DOI: 10.1111/1365-2656.13620

RESEARCH ARTICLE

Journal of Animal Ecology

Function of a multimodal signal: A multiple hypothesis test using a robot frog

Vinícius Matheus Caldart^{1,2} Vinícius Beux dos Santos³ Kierce da Caldart^{1,2}

¹LAGE do Departamento de Ecologia, Instituto de Biociências. Universidade de São Paulo, São Paulo, Brazil

²Departamento de Biologia Geral, Universidade Estadual de Ponta Grossa, Ponta Grossa, Brazil

³Departamento de Bioquímica e Biologia Molecular, Programa de Pós-graduação em Bioquímica Toxicológica, Universidade Federal de Santa Maria, Santa Maria, Brazil

Correspondence

Vinícius Matheus Caldart Email: viniciuscaldart@gmail.com

Funding information

Fundação de Amparo à Pesquisa do Estado de São Paulo, Grant/Award Number: 2015/14336-6 and 2021/00915-5; Conselho Nacional de Desenvolvimento Científico e Tecnológico, Grant/Award Number: 306550/2014-8

Handling Editor: Beatriz Willink

Abstract

- 1. Multimodal communication may evolve because different signals may convey information about the signaller (content-based selection), increase efficacy of signal processing or transmission through the environment (efficacy-based selection), or modify the production of a signal or the receiver's response to it (inter-signal interaction selection).
- 2. To understand the function of a multimodal signal (aggressive calls + toe flags) emitted by males of the frog Crossodactylus schmidti during territorial contests, we tested two hypotheses related to content-based selection (quality and redundant signal), one related to efficacy-based selection (efficacy backup), and one related to inter-signal interaction selection (context). For each hypothesis we derived unique predictions based on the biology of the study species.
- 3. In a natural setting, we exposed resident males to a robot frog simulating aggressive calls (acoustic stimulus) and toe flags (visual stimulus), combined and in isolation, and measured quality-related traits from males and local levels of background noise and light intensity.
- 4. Our results provide support to the *context* hypothesis, as toe flags (the context signal) are insufficient to elicit a receiver's response on their own. However, when toe flags are emitted together with aggressive calls, they evoke in the receiver qualitatively and quantitatively different responses from that evoked by aggressive calls alone. In contrast, we found no evidence that toe flags and aggressive calls provide complementary or redundant information about male quality, which are key predictions of the quality and redundant signal hypotheses respectively. Finally, the multimodal signal did not increase the receiver's response across natural gradients of light and background noise, a key prediction of the efficacy backup hypothesis.
- 5. Toe flags accompanying aggressive calls seem to provide contextual information that modify the receiver's response in territorial contests. We suggest this contextual information is increased motivation to escalate the contest, and discuss the benefits to the signallers and receivers of adding a contextual signal to the aggressive display. Examples of context-dependent multimodal signals are rare in the literature, probably because most studies focus on single hypotheses assuming content- or efficacy-based selection. Our study highlights the importance of considering multiple selective pressures when testing multimodal signal function.

KEYWORDS

aggressiveness motivation, complex communication, context-dependent signal, signal matching, stream-breeding frog, territorial contests

1 | INTRODUCTION

Individuals of many species exchange information using multiple signals from the same or different sensory modality, a behaviour termed complex communication (Hebets & Papaj, 2005). Multimodal communication is a form of complex communication characterized by presenting signals emitted and received through two or more sensory modalities (Partan & Marler, 2005). Despite great diversity of multimodal signals, selective pressures driving their evolution are categorized into three types (Hebets & Papaj, 2005). In *content-based* selection, combined signals convey information about the signaller, including quality, location, sex and species identity. In *efficacy-based* selection, combined signals increase the efficacy of signal transmission and processing through the environment. Finally, in *inter-signal interaction* selection, different signals interact so that one signal changes the production of a second signal or changes the receiver's response to a second signal.

Over the past decades, the isolated action of different selective pressures has been evoked to explain the evolution of multimodal signals in various taxa. For instance, content-based selection may explain the multiple signals used in agonistic interactions by male elands (Tragelaphus oryx). Each signal reflects a separate component of male fighting ability: the dominant frequency (DF) of the sound produced by knee-clicks is an honest signal of body size, the dewlap droop is an indicator of age, and facemask darkness is related to androgen-related aggressiveness (Bro-Jørgensen & Dabelsteen, 2008). Efficacy-based selection, in turn, may explain the multimodal courtship display of some wolf spiders (Schizocosa spp.). The use of both visual (leg waving) and seismic (drumming) signals allows males to achieve copulation in dark and illuminated places, as well as in substrates that transmit or not vibrations (Uetz et al., 2016). Finally, inter-signal interaction selection may explain why snapping shrimp (Alpheus heterochaelis) males respond differently to the visual stimulus of an open claw depending on a sex-specific chemical signal emitted by the sender. If the sender is a female it leads to pair formation, but if the sender is a male, it leads to agonistic interactions (Hughes, 1996).

In an influential review about complex communication, Hebets & Papaj (2005) suggested that one could narrow down possible functional explanations for a multimodal display by addressing questions about signal content, efficacy and inter-signal interactions. In another important contribution, Partan & Marler (2005) classified signals as *redundant*, when isolated signals generate a qualitatively equivalent response by the receiver, or *non-redundant*, when signals generate a qualitatively different response by the receiver. These two works provide a framework for testing hypotheses on the types of selective pressures driving the evolution of multimodal signals.

We used this framework to understand the function of a multimodal display emitted by the frog Crossodactylus schmidti (Figure 1a). Besides advertisement and aggressive calls, males have a large repertoire of visual signals, with toe flagging being the most frequently employed in social interactions (Caldart et al., 2011, 2014). Toe flags have been observed in agonistic interactions of several diurnal frogs from lotic habitats (e.g. Furtado et al., 2019; Hartmann et al., 2005), and consists of lifting toes of both feet from the substrate, performing repeated up-and-down movements, often showing contrasting colorations between ventral and dorsal parts (see video in Appendix S1). In C. schmidti, toe flags are usually emitted combined with aggressive calls during male-male contests for territory possession (Caldart et al., 2014). An outstanding question is whether toe flags convey quality information about the signaller, increase the efficacy of signal transmission, or interact with aggressive calls to elicit or modify a receiver's response.

