University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

USDA National Wildlife Research Center - Staff Publications

U.S. Department of Agriculture: Animal and Plant Health Inspection Service

The Versatility of Graded Acoustic Measures in Classification of Predation Threats by the Tufted Titmouse *Baeolophus bicolor*: Exploring a Mixed Framework for Threat Communication

Kathryn E. Sieving
Department of Wildlife Ecology and Conservation, University of Florida

Stacia A. Hetrick
Department of Wildlife Ecology and Conservation, University of Florida

Michael L. Avery
USDA/APHIS/WS/NWRC Florida Field Station, michael.l.avery@aphis.usda.gov

Follow this and additional works at: https://digitalcommons.unl.edu/icwdm_usdanwrc

Part of the Environmental Sciences Commons

Research

The versatility of graded acoustic measures in classification of predation threats by the tufted titmouse *Baeolophus bicolor*: exploring a mixed framework for threat communication

Kathryn E. Sieving, Stacia A. Hetrick and Michael L. Avery

K. E. Sieving (chucao@ufl.edu) and S. A. Hetrick, Dept of Wildlife Ecology and Conservation, Univ. of Florida, Gainesville, FL 32611-0430, USA. Present address for SAH: UF/IFAS Osceola County Extension, 1921 Kissimmee Valley Lane, Kissimmee, FL 34744, USA. – M. L. Avery, USDA/APHIS/WS/NWRC Florida Field Station, 2830 E. University Ave, Gainesville, FL 32641, USA.

Many mammal and bird species respond to predator encounters with alarm vocalizations that generate risk-appropriate responses in listeners. Two conceptual frameworks are typically applied to the information encoded in alarm calls and to associated anti-predator behaviors. 'Functionally referential' alarm systems encode nominal classes or categories of risk in distinct call types that refer to distinct predation-risk situations. 'Risk-based' alarms encode graded or ranked threat-levels by varying the production patterns of the same call types as the urgency of predation threat changes. Recent work suggests that viewing alarm-response interactions as either referential or risk-based may oversimplify how animals use information in decision-making. Specifically, we explore whether graded alarm cues may be useful in classifying risks, supporting a referential decision-making framework. We presented predator (hawk, owl, cat, snake) and control treatments to captive adult tufted titmice *Baeolophus bicolor* and recorded their vocalizations, which included 'chick-a-dee' mobbing calls (composed of chick and D notes), 'seet' notes, two types of contact notes ('chip', 'chink'), and song. No single call type was uniquely associated with any treatment and the majority of acoustic measures varied significantly among treatments (46 of 60). The strongest models (ANOVA and classification tree analysis) grouped hawk with cat and owl, and control with snake, and were based on the number or proportion of a) chick and D notes per chick-a-dee call, b) chip versus chink notes produced following treatment exposure, and c) the frequency metrics of other note types. We conclude that (1) the predation-threat information available in complex titmouse alarm calls was largely encoded in graded acoustic measures that were (2) numerous and variable across treatments and (3) could be used singly or in combinations for either ranking or classification of threats. We call attention to the potential use of mixed threat identification strategies, where risk-based signal information may be used in referential decision-making contexts.

Referential and risk-based anti-predator alarm-call systems

Anti-predator alarm calls produced in response to predator detections are extremely common among vertebrate prey species (Caro 2005) and two general schemes of coding are used to characterize their information content. 'Referential' communication identifies classes or categories of risk (e.g. different predator species or attack styles) using different types (classes) of calls, whereas, 'risk-based' communication uses graded features of the same call type to convey potentially continuous variation in degrees of perceived predation risk. If alarm calls are referential and each type of call generates a unique anti-predator response by listeners, then the communication system is defined as functionally referential (Blumstein 1999). This is commonly observed in mammalian (primate, rodent) alarm-calling systems, when distinct call types are given to denote, for example, snakes versus eagles, and each call type elicits distinctive and situationally-appropriate escape responses (Seyfarth et al.

1980, Manser 2001, Kirchhof and Hammerschmidt 2006). Birds also produce functionally-referential alarm calls (Evans et al. 1993, Templeton and Greene 2007).

In contrast, species with more generalized anti-predator alarm calls or limited vocal complexity can use the same call type for various predator species or situations, but vary the acoustic structure or production pattern to denote different threat levels. Typically described as risk-based (or urgency- or affect-based, or emotional) alarm systems, calls denote the immediacy of predation threat represented by the situation the caller perceives (Arnold et al. 2008). Here, callers vary the rate of alarm call production or acoustic structure of calls, or both. Marmots vary call rate (Blumstein 1995) and yellow-bellied marmots *Marmota flaviventris* can change call bandwidth as a function of perceived threat (Blumstein and Armitage 1997). Mexican chickadees *Poecile sclateri* vary the pitch of a single kind of alarm call (Ficken 1989) and black-capped chickadees *Poecile atricapilla* alter both the call rate of chick-a-dee (predator mobbing) calls (Baker and Becker 2002) and the

number of notes per call (Templeton et al. 2005) according to the degree of threat perceived. Variation in immediacy, degree, or urgency in a predator encounter can be caused by various factors; proximity of the caller to the predator (Leavesley and Magrath 2005), size of a predator (as an indicator of how vulnerable the prey is to the predator; Templeton et al. 2005), or by factors internal to the caller affecting its fear level or perception of urgency (Baker and Becker 2002).

Birds in family Paridae (titmice and chickadees) have complex vocalizations that have been well-characterized (Otter 2007). Most parid species use the same, highly conserved alarm-calling system, comprised of risk-based predator-mobbing alarms (chick-a-dee calls) and distinct 'seet' alarm calls (Langham et al. 2006, Templeton and Greene 2007). Mobbing calls are given to perched avian predators and other terrestrial predator species, and they attract other birds (same and different species), generating noisy, aggressive mobbing aggregations (Sieving et al. 2004). In contrast, 'seet' calls are often given when flying/attacking hawks are detected and they elicit distinct escape responses (diving into cover, flight, or freezing; Lima 1997, Hetrick 2006, Templeton and Greene 2007). Therefore, the parid alarm-calling system (like that of other vertebrates; Fichtel and Kappeler 2002) includes both risk-based and referential communication in that they can vary call structure and (or) the call types produced under different anti-predator situations.

Potential for a mixed framework of threat communication

Clear cases of one style of communication versus the other being deployed are common in the animal literature (Evans et al. 1993, Templeton et al. 2005). However, the lines between referential and risk-based communication do not often remain distinct across different threat contexts (Arnold and Zuberbühler 2005); both graded variations in acoustic structure and qualitative call type production can vary with 1) distinct situations (classes of threats) or 2) similar kinds of situations with gradations in immediacy of threat. We align with Seyfarth and Cheney (2003) and Arnold et al. (2008) in questioning whether the simple dichotomy between referential and risk-based communication creates a biased and limited view of animal signaling. Few have considered the possibility that multiple (graded or referential) acoustic cues may be used simultaneously to discriminate among risk situations in a referential manner (Hetrick 2006), and we perceive that this could limit views of the complexity of vertebrate communication systems. Animals with sophisticated perception could base appropriate types of responses on just one spectral measure of alarm calls or on several metrics at once, and choices may be context dependent (Muller and Manser 2008). Receivers and eavesdroppers may simultaneously classify and rank relevant threats so that they can fine-tune their abilities to control the risk of predation based on information extracted from others' alarm calls (Seyfarth and Cheney 2003). Graded acoustic cues are noted more commonly than referential ones across the majority of alarm-call studies, and these cues are readily shared (via eavesdropping) between

species (Langham et al. 2006, Schmidt et al. 2008). Therefore, graded cues probably comprise the majority of social information about predation threats; yet anti-predator behaviors are not simple gradations of the same type of behavior (Caro 2005).

Our objective was to assess the possibility that titmouse anti-predator vocalizations could be used to support a mixed communication system; specifically, referential classification of risks based on graded signal content. Without seeking to establish meaning of signals produced by experimental birds, our purpose was to explore the information content of alarm calls and the degree that it could support either (or both) linear rankings or categorical classification of predation risks by competent listeners.

Methods

Study design

We conducted experimental presentations of distinctly different predator species to captive tufted titmice *Baeolophus bicolor*, recorded titmouse vocal responses, quantified various acoustic measurements from sonograms, and assessed the efficacy of both ranking and classifying treatments in analyses. All predators were presented in close proximity to individual titmice to enhance visibility of the different species (or the empty cage in the control presentation), and were all presented in the same manner (in a perched position). This design was selected to encourage titmice to use vocal cue sets that would uniquely distinguish the different species, un-confounded by predator activity (attacking vs foraging) or location (flying vs perched). We expected that both graded and referential signal content would be high since titmice vary their alarm call structure in the same manner as chickadees (Templeton et al. 2005, Hetrick 2006, Templeton and Greene 2007).

