, S., & Weisman, R. (1994). A motor-driven feeder for operant training in song birds. *Behavior Research Methods, Instruments, & Computers*, *26*, 26–27.
- Palya, W. L., & Walter, D. E. (2001). *Document set for the high-performance experiment controller* [unpublished manuscript]. Jacksonville State University.
- Pyle, P. (1997). *Identification guide to North American birds*. Slate Creek Press.
- Rayburn-Reeves, R. M., & Cook, R. G. (2016). The organization of behavior over time: Insights from mid-session reversal. *Comparative Cognition & Behavior Reviews*, *11*, 103–125.
- Riley, J. R., Greggers, U., Smith, A. D., Reynolds, D. R., & Menzel, R. (2005). The flight paths of honeybees recruited by the waggle dance. *Nature*, *435*, 205–207.
- Seyfarth, R. M., & Cheney, D. L. (1993). Meaning, reference, and intentionality in the natural vocalizations of monkeys. In H. L. Roitblat, L. M. Herman, & P. E. Nachtigall (Eds.), *Language and communication: Comparative perspectives* (pp. 195–219). Psychology Press.
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, *28*, 1070–1094.
- Soard, C. M., & Ritchison, G. (2009). ‘Chick-a-dee’ calls of Carolina chickadees convey information about degree of threat posed by avian predators. *Animal Behaviour*, *78*, 1447–1453.
- Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (2000). Call-note discriminations in black-capped chickadees (*Poecile atricapillus*). *Journal of Comparative Psychology*, *114*, 357–364.
- Sturdy, C. B., & Weisman, R. G. (2006). Rationale and methodology for testing auditory cognition in songbirds. *Behavioural Processes*, *72*, 265–272.
- Suzuki, T. N. (2015). Assessment of predation risk through referential communication in incubating birds. *Scientific Reports*, *5*, 1–6.
- Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: Black-capped chickadees encode information about predator size. *Science*, *308*, 1934–1937.
- Wilson, D. R., & Mennill, D. J. (2011). Duty cycle, not signal structure, explains conspecific and heterospecific responses to the calls of black-capped chickadees (*Poecile atricapillus*). *Behavioral Ecology*, *22*, 784–790.