We built an electromechanical robot frog that simulates a *C*. *schmidti* male and programmed the robot to emit aggressive calls and toe flags, either combined or in isolation. Then, we exposed it to territorial males in the field to induce agonistic interactions (Figure 1b,c). Using a receiver-based approach, we assessed the behavioural responses of resident males to each stimulus and tested the influence of environmental variables and traits related to individual quality on the receivers' signalling responses. The data gathered in the experiment were used to test four hypotheses: two related to *content-based* selection (*quality* and *redundant signal*), one related to *efficacy-based* selection (*context*). Based on extensive knowledge about the behaviour and ecology of *C. schmidti*, we derived distinctive combinations of predictions for each hypothesis (Tables 1 and 2).

2 | MATERIALS AND METHODS

2.1 | Stimuli preparation and programming

We built an electromechanical robot frog that mimics the morphology and signalling behaviour of *C. schmidti* males. The silicon-vested robot mounts an impermeable rock mould that mimics a calling site (Figure 1; Appendix S2). We programmed the robot to emit aggressive calls and toe flags, either combined or in isolation, comprising three experimental groups: (a) *acoustic* (aggressive calls, stimulus *a*), (b) *visual* (toe flags, stimulus *v*) and (c) *multimodal* (aggressive calls + toe flags, stimulus *a* + *v*). The stimulus *a* consisted of an aggressive call from an average-sized male used in the original call description (Caldart et al., 2011), with temporal features modified in FIGURE 1 (a) Marked *Crossodactylus* schmidti male on a rock, the typical signalling site in the torrent streams inhabited by this species, and (b–c) general view of the experimental setup. The numbers indicate: (1) coloured flags marking the territories of focal males; (2) the robot frog; (3) focal male individually marked with a temporary waist belt; and (4) video-camera placed 3 m apart from the robot and 3.7 m from the focal male. The robot was always placed 70 cm away from the focal male



Adobe Audition to match population means (Appendix S2). Stimulus v consisted of toe flags as described in Caldart et al. (2014). To confirm mean duration and mean rate of toe flag emission, we analysed videos of nine males in territorial contests (Appendix S2). All stimuli programmed in the robot had the same temporal structure: a 4 min stimulus phase preceded and followed by a 4 min control phase (preand post-controls), totalling 12 min. Stimulus a + v consisted of a train of 12 s of aggressive call with a simultaneous toe flag every 12 notes, followed by 28-s silence (Appendix S2). This sequence was repeated six times during the stimulus phase. Stimuli *a* and *v* were composed of aggressive calls alone and toe flags alone respectively. The exact moments of signal emission in the stimulus phase were the same for the three types of stimuli (Appendix S2). During the preand post-control phases the robot remained silent and motionless.

2.2 | Field experiment

We conducted the experiment at Turvo State Park (27°14'34.08"S, 53°57'13.74"W), State of Rio Grande do Sul, southern Brazil, located in an Atlantic forest fragment where previous studies with *C. schmidti* have been conducted. We captured and marked males in calling activity between October 2016 and January 2017. In these months, male activity and frequency of sexually mature adults do not vary significantly (Caldart et al., 2016a, 2019). After capture,

we recorded the snout-vent length (SVL, precision 0.01 mm) and body mass (precision 0.1 g) of the males and marked them with a temporary cotton waist pelvic belt containing an individual code (Figure 1a). After releasing each male to his calling site, we tied a flag containing his individual code to the vegetation above the calling site (Figure 1b).

Because males are territorial, we could locate the marked individuals during the experiment to expose them to all experimental stimuli. One day after the marking procedure, we searched for marked males between 09:00 and 17:00 hr. We observed a marked male for 5 min and positioned the robot at a viewing distance of 70 cm, at an angle of 30° relative to the focal male (Figure 1b,c). We then waited for 5 min for acclimatization and, if no activity by the focal male was recorded, we exposed him to one of the three experimental stimuli (a, v or a + v). To avoid the influence of the order of exposure of stimulus types in males' responses, we randomized the sequence of stimuli presentation to each individual.

We recorded the entire trial for each focal male with a digital camcorder (Sony Handycam HDR-CX405, Figure 1b,c) and, immediately after the footage, we measured the levels of noise (dB) and light intensity (lux) 30 cm above the calling site with a sound-level metre (Instrutherm DEC 500; C weighting curve: 20–12,500 Hz, dB range: 35–130 dB) and a luximeter (Instrutherm LD400). After finishing a trial, we either waited for at least 30 min (plus 5 min of acclimatization) before assigning another experimental stimulus to the TABLE 1 Information on the behaviour and ecology of *Crossodactylus schmidti*, and the four hypotheses selected to understand the function of a multimodal display emitted by the males during territorial contests. Nomenclature and description of the hypotheses follow Hebets and Papaj (2005)

General information

Crossodactylus schmidti is diurnal and inhabits torrent streams with varying levels of background noise and sunlight incidence (Caldart et al., 2016a,b). Male calling activity and production of gametes in both sexes occur year-round (Caldart et al., 2019). Males compete for territories containing rocks (signalling sites) and underwater chambers (oviposition sites). Males as far as 2 m from each other exchange aggressive calls and visual signals. When an intruder approaches a resident male and emits aggressive calls, the primary aggressive response of the resident is to emit aggressive calls in return. If the intruder does not withdraw, both males exchange long-lasting aggressive calls and diverse visual signals (Caldart et al., 2011, 2014). Before escalating to physical contests, males exchange for a long time the most common multimodal signal: aggressive calls plus toe flags. Territorial contests last as much as 30 min (Caldart et al., 2014)

Content-based hypotheses

Quality: different signals provide different information about signaller quality.

This hypothesis assumes that more information is better information. As occurs with other frog species (Dyson et al., 2013), body size may be important to the outcome of contests in *C. schmidti*. Larger males are older (i.e. more experienced) and have larger testes (i.e. high levels of testosterone; Caldart et al., 2019). Body condition may also be important because long-lasting emission of toe flags may be costly. Thus, both body size and body condition may be related to male fighting ability or resource-holding potential. If the multimodal signal in *C. schmidti* conveys independent pieces of information about body size and condition (i.e. two proxies of male quality) during contests, a key prediction is: (1) the acoustic signal (i.e. dominant frequency of aggressive calls or rate of emission of aggressive notes) correlates with one proxy of quality, while the visual signal (i.e. toe flags) correlates with another proxy of quality

Redundant signal: different signals provide the same information about signaller quality.

This hypothesis assumes that different signals provide the same information about the signaller quality, allowing for increased accuracy of receiver's response. If the multimodal signal in *C. schmidti* conveys redundant pieces of information about male size or condition during contests, key predictions are: (1) the acoustic and visual signals correlate with the same proxy of male quality (i.e. body size or condition); (2) because both signals reflect the same proxy of male quality, there should be a positive correlation between acoustic and visual signals. These predictions are based on the same rationale presented above about the importance of male size and condition to the outcome of territorial contests

Efficacy-based hypothesis

Efficacy backup: one signal acts as a backup to the other in the presence of environmental variability.