Predator treatments included a live owl, cat, snake, a live-mounted (stuffed) hawk, and an empty cage (control). We selected representative species with different propensities to capture and consume small passerines that could support rankings of threat, but that varied in ways that could also support classification into several nominal categories of predation threat. For example, titmice could rank animals by body size (large, medium and small, Templeton et al. 2005), or they could classify perched animals as typically aerial- versus terrestrial-hunters, or avian, mammalian, versus reptilian predators. Therefore, the experimental design presented opportunities for both referential (e.g. different species or classes of predators) and urgency-based (e.g. predator size, attack speed) cues to be produced for use in identifying treatment differences. If a mixed alarm communication system is possible, i.e. if graded acoustic measures can be used effectively to classify predator treatments, then we predicted that information encoded in signals would support logical schemes for both categorization and ranking of treatments in analysis.

Study species

Tufted titmice are 'nuclear' species around which mixed-species foraging flocks form during the winter months in

North America (Greenberg 2000, Farley et al. 2008). Similar to other nuclear species, tufted titmice are aggressive, vocally-complex, and maintain high levels of vigilance (Munn and Terborgh 1979). Tufted titmice are active sentinels, alerting flocks to predators in the environment (Gaddis 1983, Dolby and Grubb 1998), and they exhibit both types of alarm calls widely observed across the family Paridae: the 'seet' alarm call (also known as the high whistle, see-see-see, 'hawk' call, and 'flying predator' call) and the mobbing or scold call known generally as the 'chick-a-dee' call (Marler 1955, Gaddis 1980). Titmouse mobbing calls are variants of the typical 'chick-a-dee' call, composed of introductory 'chick' notes, and subsequent D notes (dee notes, churr notes), with the number and presence of each note type being variable (Latimer 1977, Hailman 1989, Hetrick 2006, Otter 2007, Fig. 1). Tufted titmouse mobbing calls elicit large mobbing flocks, attracting more species and individuals than the calls of other sympatric passerines (Sieving et al. 2004), and their seet alarm calls result in the cessation of movement (freezing) by the caller and nearby birds, or, rapid escape to cover (Ficken 1989, Baker and Becker 2002).

Parids encode a variety of risk-related information in their alarm calls that allow discrimination among risk situations (Smith 1972, Hailman et al. 1985, Ficken et al. 1994). Many species of birds (and even chipmunks) eavesdrop on tufted titmouse (and other parid) alarm calls and respond in ways that are appropriate to situations in which the calls are given (Hetrick 2006, Langham et al. 2006, Templeton and Greene 2007, Schmidt et al. 2008). Moreover, parid alarm calls are reliably associated with detection of the risks they convey; they are rarely deceptive (Maynard-Smith and Harper 2004). Therefore, information encoded in titmouse calls is rich, utilizes both risk-based and referential signals, and reliably conveys risk.

In north–central Florida, the most common known predators of adult forest passerines (including titmice) are sharp-shinned *Accipitor striatus* (winter only) and Cooper's hawks *A. cooperii*; eastern screech-owls *Megascops asio*; red-shouldered *Buteo lineatus* and red-tailed hawks *B. jamaicensis*; American kestrels *Falco sparverius*; and house cats (Bent 1937, Morse 1973, Sieving et al. 2004). Snakes, most commonly rat snakes *Elaphe* spp., typically prey on the eggs, nestlings and occasionally adults of small birds during the summer months (Halliday and Adler 1986). We selected predators for use in this study based on this list, authors' observations in the field, and on the availability of tame individuals to use in an experimental setting.
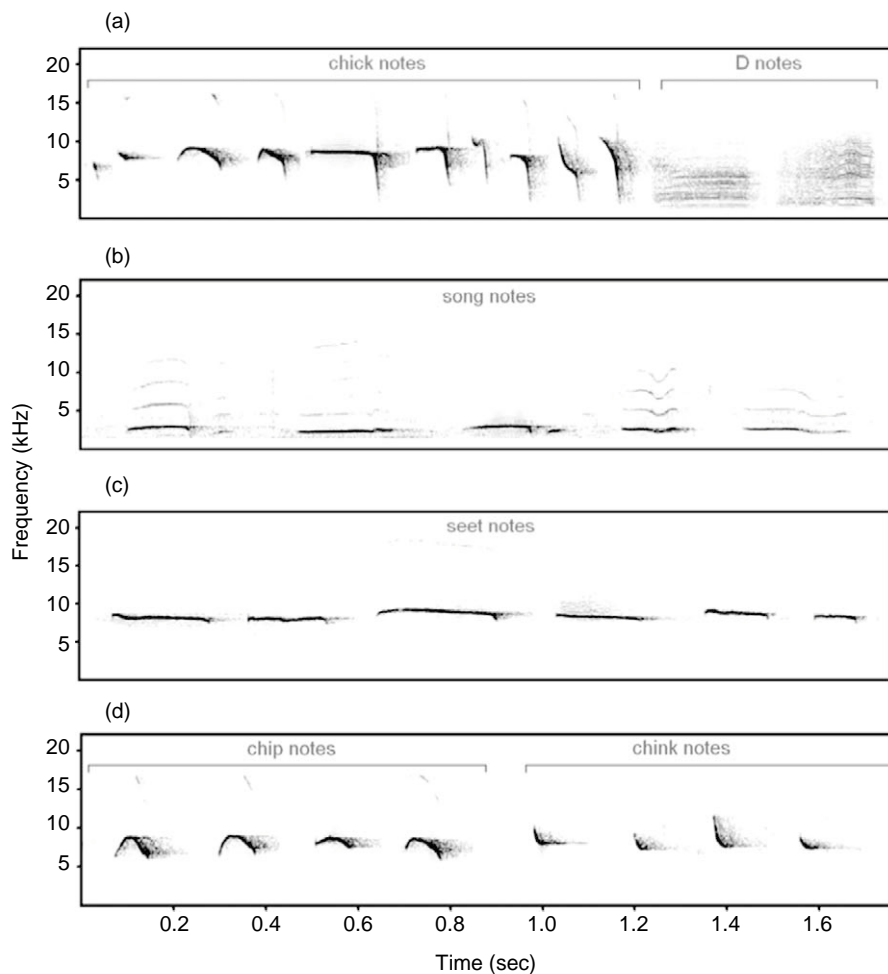

Figure 1. Examples of the main vocalizations of tufted titmice in response to presentations: (a) 'chick-a-dee' call complex showing the variation in introductory chick notes and subsequent D notes, and variations of (b) song, (c) seet, (d) chip and chink notes.

Titmouse capture and handling

During November 2004 – January 2005, we caught 15 adult titmice with mist-nets and wire cage traps on feeders at locations in north–central Florida: the Ordway Biological Station (Putnam Co.), the USDA/Wildlife Services Florida Field Station of the National Wildlife Research Center (Gainesville, FL), and various residences in the city of Gainesville (see Hetrick 2006 for titmouse housing and feeding information). At least 24 h before testing, the caged titmouse was moved to an outdoor aviary ($9 \times 3 \times 2.3$ m) for acclimation and observation – any individuals not feeding normally were released without testing. The aviary contained numerous branches and was in the shade of forest edge, providing a semi-natural environment. Both cages (for titmice and predators) were mounted on sturdy posts within the aviary (0.5 m high and 0.75 m apart). The predator presentation cage was clear plexi-glass and was covered by a sheet except during testing (Hetrick 2006). A camouflage blind for observers and recording equipment was 4.5–6 m away from the titmouse and predator cages.

Predator presentations

Each individual bird underwent a series of five treatments. We presented in random order, a stuffed sharp-shinned hawk, a live eastern screech-owl, a live domestic house cat *Felis domesticus*, a live red rat snake *Elaphe guttata*, and an empty cage (hereafter referred to as hawk, owl, cat, snake and control, respectively). A live *Accipiter* tame enough to tolerate the testing was unavailable so we used a live-mounted specimen in a natural perched pose with realistic eyes (obtained from the Florida Museum of Natural History). A tame screech owl was borrowed from a local wildlife rehabilitation center, and the other two animals were borrowed pets. Tests began at 08:00 and continued at 2-h intervals until 16:00. Each predator was placed into the predator cage 20 min before testing and titmice did not view the predator until the sheet was pulled off using a rope and pulley operated from the blind. Acoustic recordings were then made for 5 min using a shotgun microphone, a laptop computer, and Raven Interactive Sound Analysis Software ver. 1.1 with a sampling rate of 44100 at 16-bit resolution. We presented single titmice with the predator treatments to insure precise recordings of individual responses.

Acoustic response measures

Seventy-four out of 75 recordings (15 titmice × 5 treatments) were analyzed (one hawk recording was lost to equipment failure). Spectrographic analyses were performed using Avisoft SASLabPro 4.39. To edit out noise, each sound file was FIR low-pass filtered at 12 kHz and high-pass filtered at 1.4 kHz; parameters used included FFT = 512, frame size = 75%, window = hamming, overlap = 87.5%. We classified and labeled all vocalizations, which included chick-a-dee calls, songs, seet notes, and contact notes (two kinds; below). The note types in chick-a-dee calls were visually classified as introductory chick notes or subsequent D notes. In titmice, the various introductory notes grade into each other and are not easily distinguished

into natural categories; therefore, the introductory notes were classified together as chick notes (Fig. 1a). On the other hand, D notes can be reliably classified (Bloomfield et al. 2005) based on their harmonic-like structure, small frequency modulation, high entropy, low frequency, and position (always occurring at the end of a call or were the only notes comprising a call; Fig. 1a). Songs are unique and can be easily distinguished from other vocalizations (Fig. 1b). Seet notes were recognized by their high pitch ($\sim$ 8–10 kHz), narrow bandwidth, lack of sharp onset or ending, and a narrow frequency range (Fig. 1c; Marler 1955). Contact calls are single-syllable notes of two distinct types; we called them chip and chink notes. Chinks are L-shaped, whereas chip notes are inverted U-shaped, and both are shorter duration than seet notes (Fig. 1d).

We obtained gross and fine measures, representing standard descriptors of alarm vocalizations (Baker and Becker 2002, Freeberg et al. 2003, Templeton et al. 2005). Gross measures included the no. of each type of note and call (chick, D, song, seet, chip, chink) given during the 1st min and over the entire 5 min period following predator presentation, and mean duration (seconds) of notes given during the 5 min recording period. For chick-a-dee calls, we averaged the number and proportion of chick and D notes per call (5 min periods). We also quantified the proportion of chip versus chink notes produced during the 5 min sample period. Fine-scale measures were taken (using a power spectrum; FFT length = 512) on all notes produced in the 5 min period not overlapping with background noise, representing in most cases a majority of notes produced (see Table 1 for summary of note production). For each note type included in analysis, the following mean parameters were computed at the maximum spectrum of the entire note (max peak hold; after Nowicki 1989, Templeton et al. 2005); a) bandwidth (Hz), b) minimum and c) maximum frequency 1 (where the amplitude goes last below -30dB) and frequency 2 (where the amplitude goes last below -10dB; a–c measured in Hz and calculated at min and max frequency with the total option activated in Avisoft SASLabPro 4.39), d) distance to maximum frequency 1 (s), and e) entropy (a measure of sound purity that varies from 0, for a pure tone, to 1, for random noise).

Data analysis

We conducted a preliminary screening of data for effects of presentation order or individual bird on call production, using classification tree analysis (CTA). We used presentation order (1–5), bird ID (1–15), and treatment categories (cat, hawk, owl, snake, control) as predictors of presence or absence (P/A) of each note or call type produced during 5 min trial periods, and constructed models for P/A of all six note types, chick-a-dee calls and silence (when no, or fewer than five notes of any type were produced in 5 min). If presentation order or individual bird was the dominant or only predictor of the presence of a note type, then we did not include that note type in subsequent analyses.

To assess the potential risk-based versus referential information content of alarm calls, we analyzed vocalizations in two ways. For all gross and fine-scale vocalization

Table 1. Total and mean number of each note type produced in response to treatments in the first 5 min following presentations. Number of subjects producing notes of each type in parentheses.

Note type		Treatment presentation				
		Control	Snake	Owl	Hawk	Cat
Chip	n	961 (13)	1101 (14)	389 (11)	383 (9)	92 (5)
	mean	64.1	73.4	25.9	27.4	6.1
	SE	15.2	20.2	8.0	10.2	4.9
Chink	n	104 (10)	36 (4)	331 (9)	426 (11)	356 (12)
	mean	6.9	2.4	22.1	30.4	23.7
	SE	3.0	1.1	9.7	10.3	9.1
Chick	n	44 (4)	38 (3)	24 (4)	21 (4)	7 (3)
	mean	2.9	2.5	1.6	1.5	0.5
	SE	1.9	2.2	1.3	1.0	0.3
D	n	11 (2)	21 (3)	34 (7)	88 (8)	141 (5)
	mean	0.7	1.4	2.3	6.3	9.4
	SE	0.6	1.1	0.9	2.7	5.5
Seet	n	43 (4)	9 (2)	27 (5)	53 (9)	54 (7)
	mean	2.9	0.6	1.8	3.8	3.6
	SE	1.5	0.5	0.9	2.0	1.7
Song	n	4 (1)	3 (1)	41 (3)	36 (2)	50 (1)
	mean	0.3	0.2	2.7	2.6	3.3
	SE	0.3	0.2	2.2	2.4	3.3

parameters measured above (60 total), we used one-way analysis of variance (ANOVA) to test for treatment affects. If tests were significant, then we used least significant difference (LSD) post-hoc pair wise comparisons to rank treatments and groups of treatments. We transformed the data when appropriate to meet assumptions of analyses using either angular (arcsin(sqrt(n)) or log(n + 1) transformations. We used SPSS ver. 11.5 and 12.0 for Windows and alpha-level = 0.05.

We used classification tree analysis (DTREG; Sherrod 2004) to explore how potential listeners could use one or more acoustic cues in titmouse vocalizations to distinguish among the five different treatments. We constructed 10 different models varying in combinations of treatments (in two to five classes) according to different ecological assumptions about perceptions of the five treatments. Three models used nominal categorization of the five treatments (models species, taxon, hunting habit) assuming differentiation without gradation of threat; the other seven models ordered treatment classes into explicit gradations of threat (Appendix 2). We submitted all 46 variables that were significant in ANOVA to CTA, with eight additional categorical variables (P/A of chip, chink, seet, and D notes in the 1st min and in the 5 min sampling period).

CTA uses predictor variables to repeatedly split the data set into increasingly homogenous and mutually exclusive groups of cases of a categorical target variable (in this case, predator classes). Each split of the target variable uses only the single most influential variable at each node that maximizes the homogeneity within the resulting two groups (Wiles and Brodahl 2004); this minimizes the impact of collinearity among predictor variables. CTA is well-suited for diverse measures because it lacks assumptions about underlying distributions. Cross-validation (the process of repeatedly using one data subset to predict another subset) assesses the predictive value of the model as it is built. Target variable homogeneity (or impurity) within each node is expressed as misclassification rate (MC); the smaller the MC, the higher the implied predictive power. Therefore, if terminal nodes in the model have relatively low MC

values representing a significant portion of cases, new cases that fit all of the preceding node criteria have a high likelihood of being accurately classified. In this way CTA mimics step-wise cognitive processes that support discrimination among groups of observations such as the use of dichotomous keys, machine learning, and emergency room triage (Breiman et al. 1984). For these reasons, CTA was better suited for our objectives than global modeling approaches (e.g. principal components, discriminant analyses) that, both, abstract the relationships of multiple collinear variables and weigh the influences of all variables simultaneously over their ranges. Both processes are challenging cognitive procedures, considering the number of measures we analyzed (>50), and may not reflect real-time animal decision-making.

We pruned classification trees to the minimum cross-validated error using Gini node splitting criterion, and 10-fold cross-validation (training subsets of 90% were repeatedly used to predict 10% of the data). Model outputs we used included variable importance (the degree to which each variable reduces overall node impurity, scaled from 0 to 100%), gain (whether reliable classifications can be made before all cases have been processed; assessed for each treatment grouping where 50% of data were processed), and confusion matrices (Breiman et al. 1984, Farley et al. 2008). Confusion matrices yield measures of model performance including receiver–operator characteristics (ROC plots of true vs false positive identifications of target classes), accuracy (total correct classifications/all correct and incorrect classifications) and precision (true positive/true and false positive rates). We compared models using all of the above, in decreasing order of importance: ROC plots; accuracy and precision rates; gain; and the total number of variables used in each model (fewer is better). Strong models that readily support classification or ranking of treatment groupings suggest that animals sensitive to the acoustic measures used in model construction could efficiently classify predation threats (and learn to do so) that were encoded in titmouse calls, in the ways consistent with model assumptions.