This hypothesis assumes that environmental variation renders signalling and responding suboptimal or ineffective. Multimodal signals could increase the efficacy with which different signals travel through the environment and are received. We know that part of the multi-note call of *C. schmidti* is masked by the stream-generated noise and that males increase calling activity along the day as light levels reaching signalling sites increases (Caldart et al., 2016a,b). If the multimodal signal in *C. schmidti* improves signal efficacy during contests, key predictions are: (1) each signal is sufficient to elicit a receiver's response, but (2) the probability of response to the multimodal signal is higher than to the isolated signals, especially in conditions of low light intensity and high background noise, in which the transmission of the unimodal signals is more constrained; (3) the latency to respond to the multimodal signal is shorter than to the isolated signals; (4) because the acoustic and visual signals reinforce the efficacy of each other, there should be a positive correlation between them

Inter-signal interaction hypothesis

Context: the presence of one signal provides a context for the receiver to interpret and respond to a second signal.

This hypothesis assumes that receivers' response to a single signal is context dependent. Receivers often face situations in which they cannot accurately interpret a signal because its meaning is either dynamic or hard to interpret. Multimodal signals may reduce signal ambiguity by providing additional, contextual information used by the receiver to interpret and respond to a second signal. In the beginning of a contest in *C. schmidti*, the intruder male emits aggressive calls that probably communicate his intention of taking over the territory of the resident male. In some contests the exchange of aggressive calls is enough to repel the intruder. However, contestants may escalate the contest and start exchanging aggressive calls accompanied by toe flags. If toe flags provide a new context (e.g. aggressiveness motivation) for receivers to interpret aggressive calls and modify their response to it during territorial contests, key predictions are: (1) receivers do not respond to the visual signal alone (context signal), but respond to the acoustic and multimodal signals; (2) the responses to each signal are qualitatively different; (3) because the addition of toe flags provides a new context, the receivers' response to the multimodal signal is different (e.g. more intense) from the response to the acoustic signal alone

same focal male – continuing the trial only if the male had stopped interacting with the robot–or moved the robot to another site and repeated the procedure with another male. We recorded responses of 39 males, from which 25 were exposed to all stimuli and 14 to two stimuli (*a* and v = 2; *a* and a + v=5; *v* and a + v = 7). We excluded one male that did not respond to any stimulus. The final sample size was 38 males (33 exposed to *v*, 31 to *a* and 36 to a + v).

2.3 | Statistical analyses

To answer the questions presented in Table 2 we used different sets of response and predictor variables (Table 3). We used as response variables both the probability and rates of emission of the following signals by focal males: aggressive notes, which are the basic units of aggressive calls (Caldart et al., 2011); toe flags, which are the TABLE 2 Predictions of the hypotheses selected to understand the function of a multimodal display emitted by *Crossodactylus schmidti* males during territorial contests. Key predictions for supporting each hypothesis are in boldface. Questions (1-3) refer to how focal males are expected to respond to visual (v), acoustic (a) and multimodal stimuli (a + v) emitted by the robot. Questions (4-5) refer to the expected relationship between signals (4) and between signals and the proxies of male quality (5) in the responses of focal males to the stimuli emitted by the robot. Question (6) refers to how focal males are expected to respond to the multimodal stimuli emitted by the robot in the presence of environmental variability. Responses of the focal males to the stimuli emitted by the robot were assessed as: emission of aggressive notes, toe flags and other visual signals, movement towards the robot, attack to the robot, or no response. Nomenclature of the hypotheses follows Hebets and Papaj (2005)

Questions ^a	Quality	Redundant signal	Efficacy backup	Context
(1) Do <i>a</i> and <i>v</i> elicit a receiver's response on their own?	At least one signal (a or v) elicits a response	At least one signal (<i>a</i> or <i>v</i>) elicits a response	Both signals (<i>a</i> and v) elicit a response	Context signal (v) alone does not elicit a response, the other signal (a) does
(2) Are the receiver's responses to <i>a</i> and <i>v</i> qualitatively different?	Responses to each signal alone may or may not be different	Responses to each signal alone may or may not be different	Responses to each signal alone may or may not be different	Responses to each signal alone are different (no response to v, response to a)
(3) What are the effects of <i>a</i> + <i>v</i> on the receiver's response?	Combined signals can increase or decrease the intensity of response	Combined signals can increase or decrease the intensity of response	Combined signals increase probability of response and reduce latency to response	Combined signals increase the intensity of response compared to <i>a</i>
(4) Is there a correlation between <i>a</i> and <i>v</i> in the receiver's response?	Signals can or cannot be correlated	Signals are positively correlated	Signals are positively correlated	Signals can or cannot be correlated
(5) Do <i>a</i> and/or v in the receiver's response covary with its quality?	Signal <i>a</i> covaries with one proxy of quality; signal <i>v</i> covaries with another proxy	Both signals (<i>a</i> and v) covary with the same proxy of quality	No	No
(6) Is the probability of receiver's response, across the environmental gradients, higher to $a + v$?	No	No	Yes (especially in conditions in which the transmission of unimodal signals is more constrained)	No

^a(1) and (4-6) follow Hebets & Papaj (2005); (2-3) follow Partan & Marler (2005).

most frequent visual signal emitted in social interactions (Caldart et al., 2014); and other visual signals, which includes three displays (i.e. limb lifting, running–jumping and body jerking) that were pooled together because they are emitted less frequently than toe flags (Caldart et al., 2014), but may play a role on the multimodal communication between males during agonistic interactions. We also used as response variables the DF and the modulation of DF (i.e. the difference between the higher and the lower DF values) of focal males aggressive calls, latency to emit aggressive notes, toe flags, and other visual signals, latency to move towards the robot after the onset of the stimulus, and frequency of attacks to the robot. Different combinations of these variables were used in statistical models as described below.

Questions (1–3): We built generalized mixed-effects models (GLMMs) to test the effects of the experimental phases of each stimulus emitted by the robot on the rates of emission of different signals by focal males, fitting male identity as a random factor to control for repeated exposure of the same individual to different experimental phases and stimuli. We used the rates of signal emission as a continuous response variable and the pre- and post-control phases as categorical predictors. To test post-hoc differences in the rates of signal emission between experimental phases, we computed the estimated

marginal means of the response variable for each experimental phase based on the fitted models and calculated customized contrasts to compare the estimated marginal means with one another using the package EMMEANS (Lenth, 2019). We used the *mvt* alpha adjustment method for multiple contrasts in the MVTNORM package (Genz et al., 2020). Finally, we tested if there is any association between the types of stimuli emitted by the robot and the frequency of attacks to the robot by the focal males using a Chi-square test.

Question (4): To test for correlations between signals in the response of focal males, we obtained Spearman correlation coefficients for the relationships between the rate of emission of aggressive notes and toe flags, considering signals emitted during the stimulus phases of each experimental group separately.

Question (5): We built linear models and GLMMs to test the effects of two proxies of male quality on their signalling response during the stimulus phases of each experimental group. Our first proxy of quality was body size (SVL) because larger males have fight advantages in anurans (Dyson et al., 2013). Body size also constrains the variation in the call DF, meaning that this acoustic trait provides honest size-related information. Our second proxy of quality was body condition, estimated as the residuals of an ordinary least square regression between log₁₀-transformed body mass

Questions	Model type	Response variables	Error distribution (link function)	Predictor variables	Random variable	Model results
(1-3)	GLMM	Emission rates ^a of aggressive notes, toe flags, other visual signals	Tweedie-Poisson ^d (log-link)	Experimental phases	Male identity	Appendix S5
(5)	Linear regression	Dominant frequency (DF) ^b and modulation of DF ^b (in Hz)	Gaussian	Body size, body condition	Not applicable	Appendix S3
	GLMM	Emission rates ^b of aggressive notes, toe flags, other visual signals	Gamma (log-link)	Body size, body condition, type of stimulus	Male identity	Appendix S3
	GLMM	Probability of emission ^b of aggressive notes, toe flags, other visual signals	Binomial (cloglog ^e and logit)	Body size, body condition, type of stimulus	Male identity	Appendix S3
(6)	GLMM	Emission rates ^b of aggressive notes, toe flags, other visual signals	Gamma (log-link)	Background noise, light intensity, type of stimulus	Male identity	Appendix S7
	Cox-Proportional Hazards	Latency ^c to emit aggressive notes, toe flags, other visual signals Latency ^c to move towards the robot	Not applicable	Background noise, light intensity, type of stimulus	Male identity	Appendix S7

TABLE 3 Specifications of the models used to answer questions (1–3), (5) and (6) presented in Table 2

^aEmitted during all experimental phases of the visual, acoustic and multimodal stimuli.

^bEmitted only during the stimulus phase.

^cLatency to respond after the stimulus onset.

^dTo account for underdispersion, overdispersion and/or zero-inflation in count response variables.

^eTo account for unbalanced number of zeros and ones in binary response variables.

and \log_{10} -transformed SVL (Appendix S3). The residual index has been used to estimate condition in amphibians (reviewed in Brodeur et al., 2020), representing a good proxy of quality because call production in ectotherms is energetically costly (Ophir et al., 2010). Positive residuals indicate males in better body condition than males with negative residuals.

We fitted two linear models using as response variables the mean DF and modulation of DF of focal males aggressive calls, with the male SVL and body condition as predictor variables. Data used for these models are based on 29 males from which DF values were extracted from the videos and do not include repeated measures of the same individual (Appendix S3). Then, we fitted GLMMs to explore the possibility that SVL and body condition influence the probability and rates of signal emission. In these models the predictor variables were the interaction between SVL and type of stimulus, and between body condition and type of stimulus, because the signalling rates in the response of focal males varied between the stimuli emitted by the robot (see Section 3). Finally, we tested whether SVL (relative to robot size) and body condition differed between males that have and have not attacked the robot using a Mann-Whitney test and a *t*-test respectively.

Question (6): to test the effects of environmental variables on male response to the stimulus phases of each experimental group, we performed GLMMs and Cox proportional hazards (CPH) models. We first fitted three GLMMs using as response variables the rates of signal emission, and as predictor variables we used background noise and \log_{10} -transformed light intensity at the signalling sites interacting with the experimental stimulus males were exposed to. Finally, we built random-effect (frailty) CPH models using as response variables the latencies to emit a signal and to move towards the robot after the onset of the experimental stimulus. As predictor variables, we used background noise and light intensity interacting with the type of stimulus. Because some males did not respond during the stimulus phase, the latency data were right censored (1 = non-respondent males, 2 = respondent males). In the GLMMs and CPH models, male identity was included as a random factor and continuous predictors were standardized and centred to zero to produce comparable effect sizes (Schielzeth, 2010).

GLMMs were built using the package GLMMTMB (Brooks et al., 2017) considering different error distributions and link functions (Table 3). After fitting a model, we tested its goodness-offit, the significance of the dispersion parameter and the presence of zero-inflation (when applicable) using the package DHARMA (Hartig, 2020). The CPH models were fitted using the package SUR-VIVAL (Therneau, 2020). For diagnostics of the CPH models, we compared the log-likelihood and AIC values using the package STATS (R Core Team, 2020). We also compared the C-index between models and checked the proportional hazards assumption using the package SURVIVAL (Therneau, 2020). Results of the diagnostics for all models are presented in Appendices S3, S5 and S7. All statistical analyses were performed in R version 4.0.2 (R Core Team, 2020).

3 | RESULTS

Focal males responded to the robot using aggressive notes, toe flags and other visual signals (see video in Appendix S4), with a higher proportion of males responding to the acoustic (*a*) and multimodal



FIGURE 2 Percentage of *Crossodactylus schmidti* males that responded with aggressive notes, toe flags, and other visual signals after the onset of visual, acoustic, and multimodal stimuli emitted by a robot frog

(a + v) stimuli. Responses via aggressive notes and other visual signals were more frequent than via toe flags (Figure 2).

3.1 | Question (1): Do *a* and *v* elicit a receiver's response on their own?

There were significant differences between experimental phases in the rate of aggressive notes ($F_{8,288} = 25.59$, p < 0.001), toe flags ($F_{8,198} = 6.376$, p < 0.001) and other visual signals emitted by focal males ($F_{8,279} = 8.333$, p < 0.001). The stimulus phase of *a* elicited higher rates of aggressive notes, toe flags and other visual signals than pre- and post-control phases (Figure 3). The stimulus phase of a + v elicited higher rates of aggressive notes than the control phases, and higher rates of toe flags and other visual signals than the pre-control phase (Figure 3). Finally, the stimulus phase of *v* did not elicit higher rates of aggressive notes than the pre-control phase and did not elicit higher rates of toe flags and other visual signals than both control phases (Figure 3). Thus, both *a* and a + vwere sufficient to elicit a receiver's signalling response, while *v* was not.