Results

Titmice typically gave a combination of call types in response to all predator and control treatments, usually consisting of many chip or chink notes and fewer chick-a-dee and seet calls; song and complete silence were infrequent (Table 1). An association between song production and presentation order suggested singing was associated with captivity length or time of day, not treatment effects. Despite an apparent increase in song with treatment level threat (Table 1), this is apparently an artifact because only four individuals produced song, only one produced 67% of all songs (88/134), and 50 of those were produced in front of the cat. While these results suggest that threat or fear may cause song production in some individuals, the use of songs in a threat context has not been addressed and we do not know how to interpret their use in this study. Therefore songs were not included in analysis. Chip and chink notes were the most common vocalization; 68 of 74 recordings contained chip notes. Less than a quarter of birds responded with chick-a-dee calls to the control, snake and cat treatments, and about half did so to the owl and hawk. More titmice gave seet calls in response to the hawk and cat than to the control, snake and owl treatments.

Ranking of predator presentations

In ANOVA models with LSD comparisons, fifteen of 25 gross spectrographic variables and 31 of 35 fine-scale acoustic measures varied significantly with treatment ($p < 0.05$). Of those 46 models, 28 had adjusted R^2-values > 0.10 (Appendix 1). Several measures varied across treatments, including no. notes given during the first minute, note duration and entropy. In multiple comparisons tests, the clearest and most frequent groupings based on individual variables ranked cat before hawk before owl versus a snake–control group, or included cat with hawk (and sometimes the owl) versus a group with the snake and control (and sometimes the owl). Both gross and fine spectral measures produced the latter ranking of two groups, but the proportion of D (or chick) notes per chick-a-dee call produced the strongest model. It was the only variable to separate cat, hawk and owl from the snake–control group and provided a near linear ranking of the five treatments and a complete ranking of the four animals. The precision of ranking using this variable is due to the smooth gradation of chick and D notes per chick-a-dee call across treatments; the control and snake treatments elicited significantly more chick notes per call and fewer D notes per call than the owl, cat and hawk, and the cat elicited the fewest chick and most D notes, followed by the hawk and owl, respectively (Fig. 2, top).

Two other strong ANOVA models distinguished cat–hawk–owl versus snake–control groups, but based on the proportion of chink and chip notes produced in the 1st min and 5 min following predator presentation (Fig. 2, bottom). Other strong models reflecting this same basic pattern used spectral features of the seet (max freq. 1 and 2, min freq. 2) and D notes (max freq. 2 and entropy; Appendix 1). A few spectral measures singled out predators from the other treatments. Chip note min. freq. 1 was much lower for the owl than any other treatment (Appendix 1, Fig. 3A) and
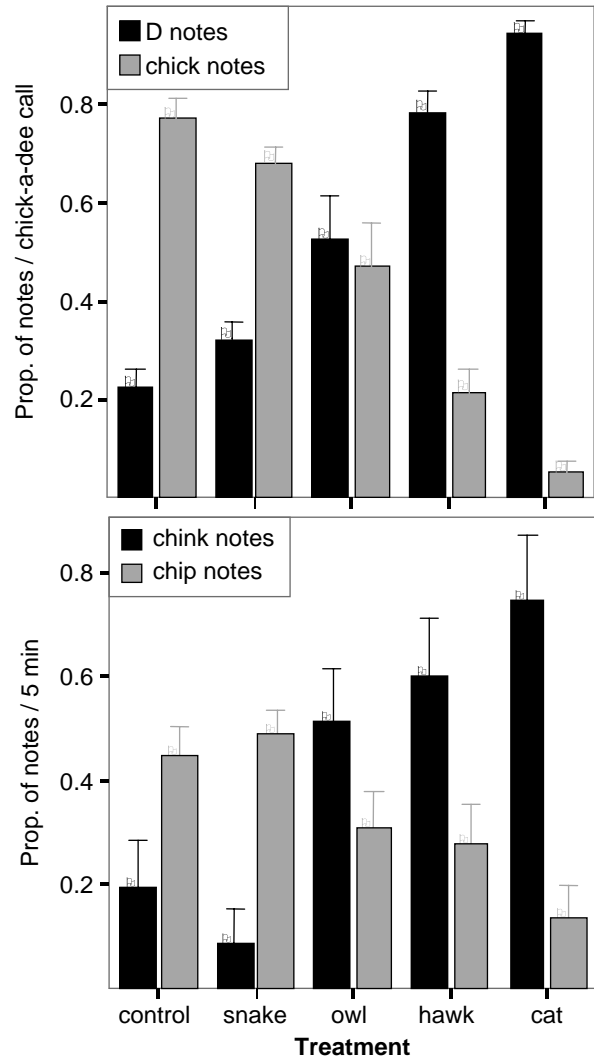

Figure 2. Mean proportion of chick versus D notes per 'chick-a-dee' call (top) and of chip versus chink notes per 5 min period (bottom) given in response to predator treatment presentations. See Appendix 1 for ANOVA/LSD results indicating significant differences among treatments using these measures. Bars: ±1 SE.

distance to maximum frequency of the seet notes was significantly shorter for the cat than the other four treatments (Fig. 3B). Finally we note an interesting contrast between the use of D notes and seet notes. The best-fit ANOVA model shows that seet note duration was longest for low risk (control) and shortest for the cat, whereas, D note duration varied the opposite way – longer for the high risk hawk, and shorter for the low risk snake (Appendix 1). Therefore, ample signal strength and diversity was available to support either ranking or classification of treatments.

Ranking versus classification of predator groups

We fit classification tree models to ten different treatment groupings that assumed different criteria for separating the groups. Three models assumed nominal classifications of treatments (species, taxon, hunting habit) and the other seven assumed ordinal rankings of treatment groups based on biological attributes of the predator species (size, % birds
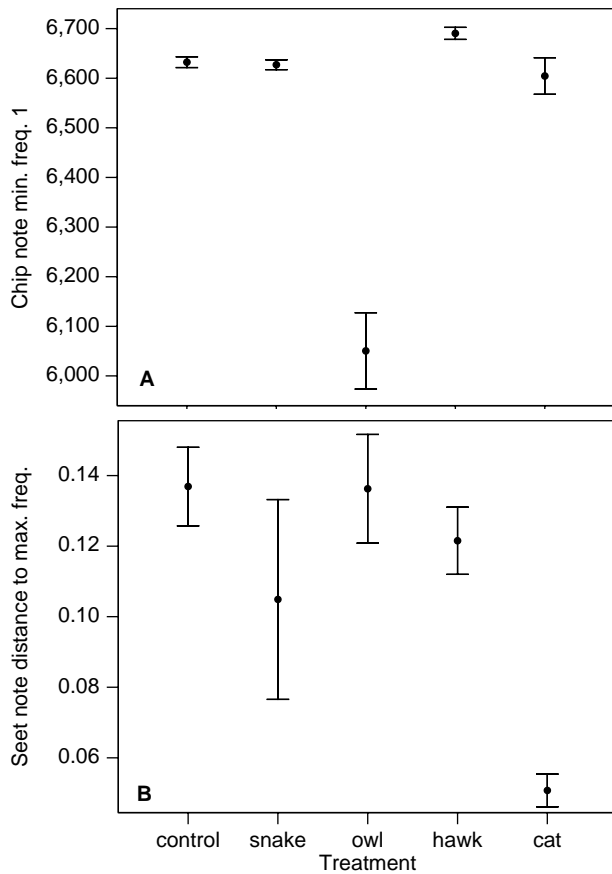
Figure 3. (A) mean of min. frequency 1 (KHz; where amplitude goes last below −30 dB) of chip notes and (B) mean distance (sec) to max. frequency of seet notes produced during 5 min following treatment presentations. See Appendix 1 for ANOVA/LSD results indicating significant differences among treatments using these measures. Bars: ±1 SE.

in diet, predator, risk1–4: Appendix 2). All models except species grouped two or more treatments together into classes (see Appendix 2 for underlying assumptions used). Either the ratio of chink relative to chip notes produced or the total number or presence of chip notes had the highest importance values in every model. Secondarily, note frequency, duration, and entropy measures were influential in several models.