FIGURE 3 Emission rates of aggressive notes (a), toe flags (b), and other visual signals (c) by Crossodactylus schmidti males in response to visual, acoustic, and multimodal stimuli emitted by a robot frog. White dots represent the mean response estimated from the models for each experimental phase (pre-stimulus, stimulus, and post-stimulus). Vertical bars show 95% confidence intervals. Triangles in the right panels show between-phase (black, blue, and red) and between-stimuli (yellow) contrasts. Thick solid lines connecting dots indicate significant posthoc differences, dashed lines indicate non-significant differences, and the thin solid line in (b) indicates a marginally significant difference (p = 0.054)



FIGURE 4 Percentage of *Crossodactylus schmidti* males attacking a robot frog (a) in response to visual (n = 1 of 33 males), acoustic (n = 2 of 31 males) and multimodal (n = 6 of 36 males) stimuli emitted by the robot. Comparisons of (b) snout-vent length (relative to robot size) and (c) body condition between males attacking or not attacking the robot

3.2 | Question (2): Are the receiver's responses to *a* and *v* qualitatively different?

Responses of focal males to a and v were qualitatively different because the stimulus phase of a elicited more signals than the control phases, but the stimulus phase of v failed to do so (Figure 3). Moreover, the stimulus phase of a elicited higher rates of aggressive notes and toe flags than the stimulus phase of v. Thus, the receiver's responses to a and v were qualitatively and quantitatively non-redundant.

3.3 | Question (3): What are the effects of a + v on the receiver's response?

Focal males more than doubled the rate of emission of aggressive notes in the post-control phase of a + v compared to the precontrol phase, but this increase did not occur in response to a or v (Figure 3a). Compared to the stimulus phase of v, the stimulus phase of a + v elicited higher rates of emission of other visual signals, while the stimulus phase of a did not (Figure 3c). Nine of 38 focal males attacked the robot (see video in Appendix S4), most of them (n = 6) during the stimulus phase of a + v (Figure 4a). However, the frequency of attacks was not influenced by stimulus type ($\chi^2 = 4.265$, df = 2, p = 0.119).

3.4 | Question (4): Is there a correlation between *a* and *v* in the receiver's response?

During the stimulus phase of a + v, the rate of emission of aggressive notes covaried positively with the rate of emission of toe flags ($r_s = 0.56$, n = 34, p < 0.001). During the stimulus phase of a, however, the rate of emission of aggressive notes did not correlate with the rate of emission of toe flags ($r_s = 0.24$, n = 30, p = 0.202, Appendix S6). We did not test correlations between signals for the stimulus phase of v because only few males responded via toe flags.

3.5 | Question (5): Do *a* and/or *v* in the receiver's response covary with its quality?

Snout-vent length of focal males had a negative effect on the mean DF of their aggressive calls ($\beta \pm SE = -98.17 \pm 22.75$, t = -4.315, df = 24, p < 0.001, Figure 5a), with SVL explaining part of the variation in mean DF ($F_{2,24} = 9.50$, $R^2 = 0.44$, p < 0.001). SVL had a positive effect on the modulation of DF ($\beta \pm SE = 253.74 \pm 94.87$, t = 2.675, df = 24, p = 0.013, Figure 5b), with SVL explaining part of the variation in the modulation of DF ($F_{2,24} = 4.09$, $R^2 = 0.25$, p = 0.029). Body condition, in turn, had no effects on mean DF ($\beta \pm SE = -99.60 \pm 160.83$, t = -0.619, df = 24, p = 0.541, Figure 5a) and modulation of DF ($\beta \pm SE = 677.88 \pm 670.60$, t = 1.011, df = 24, p = 0.322, Figure 5b).

Mean DF and modulation of DF in aggressive calls of focal males did not differ between experimental stimuli (Appendix S3). Neither the rate nor the probability of emission of aggressive notes, toe flags and other visual signals were influenced by SVL or body condition, irrespective of the type of stimulus (Figure 6). SVL and body condition did not differ between males attacking or not attacking the robot (SVL: U = 113.5, $n_1 = 9$ and $n_2 = 29$, p = 0.57; body condition: t = -1.21, df = 36, p = 0.23, Figure 4b,c).

FIGURE 5 Relationships between two proxies of male quality (snout-vent length and body condition) and the mean dominant frequency (a) and modulation of dominant frequency (b) of aggressive calls emitted by *Crossodactylus schmidti* males in response to a robot frog. Solid and dashed lines indicate significant and non-significant effects, respectively



FIGURE 6 Forest plot of regression coefficients (β estimates) from the GLMMs fitted to estimate the effects of the interaction between two proxies of male quality (snout-vent length and body condition) and the type of stimuli (visual, acoustic and multimodal) emitted by a robot frog on the rates of emission of aggressive notes, other visual signals and toe flags by *Crossodactylus schmidti* males. Circles represent means, and horizontal lines, the 95% confidence interval. For toe flagging, we do not have estimates for the visual stimulus because few males responded to it using this signal

3.6 | Question (6): Is the probability of receiver's response, across the environmental gradients, higher to a + v?

Background noise positively affected the rate of toe flag emission in response to a + v ($\beta \pm SE = 0.544 \pm 0.227$, z = 2.389, p = 0.017, Figure 7b), while light intensity positively affected the rate of other visual signals in response to v ($\beta \pm SE = 0.419 \pm 0.189$, z = 2.219, p = 0.026, Figure 7c). The CPH models (Appendix S7) indicated that light intensity was negatively associated with the probability of a focal male to emit aggressive calls, but only in response to *a* $(\beta \pm SE = -0.448 \pm 0.214, \chi^2 = 4.36, p = 0.037; HR = 0.64, 95\%$ CI: 0.419–0.972). Pairwise comparisons of event probability curves for latency to signalling and moving towards the robot (Figure 8) indicated that the probability of response to *v* is always lower than to *a* and *a* + *v* (log-rank test, *p* < 0.05 for all response variables), and that the probability of response did not differ between *a* and *a* + *v* (logrank test, *p* > 0.05 for all response variables).



FIGURE 7 Relationships between background noise and light intensity at signalling sites of *Crossodactylus schmidti* males and rates of emission of aggressive notes (a), toe flags (b) and other visual signals (c) in response to visual (black dots and lines), acoustic (blue dots and lines), and multimodal stimuli (red dots and lines) emitted by a robot frog. Solid and dashed lines indicate significant and non-significant effects, respectively. For toe flagging, we do not have estimates for the visual stimulus because few males responded to it using this signal

4 | DISCUSSION

We used a robot frog simulating a C. schmidti male to test four hypotheses on the function of a multimodal display emitted during territorial contests. Our main results are: (a) signalling responses of focal males (i.e. territory owners) differ gualitatively and guantitatively between uni- and multimodal stimuli emitted by the robot (i.e. territory invader); (b) the multimodal stimulus elicits higher signalling rates and long-lasting responses by focal males; (c) body size influences the DF of aggressive calls; (d) body condition does not influence signalling rates or call DF; and (e) background noise and light intensity affect the emission of visual signals, but the multimodal stimulus does not increase the probability of response by focal males nor reduce their latency to respond across the environmental gradient. In what follows, we relate these results to the questions and predictions presented in Table 2 and explore the implications of our findings for understanding the function of multimodal signalling in animal contests.