The strongest model (risk1; Fig. 4) organized treatments into three classes – (hawk, cat) versus (owl) versus (snake, control). This model assumed ordinal ranks based on assumptions about the potential for attack at close range (high, medium, low threat, respectively). Identification of the high threat group versus the other two was accomplished with the highest ratio of true positive to false positive classifications (Fig. 5), and with second highest overall accuracy and precision across both training and validation data sets (Fig. 6). This model also had the highest overall gain (94%) but also the largest total number of variables (Appendix 2); its groupings were consistent with the best ANOVA models that placed cat, hawk and owl versus control and snake group (Appendix 1). The next strongest CTA models, risk3 (two groups: cat, hawk, owl vs snake, control) and size (four groups: cat vs hawk, snake vs owl vs control) had excellent receiver–operator reliability

(Fig. 5) and excellent precision and accuracy (Fig. 6) regarding distinction of the highest threat groups (c, h, o and c, respectively) from the others (Appendix 2). Risk3 included only six variables (vs 22 by risk1) and had the highest overall accuracy and precision (Fig. 6); therefore, this model may be the most realistic (easiest for diverse prey species to use). The best model assuming purely nominal groupings was taxon (four groups: birds vs mammal vs reptile vs control), but its performance was intermediate by all measures (Appendix 2). The other two nominally categorized models (species, hunting habit) were among the worst by all measures (Fig. 5, 6, Appendix 2).

Discussion

Encoding of risk in titmouse alarm calls

Chick-a-dee calls are among the best characterized of animal signals (Otter 2007). The basic structure and function of these calls is conserved across the Paridae family (Langham et al. 2006); they are used most commonly in a social signaling context (to communicate perceived threat to others) and they induce mobbing behavior by multiple species. It has long been known that greater numbers of more strident mobbing calls are given with increased levels of threat posed by different predation-risk situations (Latimer 1977). For both tufted titmice (Hetrick 2006) and a related species (black-capped chickadees) it was recently established that the pattern of chick and D note production variation within chick-a-dee calls is a direct measure of perceived predation threat (more D notes = higher, more chick notes = lower; Templeton et al. 2005). Therefore, since parid alarm calls are functionally and structurally conserved across the Paridae (Langham et al. 2006), and chick-a-dee call structure in titmice varies similarly to chickadees in relation to threat level (Hetrick 2006), we conclude that titmice in this experiment perceived a gradation of threat level in the following way; cat > hawk > owl > (snake = control; Appendix 1). This finding gave us a solid reference for interpreting analyses.

Surprisingly, the contact call structure (chinks vs chips) mimicked the alarm call structure closely, by separating a high risk (cat–hawk–owl) from a low risk group (control–snake; Fig. 2; Appendix 1) based on the proportion of chink to chip notes. The existence of a risk signaling system based on contact calls that parallels the information in alarm calls is significant in several ways. First, the flow of risk-specific information coming from highly vocal titmice is not limited to the times when they are mobbing predators or signaling attacks, but may also include the many hours/day when contact calls are the principal vocalizations produced. This suggests that the rich diversity of vertebrates that utilize parid alarm calls in anti-predator decision-making (Hetrick 2006, Langham et al. 2006, Templeton and Greene 2007, Schmidt et al. 2008) might also obtain threat-relevant information from their contact calls. Parids are widespread social dominants in the Holarctic (Harrap and Quinn 1995) that influence movement (Sieving et al. 2004, Otter 2007), settlement patterns (Mönkkönen and Forsman 2002), foraging (Dolby and Grubb 1998), and possibly reproductive success (Seppänen et al. 2005) of heterospecifics that
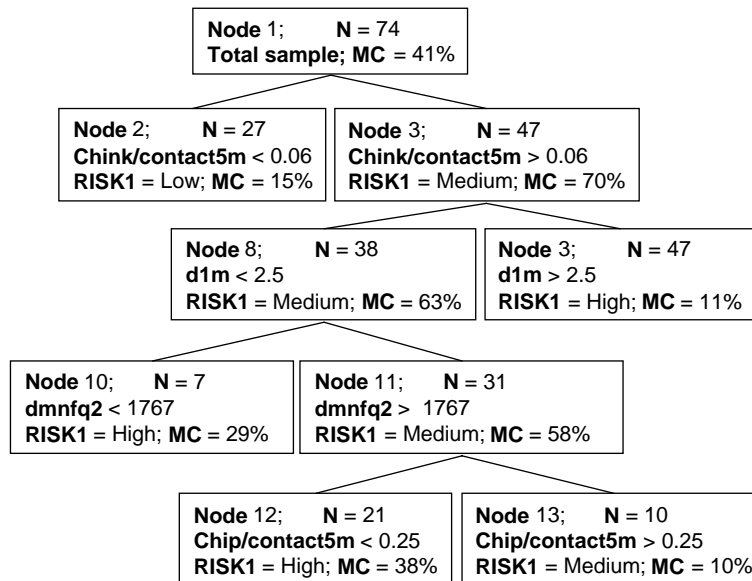
Figure 4. RISK4, the strongest classification tree model according to receiver operator characteristics (Fig. 5). Nodes with misclassification rates (MC) of 15% or less and sample sizes of 10 or greater carry high predictive power, and all nodes above a terminal node contribute to the classification of cases in that node. Here, the owl (medium threat) was distinguished from high threat species (cat and hawk) by eliciting all of the following; a higher ratio of chip to chink notes (node 13), D notes with higher min. freq. 2 (node 11), and fewer D notes in the first minute (node 8). Medium was distinguished from low (snake and control) by eliciting fewer chink to chip notes during the 5 min response period (node 3).

share predators with them. If contact calls, the most common vocalizations of parids, express similar detail about risk as their social alarm calls, then information transfer becomes an even more compelling hypothesis to explain their pervasive social dominance and positive influences on heterospecifics.

A second implication of this finding is that species with less vocal complexity than titmice may be communicating risk, or other information, with a greater accuracy, precision and sophistication than previously thought. For example, most bird species have two main vocalizations – songs and contact calls (Zeigler and Marler 2004) – and in all contexts other than sexual selection and territory advertisement, calls are typically used. Our findings suggest that most songbird species, therefore, could be capable of conveying sophisticated precision in threat perception using contact calls alone. For example, Langham et al. (2006) found that the single-note calls of both vireos and wrens were nearly as strong in generating predator-mobbing as parid 'chick-a-dee' calls. Thus, contact calls in general may provide more than just location information and, across species, may represent a near constant flow of information conveying predation threat perception within bird communities.

We detected both, significant redundancy in signal content in titmouse vocalizations, including parallel contact and alarm call encoding, and strikingly distinctive patterns that identified unique treatment groupings. Measures varying in parallel with acoustic patterns in alarm and contact calls included no. chip notes/5 min, frequency characteristics of the D and seet calls, and entropy of chip, chick and D notes (Appendix 1). The more dangerous cat, hawk, and owl elicited higher note entropy than the snake and control treatments, max. frequency (pitch) of seet notes were higher in front of the cat and hawk, and distance to the max. pitch of seet calls was shorter for the cat than other treatments

(Appendix 1, Fig. 2). Entropy and pitch patterns suggest that titmice signaled both high aggression and fear toward the cat, and to a lesser extent toward the hawk and owl (Latimer 1977, Leavesley and Magrath 2005). Acoustic metrics that produced contrasting patterns of treatment discrimination included entropy of chip notes, chip max. frequency 1, chip note bandwidth and min. frequency 1 all distinguished the owl from other treatments (Appendix 1, Fig. 3A). Patterns in these metrics were responsible for the variable group associations of the owl (Appendix 1, 2). Additionally, the cat could be distinguished from the hawk based on distance to max. frequency 1 of the seet notes (Fig. 3B). Thus, in titmouse alarm vocalizations we found a predominance of classically risk-based signal content with significant redundancy in predation threat identification; yet there were a number of acoustic metrics that could support groupings of treatments at odds with the predominant rankings.