The emission of aggressive calls by the robot elicited signalling responses from resident males, but the emission of toe flags did not (question 1), which implies that the responses to acoustic and visual signals are non-redundant (question 2). The emission of multimodal signals by the robot also elicited signalling responses from resident males, which responded with higher rates of visual signalling and an acoustic response that took longer to dissipate. The non-redundant responses to the unimodal signals and the increased investment in aggressive signalling (i.e. aggressive notes and other visual signals) in response to multimodal signals indicate an upward modulation (sensu Partan & Marler, 2005) in the receiver's response that occurs only when toe flags are received coupled to aggressive calls (question 3). In a similar experiment in which a robot frog was placed inside the territory of *Allobates femoralis* males, individuals showed non-redundant responses to visual and acoustic stimuli and a modulation in aggressiveness, increasing attacks to the robot when it emitted multimodal signals (Narins et al., 2003). Thus, in territorial contests, multimodal signals may be interpreted by the receiver as the willingness of the sender to escalate the contest.

The emission of aggressive notes and toe flags was positively correlated in the response of resident males to the multimodal signal emitted by the robot, but not in their response to the robot's acoustic signal (question 4). Thus, the covariance between signals emitted by resident males seems to be modulated by the type of stimulus they received from the robot. In fact, during natural contests, males first exchange long, repetitive aggressive calls, and rarely emit visual signals of any type. As the contest escalates, the emission of aggressive calls and toe flags by both contestants increases (Caldart et al., 2011, FIGURE 8 Cumulative event probability curves for the latency to emit signals (a-c) and to move towards a robot frog (d) by *Crossodactylus schmidti* males after the onset of visual (black lines), acoustic (blue lines) and multimodal stimuli (red lines) emitted by a robot frog. A vertical rise in the curves indicates that a response occurred at that time. Coloured areas indicate the 95% confidence interval for each curve. Dashed lines indicate the median latency to respond for each experimental stimulus (not given if cumulative events <50%)



2014), suggesting some level of signal matching. Examples of signal a matching in animal contests are almost entirely limited to songbirds, (where this behaviour is interpreted as an indication of aggressive intentions (e.g. Vehrencamp, 2001; Peake et al., 2005). In *C. schmidti*, thowever, the simultaneous emission of multimodal signals by the grobot and resident male did not increase the chance of attacks py the latter. We suggest that multimodal signal matching in frog the contests is a form of mutual assessment. Because the emission of the multimodal signals is apprendix of the contests is a form of mutual assessment. Because the emission of the multimodal signals are apprendix of the multimodal signals in apprendix of the multimodal signals are apprendix of the multimodals are apprendix of the multimodal signals are apprendix of the multimodal signals are apprendix of the multimodals are apprendix of the multimodal signals are apprendix of the multimodals are apprendix of the multimodal signals are apprendix of the multimodals ar

contests is a form of mutual assessment. Because the emission of multiple signals is energetically costly (e.g. Mowles et al., 2017), a contestant that emits multimodal signals may acquire information on the strength or stamina of the rival by inducing him to do the same. Likewise, competing stags of the red deer (*Cervus elaphus*) emit a costly acoustic signal, and mutually assess their fighting abilities by roaring to each other at similar rates (Clutton-Brock & Albon, 1979).

The DF of aggressive calls was inversely related to male body size (question 5), meaning that DF carries reliable information about signaller size (Searcy & Nowicki, 2005). Given that larger males have fight advantages in anurans (Dyson et al., 2013), contestants should assess the size of the rival before escalating a contest (van Staaden et al., 2011). According to content-based hypotheses, the *rate* of emission of signals should also increase with male body condition, which is related to energy reserves necessary for sustaining costly activities. However, body condition did not affect the rate of emission of aggressive notes and toe flags. Because we used only body size and condition as proxies of male quality, we cannot exclude the possibility that the rates of emission of these two signals reflect other aspects of male quality. Nevertheless, we show that the rates of emission of acoustic and visual signals do not convey independent or redundant information about male condition (question 5), which 711

Journal of Animal Ecology

are key predictions for the content-based hypotheses tested here (Table 1).

The calling activity of C. schmidti increases with light levels throughout the day and calls are partially masked by streamgenerated noise (Caldart et al., 2016a,b). If multimodal signals play an efficacy backup function, the probability of response to the multimodal signal emitted by the robot should be higher than to the isolated signals across the environmental gradient (question 6). Moreover, if signal efficacy is improved by multimodality, the latency to respond to the multimodal signal emitted by the robot should be shorter than to the unimodal signals. Contrary to these predictions, the event probability curves for the latency to responses of resident males did not differ between acoustic and multimodal signals. Interestingly, the rate of toe flags in response to the multimodal signal increased with background noise, the rate of other visual signals in response to toe flags increased with light intensity, and the probability of emission of aggressive calls in response to aggressive calls decreased with light intensity. Environmental effects on signalling behaviours have been reported for diurnal stream-breeding frogs of the genus Staurois in which males increase visual signal emission as light levels increase (Grafe & Wanger, 2007) and decrease calling activity under noise conditions (Grafe & Tony, 2017). However, as reported here for C. schmidti, experimental evidence shows that multimodal signals in Staurois do not play an efficacy-related function.

The multimodal display exhibited by *C. schmidti* males in territorial contests may be targeted by multiple selective pressures. Which hypothesis tested here better explains its function? The context hypothesis received stronger support by our data.

According to this hypothesis, the function of toe flags is to provide a new context that is used by the receiver to interpret the aggressive call and modify the response to it (Table 1). Leger (1993) defined context as the "set of events, conditions, and changeable recipient characteristics that modify the effect of a signal on recipients' behaviour". He argued that a contextual relationship between co-occurrent signals exists either when signals have a synergistic effect when combined, producing a greater change in the receiver than that produced by isolated signals, or when the combined signals generate a qualitatively different response than that evoked by either signal alone. Our results agree with Leger's definitions of context and contextual relationship, and support all predictions of the context hypothesis (Table 2). Toe flags alone (the context signal) do not elicit a receiver's response, while the aggressive call and the multimodal signal do so. Moreover, the multimodal signal increases the intensity of the receiver's response compared to the aggressive calls alone. Although these findings are compatible with predictions of the context hypothesis, we stress that alternative explanations are possible. For example, multimodality may increase the intensity of the aggressive display and, as a result, it may produce more intense responses that take longer to dissipate simply because they are more intense. This is a testable hypothesis that can be experimentally explored in future studies.