Ranking versus categorization of predator treatments

Overall, ranking of groups of treatments was more useful in achieving reliable identifications of treatments in our study. In general this suggests that a 'risk-based' framework would function best for competent eavesdroppers to distinguish among the stimuli presented to titmice in our experiment. Both chick-a-dee (ANOVA models) and contact call structure variation (CTA models) provided the most precise treatment ordination schemes (Fig. 2), confirming the general characterization of parid alarm calls as 'risk-based signals' (Templeton et al. 2005). Yet, both ANOVA and CTA models assuming ordinal rankings of threats incorporated groupings of two or more treatments (based on single or several metrics, respectively), allowing greater accuracy
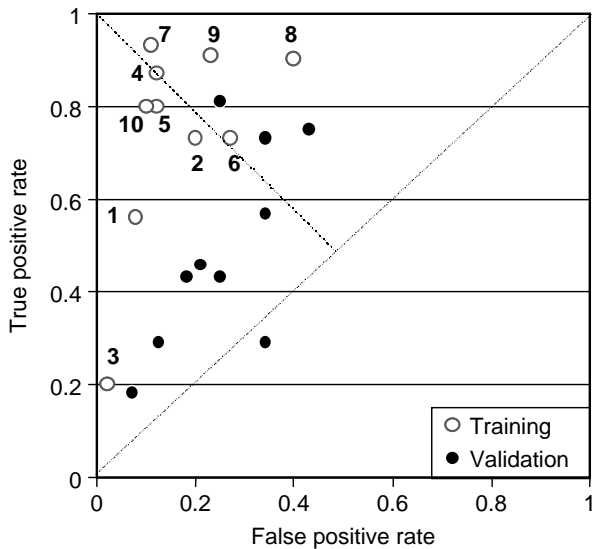
Figure 5. Receiver–operator characteristic (ROC) plot based on confusion matrices for the 10 classification tree models (see Appendix 2, Fig. 6 for model numbers). Confusion matrices were constructed using the most dangerous predator class versus other classes (combined if >1). Since model 1–3 were constructed using nominal classes (no gradations in threat were assumed), we selected whichever class included the cat as the 'most dangerous', given that titmice identified the cat as the greatest threat based on the note composition of chick-a-dee and contact calls (Fig. 2). Confusion matrices for all other models were constructed using the class with the highest assumed threat level (determined a priori; see text, Appendix 2). The best models (top left corner) correctly identified the most dangerous predator classes with either less (left of shorter line; conservative) or more false positives (right; liberal). The two best models were risk1 and risk3 (7 and 9, respectively). Size (4) and risk2 (8) are strong models (>90% correct ID of most dangerous predator) but the latter was quite liberal, generating many false positives. Both risk4 (10) and % birds in diet (5) were very conservative. The line from bottom left to top right marks random model performance (false to true positive responses =1:1). The strongest validation data, indicative of how reliably new cases could be correctly classified, were generated by risk3, and the worst by % birds in diet (Fig. 6).
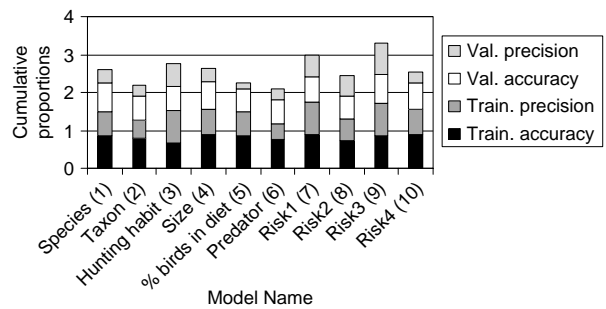

Figure 6. Summary of accuracy and precision (calculated from confusion matrices) for both training and validation data used in classification tree models with respect to identification of the most dangerous predator class versus the others. Y axis reflects the sum of four proportions. Risk1 and risk3 had the highest, and predator and taxon the lowest, overall performance across training and validation sets.

and precision in classification of all data than a straight ranking of all five treatments (Fig. 4, 5, Appendix 1, 2). No model based on single (Appendix 1) or multiple (Appendix 2) acoustic characteristics uniquely distinguished all five treatments, and the very best models consistently put snake with control and cat with hawk. We can interpret this to mean that in the context of our experiment, a mixed communication strategy allowing simultaneous ranking and grouping, could achieve the most accurate, precise, and reliable threat identification schemes. We note that our analyses assume nothing about how these calls may be used by listening animals in the wild (eavesdroppers or receivers); only how an individual that can perceive the acoustic measures we used might process the information to inform threat-dependent decisions (Seyfarth and Cheney 2003).

Unique findings

Several findings generated by our experimental set up provide unique insights into anti-predator vocalizations.

First, that the cat was perceived as the highest threat runs counter to two conclusions that are common in the literature; that higher risk predators are closer in size to their prey, and terrestrial predators are less dangerous for birds than aerial ones. The cat was much larger than the screech owl which was closer in size to the titmice; this is the reverse pattern of size and perceived threat to that demonstrated by chickadees in Templeton et al. (2005), where larger avian predators elicited lower threat calls than screech owls. The difference is that we constrained animals to face predators at very close range, simulating 'predator inspection' (Koboroff and Kaplan 2006). We interpret both findings, the reversal of size and of aerial versus terrestrial threat perception, to be related to the close proximity of test subjects to predators. Titmice reacted to their immediate condition, rather than threats deriving from typical encounters. As a terrestrial predator, a cat is not normally threatening to a midstory bird, but the species possesses a very quick attack style at close range (lashes out and leaps), faster than the avian predators confined to a perch. Attack by either bird would necessitate first flapping, lifting, and then extending the talons (Evans et al. 1993). Further, it is logical that the accipiter ranked next in threat (ANOVA) or with the cat (risk1 model) because it has longer legs and foot-strike range than the owl.

Marler (1955) has called the parid seet call the 'hawk' call or 'flying predator' call, suggesting that it should be produced only when a predator on the wing is spotted at close range; and this does happen frequently in our system (unpubl.). Moreover, playing this call causes distinct anti-attack responses; birds freeze or dive quickly into cover, ceasing all movements (feeding, preening, etc.; Lima 1997, Hetrick 2006). Together with the chick-a-dee call (given to perched predators), the seet call (for attacking predators) describes a functionally-referential signaling system – unique calls for unique situations with unique responses (Evans et al. 1993). But each of our experimental treatments elicited at least some seet calls (Table 1); even the controls (Table 1). Therefore, seet calls are not strictly referential 'flying attack' calls, just as D and chick note production occurs in various contexts (Freeberg et al. 2007). And like the latter notes, seet calls also varied in acoustic structure across treatments (Appendix 1), marking

them as potentially risk-based signals. Given the reputation of the seet call as a high danger warning call, a surprisingly high number were produced during controls (Table 1). But because we documented that the acoustic structure of seet calls produced during control trials was different from those given during higher risk trials we can now interpret the acoustic encoding of lower and higher threat in this call type as well (based on note duration and max. frequency measures; Appendix 1).

Finally, titmice are highly social and the lack of familiars to solicit (either kin or heterospecific flock mates) could have influenced vocal production (Evans et al. 1993). Most warning calls, including parid chick-a-dee and seet calls, are overtly social signals sent to kin and perhaps to heterospecific associates (Caro 2005, Templeton and Greene 2007). Relatively few alarm (mobbing) calls were produced relative to the number of contact calls (Table 1), unlike mobbing aggregations in the wild where alarm calls far outnumber contact calls (Langham et al. 2006). Titmice facing predators alone at close range may have emphasized expressions of emotive (motivational) state rather than social warning (Smith 1991). Contact calls are typically used for communicating location and other information to nearby mates and offspring (Zeigler and Marler 2004). However, since public alarm signals are normally honest (Maynard-Smith and Harper 2004), and because the few we did record were consistent across treatments and individuals, we are quite confident that our interpretations of perceived threat are correct.

Utilizing predation threat information in a dangerous world

Information availability and use by vertebrates in decision-making is complex and context-dependent (Danchin et al. 2008) and our findings support others' contentions that a dichotomous characterization of alarm communication is an oversimplification. Both categorization and ranking of experimental treatments were possible with acoustic metrics we recorded, many of which are known to be perceived and used by species that associate with parids (Templeton and Greene 2007). The strongest classification models suggested mixed approaches were best because they ranked groupings (classes) of treatments. In the Holarctic where parids are numerous, ubiquitous in woodlands, and produce significant information with relevance to predation threats, social information may be of primary importance to a majority of small prey that respond appropriately to titmouse alarms and calls (Langham et al. 2006, Schmidt et al. 2008). Relative to the richness of cues encoded in parid antipredator vocalizations, the number of typical predator encounters experienced by a given temperate-zone prey species is not likely to be very high (we estimate <10 common types of encounters). Therefore, the use of information encoded in titmouse vocalizations to group like threats and uniquely distinguish key threats is probably common among eavesdroppers and receivers, as suggested by Seyfarth and Cheney (2003) and Arnold et al. (2008); our analysis lends quantitative support to their conclusions. We emphasize that since simultaneous processing of multiple informational cues is necessary in complex decision-making (e.g. Dawson et al. 2006), studies of communication systems in vertebrates should involve appropriate multivariate analytical approaches that do not constrain human understanding to simplified (e.g. risk-based) frameworks. In sum, we propose that characterization of the availability, perception, cognitive processing, and use of information in animal decision-making needs a broader operational framework and greater empirical focus on identifying underlying causal relationships.

References

Arnold, K. and Zuberbühler, K. 2005. Semantic combinations in primate calls. – Nature 441: 303.

Arnold, K. et al. 2008. A forest monkey's alarm call series to predator models. – Behav. Ecol. Sociobiol. 62: 549–559.

Baker, M. C. and Becker, A. M. 2002. Mobbing calls of black-capped chickadees: effects of urgency on call production. – Wilson Bull. 114: 510–516.

Bent, A. C. 1937. Life histories of North American birds of prey (part 1). – US Natl Mus. Bull. 167: 95–111.

Bloomfield, L. L. et al. 2005. Note types and coding in parid vocalizations. III: The chick-a-D call of the Carolina chickadee (*Poecile carolinensis*). – Can. J. Zool. 83: 820–823.

Blumstein, D. T. 1995. Golden marmot alarm calls. I. The production of situationally specific vocalizations. – Ethology 100: 113–125.

Blumstein, D. T. 1999. The evolution of functionally referential alarm communication: multiple adaptations; multiple constraints. – Evol. Comm. 3: 135–147.

Blumstein, D. T. and Armitage, K. B. 1997. Alarm-calling in yellow-bellied marmots. I. The meaning of situationally variable alarm calls. – Anim. Behav. 53: 143–171.

Breiman, L. et al. 1984. Classification and regression trees. – Chapman and Hall/CRC Press.

Caro, T. 2005. Antipredator defenses in birds and mammals. – Univ. of Chicago Press.

Danchin, E. et al. (eds) 2008. Behavioural ecology. – Oxford Univ. Press.

Dawson, M. R. W. et al. 2006. Statistical classification of black-capped (*Poecile attricapillus*) and mountain chickadee (*Poecile gambeli*) call notes. – J. Comp. Psychol. 120: 147–153.

Dolby, A. S. and Grubb, T. C. Jr 1998. Benefits to satellite members in mixed species foraging groups: an experimental analysis. – Anim. Behav. 56: 501–509.

Evans, C. S. et al. 1993. On the meaning of alarm calls: functional reference in an avian vocal system. – Anim. Behav. 46: 23–38.

Farley, E. A. et al. 2008. Characterizing complex mixed-species bird flocks using an objective method for determining species participation. – J. Ornithol. 149: 451–468.

Fichtel, C. and Kappeler, P. M. 2002. Anti-predator behavior of group-living Malagasy primates: mixed evidence for a

referential calling system. – Behav. Ecol. Sociobiol. 51: 262–275.

Ficken, M. S. 1989. Acoustic characteristics of alarm calls associated with predation risk in chickadees. – Anim. Behav. 39: 400–401.

Ficken, M. S. et al. 1994. The chick-a-dee call system of the Mexican chickadee. – Condor 96: 70–82.

Freeberg, T. M. et al. 2003. Variation in chick-a-dee calls of a population of Carolina chickadees, *Poecile carolinensis*: identity and redundancy within note types. – J. Acoust. Soc. Am. 113: 2127–2136.

Freeberg, T. M. et al. 2007. Complexities in vocal communication. – In: Otter, K. A. (ed.), The ecology and behavior of chickadees and titmice: an integrated approach. Oxford Univ. Press, pp. 235–240.

Gaddis, P. K. 1980. Mixed flocks, accipiters, and antipredator behavior. – Condor 82: 348–349.

Gaddis, P. K. 1983. Composition and behavior or mixed-species flocks of forest birds in north–central Florida. – Fla Field Nat. 11: 25–44.

Greenberg, R. 2000. Birds of many feathers: the formation and structure of mixed-species flocks of forest birds. – In: Boinski, S. and Garber, P. A. (eds), On the move: how and why animals travel in groups. Univ. of Chicago Press, pp. 521–558.

Hailman, J. P. 1989. The organization of major vocalizations in the Paridae. – Wilson Bull. 101: 305–343.

Hailman, J. P. et al. 1985. The 'chick-a-dee' calls of *Parus atricapillus*: a recombinant system of animal communication compared to written English. – Semiotica 56: 191–224.

Halliday, T. R. and Adler, K. (eds) 1986. The encyclopedia of reptiles and amphibians. – Facts on File Inc., New York.

Harrap, S. and Quinn, D. 1995. Chickadees, tits, nuthatches and treecreepers. – Princeton Univ. Press.

Hetrick, S. A. 2006. Investigation of tufted titmouse (*Baeolophus bicolor*) anti-predator vocalizations. MSc thesis. – Univ. of Florida, Gainesville, FL, USA.

Kirchhof, J. and Hammerschmidt, K. 2006. Functionally referential alarm calls in tamarins (*Saguinus fuscicollis* and *Saguinus mystax*) – evidence from playback experiments. – Ethology 112: 346–354.

Koboroff, A. and Kaplan, G. 2006. Predator inspection by birds. – J. Ornithol. 147: 195.

Langham, G. M. et al. 2006. Why pishing works: titmouse (Paridae) scolds elicit a generalized response in bird communities. – Ecoscience 13: 485–496.

Latimer, W. 1977. A comparative study of the songs and alarm calls of some *Parus* species. – Z. Tierpsychol. 45: 414–433.

Leavesley, A. J. and Magrath, R. D. 2005. Communicating about danger: urgency alarm calling in a bird. – Anim. Behav. 70: 365–373.

Lima, S. L. 1997. Ecological and evolutionary perspectives on escape from predatory attack – a survey of North American birds. – Wilson Bull. 105: 1–47.

Manser, M. B. 2001. The acoustic structure of suricate's alarm calls varies with predator type and level of response urgency. – Proc. R. Soc. Lond. B 268: 2315–2324.

Marler, P. 1955. Characteristics of some animal calls. – Nature 176: 6–8.

Maynard-Smith, J. and Harper, D. 2004. Animal signals. – Oxford Univ. Press.

Mönkkönen, M. and Forsman, J. T. 2002. Heterospecific attraction among forest birds: a review. – Ornithol. Sci. 1: 41–51.

Morse, D. H. 1973. Interactions between tit flocks and sparrowhawks *Accipiter nisus*. – Ibis 115: 591–593.

Muller, C. A. and Manser, M. B. 2008. The information banded mongooses extract from heterospecifics alarms. – Anim. Behav. 75: 897–904.

Munn, C. A. and Terborgh, J. W. 1979. Multi-species territoriality in neotropical foraging flocks. – Condor 81: 338–347.

Nowicki, S. 1989. Vocal plasticity in captive black-capped chickadees: the acoustic basis and rate of call convergence. – Anim. Behav. 37: 64–73.

Otter, K. (ed.) 2007. The ecology and behavior of chickadees and titmice: an integrated approach. – Oxford Univ. Press.

Schmidt, K. A. et al. 2008. Eastern chipmunks increase their perception of predation risk in response to titmouse alarm calls. – Behav. Ecol. 19: 759–763.

Seppänen, J.-T. et al. 2005. Presence of other species may counter seasonal decline in breeding success – a field experiment. – J. Avian Biol. 36: 380–385.

Seyfarth, R. M. and Cheney, D. L. 2003. Signalers and receivers in animal communication. – Annu. Rev. Psychol. 54: 145–173.

Seyfarth, R. M. et al. 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. – Science 210: 801–803.

Sherrod, P. H. 2004. DTREG, classification and regression trees for data mining and modeling, rel. 3.5. – P. H. Sherrod, Brentwood, TN.

Sieving, K. E. et al. 2004. Heterospecific facilitation of forest-boundary crossing by mobbing understory birds in north–central Florida. – Auk 121: 738–751.

Smith, S. T. 1972. Communication and other social behavior in *Parus carolinensis*. – Nuttall Ornithol. Club, Cambridge, UK.

Smith, S. M. 1991. The black-capped chickadee: behavioral ecology and natural history. – Cornell Univ. Press, Ithaca.

Templeton, C. N. and Greene, E. 2007. Nuthatches eavesdrop on variations in heterospecific chickadee mobbing alarm calls. – Proc. Natl Acad. Sci. USA 104: 5479–5482.

Templeton, C. N. et al. 2005. Allometry of alarm calls: black-capped chickadees encode information about predator size. – Science 308: 1934–1937.

Wiles, L. and Brodahl, M. 2004. Exploratory data analysis to identify factors influencing spatial distributions of weed seed banks. – Weed Sci. 52: 936–947.