During contests, individuals are expected to extract from each other information about resource-holding potential (RHP) and aggressiveness motivation (van Staaden et al., 2011). As discussed above, DF is an index of size-related fighting ability that can be promptly assessed by the contestants. Then, if the contest escalates, adding toe flags to the acoustic signal may provide a new context to the receiver, using parameters that are arbitrary with respect to RHP (the so-called conventional signals, Guilford & Dawkins, 1995). Because the emission of toe flags increases as the contests escalate (Caldart et al., 2014), we suggest that the contextual information provided by toe flags accompanying aggressive calls is motivation to persist or to escalate. The context function of the multimodal signal also provides an explanation for the covariance between aggressive notes and toe flags in the response of resident males to the multimodal signal emitted by the robot. If resident males interpret the multimodal signal emitted by the robot as an indication of the intruder's motivation to invade the territory, and if multimodal signal matching in frog contests is a form of mutual assessment, resident males should emit multimodal signals in response to communicate their own motivation to defend the territory. Moreover, because toe flags are not a long-lasting signal, they should be repeatedly emitted to modify the meaning of the accompanied aggressive call (and the receiver's response to it), which generates an inter-signal interaction.

5 | CONCLUSIONS

A context-dependent multimodal display, composed of nonredundant signals that provide a context capable of changing the receiver's response, as the one of *C. schmidti*, seems particularly

suitable for systems in which: (i) the information conveyed by a signal changes over the course of the interaction, as occurs in graded aggressive calls of anurans, and (ii) the information contained in a signal is difficult to ascertain due to signal properties that are less stereotyped, such as anuran aggressive calls, whose temporal features (e.g. number of notes, note duration) are more variable than in advertisement calls (Wells & Schwartz, 2007). In such systems, signallers that add a contextual signal to the display may be more likely to reduce the ambiguity of the other signal. From the receivers' perspective, the exposure to both signals may also be beneficial because they acquire less ambiguous information about the signaller. We argue that contextual signals are particularly important in contests because individuals that are unbale to acquire proper information on the aggressiveness of the rival may incur injury costs. Future studies should explore how prevalent context-dependent multimodal signals are in animal contests and how individuals balance benefits and costs of multimodal signal matching. As a final remark, we stress that examples of context-dependent multimodal signals are rare in the literature (e.g. Hughes, 1996), probably because most studies focus on single hypotheses based on content- or efficacy-based selection. Our study, therefore, highlights the importance of considering multiple selective pressures when testing the function of multimodal signals.

ACKNOWLEDGEMENTS

We thank Tiago Theis, Álvaro Mainardi, Lívia Miron, Samanta lop, Maiquel Elsenbach and Maicon Elsenbach for helping in fieldwork, Sonia Cechin for lending field equipment, Solimary García-Hernández and André Rodrigues for assistance in statistical analyses, and Eileen Hebets, Alexandre Palaoro and two reviewers for suggestions on the manuscript. V.M.C. was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (2015/14336-6). G.M. is supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (306550/2014-8) and Fundação de Amparo à Pesquisa do Estado de São Paulo (2021/00915-5). This project was authorized by the license #21/2016 and authorization #173/2016 for temporary marking of individuals, provided by the Divisão de Unidades de Conservação da Secretaria do Ambiente e Desenvolvimento Sustentável do Estado do Rio Grande do Sul. No individual was collected or held in captivity during this study.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

AUTHORS' CONTRIBUTIONS

V.M.C. and G.M. conceived the study and designed methodology; V.M.C. and M.B.S. built the robot; V.M.C. collected and analysed the data; V.M.C. wrote the first version of the manuscript. All authors revised the statistical analyses and multiple versions of the manuscript. All authors gave final approval for publication.

DATA AVAILABILITY STATEMENT

Data available from the Dryad Digital Repository https://doi. org/10.5061/dryad.41ns1rnfd (Caldart et al., 2021).

ORCID

Vinícius Matheus Caldart ២ https://orcid.

org/0000-0002-1140-8189

Mauricio Beux dos Santos ២ https://orcid.

org/0000-0002-0555-2037

Glauco Machado (D https://orcid.org/0000-0002-0321-128X

REFERENCES

- Brodeur, J. C., Damonte, M. J., Candioti, J. V., Poliserpi, M. B., D'Andrea, M. F., & Bahl, M. F. (2020). Frog body condition: Basic assumptions, comparison of methods and characterization of natural variability with field data from *Leptodactylus latrans. Ecological Indicators*, 112, 106098. https://doi.org/10.1016/j.ecolind.2020.106098
- Bro-Jørgensen, J., & Dabelsteen, T. (2008). Knee-clicks and visual traits indicate fighting ability in eland antelopes: Multiple messages and back-up signals. BMC Biology, 6, 47. https://doi. org/10.1186/1741-7007-6-47
- Brooks, M. E., Kristensen, K., Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., & Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zeroinflated generalized linear mixed modeling. *The R Journal*, *9*, 378– 400. https://doi.org/10.32614/RJ-2017-066
- Caldart, V. M., dos Santos, M. B., & Machado, G. (2021). Data from: Function of a multimodal signal: A multiple hypothesis test using a robot frog. *Dryad Digital Repository*, https://doi.org/10.5061/dryad.41ns1rnfd
- Caldart, V. M., Iop, S., & Cechin, S. Z. (2011). Vocalizations of Crossodactylus schmidti Gallardo, 1961 (Anura, Hylodidae): Advertisement call and aggressive call. North-Western Journal of Zoology, 7, 118–124.
- Caldart, V. M., Iop, S., & Cechin, S. Z. (2014). Social interactions in a neotropical stream frog reveal a complex repertoire of visual signals and the use of multimodal communication. *Behaviour*, 151, 719–739.
- Caldart, V. M., Iop, S., Lingnau, R., & Cechin, S. Z. (2016a). Calling activity of a stream-breeding frog from the austral Neotropics: Temporal patterns of activity and the role of environmental factors. *Herpetologica*, 72, 90–97. https://doi.org/10.1655/HERPETOLOG ICA-D-15-00029
- Caldart, V. M., Iop, S., Lingnau, R., & Cechin, S. Z. (2016b). Communication in a noisy environment: Short-term acoustic adjustments and the underlying acoustic niche of a Neotropical stream-breeding frog. *Acta Ethologica*, *19*, 151–162. https://doi.org/10.1007/s1021 1-016-0235-2
- Caldart, V. M., Loebens, L., Brum, A. J. C., Bataioli, L., & Cechin, S. Z. (2019). Reproductive cycle, size and age at sexual maturity, and sexual dimorphism in the stream-breeding frog *Crossodactylus* schmidti (Hylodidae). South American Journal of Herpetology, 14, 1-11. https://doi.org/10.2994/SAJH-D-17-00060.1
- Clutton-Brock, T. H., & Albon, S. D. (1979). Roaring of red deer and the evolution of honest advertisement. *Behaviour*, *69*, 145–170.
- Dyson, M. L., Reichert, M. S., & Halliday, T. R. (2013). Contests in amphibians. In I. C. W. Hardy, & M. Briffa (Eds.), Animal contests (pp. 228-257). Cambridge University Press.
- Furtado, R., Lermen, L. N., Márquez, R., & Hartz, M. S. (2019). Neotropical dancing frog: The rich repertoire of visual displays in a hylodine species. *Journal of Ethology*, 37, 291–300. https://doi.org/10.1007/ s10164-019-00600-x
- Genz, A., Bretz, F., Miwa, T., Mi, X., Leisch, F., Scheipl, F., & Hothorn, T. (2020). mvtnorm: Multivariate normal and t distributions. R package version 1.1-0. https://CRAN.R-project.org/package= mvtnorm