Zeigler, H. P. and Marler, P. 2004. Bird calls-their potential for behavioral neurobiology. Behavioural neurobiology of bird song. – Ann. NY Acad. Sci. 1016: 31–44.

Appendix 1. List of gross and fine-scale spectrographic measures taken from chick-a-dee calls chick, D, seet, song, chip, and chink notes produced by titmice during trials that varied significantly (p <0.05) in ANOVA among treatments (control, cat, hawk, owl and snake) and that had model adj. R^2-values >0.10 (18 significant models had lower R^2-values). Variable abbreviations in bold italics. Adjusted R^2 and pairwise comparisons (LSD; p <0.05) are shown. Gross measures were calculated as mean values for note types produced in either the first minute following exposure to treatments or over the 5 min response period. All fine measures were summarized from all distinct notes produced during the 5 min response period.

Gross measures; variable ID	R^2	LSD pair-wise comparisons
No. chip notes in min 1; *chip1m*	0.19	[co =sn] >[ca =ow =ha]
No. chink notes in min 1; *chink1m*	0.12	ha >[ow =sn =co]; [ca =ha]; [ca =ow =sn =co]
No. chick notes in min 1; *chick1m*	0.12	ha >[co, ow, sn]; [co =ca =ow =sn]
No. D notes in min 1; *d1m*	0.13	ha >[co, ow, sn]; ca =ha; [co =ca =ow =sn]
No. chip notes in 5 min; *chip5m*	0.17	[co =sn] >[ca =ow =ha]
Chink duration 5 min (s); *chinkdur*	0.18	[sn =ow] >[co =ha =ca]
Chick duration 5 min (s); *chickdur*	0.18	[sn =ow] >[co =ca =ha]
D duration 5 min (s); *ddur*	0.45	ha >[co =ca =ow], sn; ca >sn
Seet duration 5 min (s); *seetdur*	0.80	co >ow >ca
Pr. dees per chick-a-dee call; *d/cd5m**	0.59	ca >ha >ow >[sn =co]
Pr. chinks/chips, 1 min; *chink/cont1m**	0.44	[ca =ha =ow] >[sn =co]
Pr. chinks/chips, 5 min; *chink/cont5m**	0.34	[ca =ha =ow] >[sn =co]

Fine measures (means, 5 min)	R^2	LSD pair-wise comparisons
Chip note bandwidth; *chipbawi*	0.13	ow >[co =ca =sn =ha]
Chink note bandwidth; *chinkbawi*	0.12	ca >[co =ow] >sn; ow >ha >sn
D note dist. to max. freq. (mm); *ddmxfq*	0.12	[ca =ha] >[ow =sn =co]
Seet note dist. to max. freq.; *seetdtmxf*	0.23	[ha =ow =sn =co] >ca
Chip note min. freq. 1; *chipmnfq1*	0.10	[co =ca =sn =ha] >ow
Chip note max. freq. 1; *chipmxfq1*	0.11	ow >ca >[co =sn =ha]
Chink note max. freq. 1; *chinkmxfrq1*	0.12	ca >[ha =ow =co] >sn
Seet note max. freq. 1; *seetmxfq1*	0.32	ca >ha >[ow =sn =co]
Chick note min. freq. 2; *chickmnfq2*	0.34	ha >[co, ow]; [ca, co, sn] >ow; [co =ca =sn]
Seet note min. freq. 2; *seetmnfq2*	0.42	ca >ha >[ow =sn =co]
D note max. freq. 2; *dmxfq2*	0.12	[ca =ha] >[ow =sn =co]
Seet note max. freq. 2; *seetmxfq2*	0.34	ca >ha >[ow =sn =co]
Chip note entropy; *chipent*	0.11	ow >ca >ha >[sn =co]
Chink note entropy; *chinkent*	0.12	ca >co >ow >[ha =sn]
Chick note entropy; *chickent*	0.26	ow >ca >ha; [co =ca =sn]; [co >ha]
D note entropy; *dent*	0.20	ha >ca >[co =ow =sn]

*each of these models has parallel ANOVA models (with equivalent F, P, LSD values) for chicks/chick-a-dee call, chips/contact calls in min 1, and 5 min., respectively (not reported).

Appendix 2. Classification tree model summaries, with model names (and ID numbers) in the first column. Models 1–3 assume a nominal classification of treatments (different in kind; species, taxon, hunting habit) and the rest assume ordinal rankings (different in degree; size, diet, predator and the four risk models). Species simply separates the five treatments. Taxon groups avian predators distinct from the mammal and reptile, hunting habit groups the aerial versus terrestrial hunters. Size groups represent mean published mass for adults of species used; a factor known to be correlated with parid alarm call structure (Templeton et al. 2005). Domestic cats weigh the most, rat snake and sharp shinned hawks are intermediate, and screech owls are much smaller. % birds in diet reflects the frequency that predators hunt successfully for adult passerine birds (cat, hawk and owl, frequently; snakes, not so often), and predator separates all animals from the control. The four risk models are predicated on how the experimental set-up may have influenced birds' perception of direct attack probability from the predators they faced (<1 m away). In risk1, cat and hawk pose equally high attack risk (swift/powerful), versus the smaller, weaker owl versus slow snake; in risk2 the owl is placed with the low risk snake, but included in the high risk group in risk3. Risk4 assumes that the cat can strike quickest by leaping with its paws outstretched; both birds must first jump, flap, and then pounce, and the snake must first crawl and then strike (and only very slowly because of low winter temperatures). Class indicates how treatments were grouped. Gain for each treatment class and the mean of classes is presented. The important variables column lists: the total number of variables included in the model (no. in bold); those with >45% importance (see Appendix 1 for abbreviations); and the first split in each model (italics).

Model	Class	Gain	Mean (SD)	Important variables
Species (1)	cat (c)	78%	83% (7.1)	**10:** chink/cont5m; chip/cont5m; chip/cont1m; chippres1m; seetdtmfq; chip1m (49) *1st split: snake <owl; chink/cont5m*
	hawk (h)	95%		
	owl (o)	79%		
	snake (s)	84%		
	control (con)	79%		
Taxon (2)	bird (o, h)	80%	86% (6.6)	**9:** chink/cont5m; chip/cont5m; chip5m; chippres1m; chip1m; chip/cont1m *1st split: reptile <bird; chink/cont5m*
	mammal (c)	82%		
	reptile (s)	95%		
	control	86%		
Hunting habit (3)	ground (c, s)	61%	73% (10.2)	**6:** chip1m; chink/cont5m; chip/cont5m *1st split: control <aerial; chink/cont5m*
	aerial (h, o)	75%		
	control	82%		
Size (4)	large (c)	89%	85% (9.8)	**16:** chip5m; chinkmnfq2; chink/cont5m; chinkent; dent; chinkdur; chinkmxfq1; dmxfq2; seetmnfq2 *1st split: large <control; chip5m*
	medium (h, s)	71%		
	small (o)	89%		
	control	93%		

Model	Confidence of treatment ID		Model performance	
	Class	Gain	Mean (SD)	Important variables
% birds in diet (5)	high (c, h, o)	80%	91% (10.0)	**19:** chink/cont5m; chip/cont5m; chippres1m; chip1m; chinkent; chip/cont1m *1st split: low <high; chink/cont5m*
	low (s)	94%		
	control	100%		
Predator (6)	yes (h, o, s, c)	58%	68% (15.0)	**1:** chip1m *1st split: yes <no; chip1m*
	no (con)	79%		
Risk1 (7)	high (h, c)	99%	94% (9.5)	**22:** chink/cont5m; chip/cont5m; chippres1m; chip/cont1m *1st split: low <medium; chink/cont5m*
	medium (o)	100%		
	low (s, con)	83%		
Risk2 (8)	high (h, c)	76%	75% (0.7)	**6:** chip/cont1m; chippres1m; chip1m; chink/cont5m; chip/cont5m *1st split: low >high; chip/cont1m*
	low (o, s, con)	75%		
Risk3 (9)	high (c, h, o)	71%	76% (7.8)	**6:** chink/cont5m; chip/cont5m; chip1m; chippres1m; chip/contact1; chip5m *1st split: low <high; chink/cont5m*
	low (s, con)	82%		
Risk4 (10)	strike (c)	84%	84% (1.0)	**9:** chink/cont5m; chip/cont5m; chip5m; chippres1m; chip/cont1m; chip1m; seetdtmfq *1st split: flap >crawl; chink/cont5m*
	flap (h, o)	85%		
	crawl (s, con)	83%		