- Grafe, T. U., & Tony, J. A. (2017). Temporal variation in acoustic and visual signalling as a function of stream background noise in the Bornean foot-fagging frog, *Staurois parvus. Journal of Ecoacoustics*, 1, #X74QE0.
- Grafe, T. U., & Wanger, T. C. (2007). Multimodal signaling in male and female foot-flagging frogs *Staurois guttatus* (Ranidae): An alerting function of calling. *Ethology*, 113, 772–781. https://doi. org/10.1111/j.1439-0310.2007.01378.x
- Guilford, T., & Dawkins, M. S. (1995). What are conventional signals? Animal Behaviour, 49, 1689–1695. https://doi. org/10.1016/0003-3472(95)90090-X
- Hartig, F. (2020). DHARMa: Residual diagnostics for hierarchical (multilevel/mixed) regression models. R package version 0.2.7. https:// CRAN.R-project.org/package=DHARMa
- Hartmann, M. T., Giasson, L. O. M., Hartmann, P. A., & Haddad, C. F. B. (2005). Visual communication in Brazilian species of anurans from the Atlantic forest. *Journal of Natural History*, *39*, 1675–1685. https://doi.org/10.1080/00222930400008744
- Hebets, E. A., & Papaj, D. R. (2005). Complex signal function: Developing a framework of testable hypotheses. *Behavioral Ecology and Sociobiology*, 57, 197–214. https://doi.org/10.1007/s0026 5-004-0865-7
- Hughes, M. (1996). The function of concurrent signals: Visual and chemical communication in snapping shrimp. *Animal Behavior*, *52*, 247-257.
- Leger, D. W. (1993). Contextual sources of information and responses to animal communication signals. *Psychological Bulletin*, 113, 295–304.
- Lenth, R. (2019). emmeans: Estimated marginal means, aka least-squares means. R package version 1.4. https://CRAN.R-project.org/packa ge=emmeans
- Mowles, S. L., Jennions, M., & Backwell, P. R. Y. (2017). Multimodal communication in courting fiddler crabs reveals male performance capacities. *Royal Society Open Science*, 4, 161093. https://doi. org/10.1098/rsos.161093
- Narins, P. M., Hödl, W., & Grabul, D. S. (2003). Bimodal signal requisite for agonistic behavior in a dart-poison frog, *Epipedobates femoralis*. *Proceedings of the National Academy of Sciences of the United States* of America, 100, 577–580. https://doi.org/10.1073/pnas.02371 65100
- Ophir, A. G., Schrader, S. B., & Gillooly, J. F. (2010). Energetic cost of calling: General constraints and species-specific differences. *Journal of Evolutionary Biology*, 23, 1564–1569. https://doi. org/10.1111/j.1420-9101.2010.02005.x
- Partan, S., & Marler, P. (2005). Issues in the classification of multimodal communication signals. The American Naturalist, 166, 231–245. https://doi.org/10.1086/431246
- Peake, T. M., Matessi, G., McGregor, P. K., & Dabelsteen, T. (2005). Song type matching, song type switching and eavesdropping in male great tits. *Animal Behaviour*, 69, 1063–1106. https://doi. org/10.1016/j.anbehav.2004.08.009
- R Core Team. (2020). R: A language and environment for statistical computing. v.4.0.2. R Foundation for Statistical Computing.
- Schielzeth, H. (2010). Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution*, 1, 103–113. https://doi.org/10.1111/j.2041-210X.2010.00012.x
- Searcy, W. A., & Nowicki, S. (2005). The evolution of animal communication: Reliability and deception in signaling systems. Princeton University Press.
- Therneau, T. (2020). A package for survival analysis in R. R package version 3.2-7. https://CRAN.R-project.org/package=survival
- Uetz, G. W., Clark, D. L., & Roberts, J. A. (2016). Multimodal communication in wolf spiders (lycosidae): An emerging model for study. Advances in the Study of Behavior, 48, 117–159.

- van Staaden, M. J., Searcy, W. A., & Hanlon, R. T. (2011). Signaling aggression. In R. Huber, D. L. Bannasch, & P. Brennan (Eds.), Advances in genetics (Vol. 75, pp. 23–49). Academic Press.
- Vehrencamp, S. L. (2001). Is song-type matching a conventional signal of aggressive intentions? Proceedings of the Royal Society of London B: Biological SciencesB, 268, 1637–1642. https://doi.org/10.1098/ rspb.2001.1714
- Wells, K. D., & Schwartz, J. J. (2007). The behavioral ecology of anurancommunication. In P. M. Narins, A. S. Feng, R. R. Fay, & A. N. Popper (Eds.), *Hearing and sound communication in amphibians*. Springer handbook of auditory research book series (Vol.28, pp. 44–86). Springer. https://link.springer.com/chapter/10.1007/978-0-387-47796-1_3

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Caldart, V. M., dos Santos, M. B., & Machado, G. (2022). Function of a multimodal signal: A multiple hypothesis test using a robot frog. *Journal of Animal Ecology*, 91, 701–714. https://doi.org/10.1111/1365-2656.13620