Animal sexual displays are typically repeated over time and consist of components that are also repeated within a display, creating potential for within-individual variation in signal production. Across taxa, patterns of variation in and female preferences for temporal properties of signals are well documented, but data describing how within-individual variation functions in communication are scarce. Is within-individual variation itself a signal of male quality, or noise that obscures another signal encoded by a temporal pattern? In this study of Cope’s gray treefrog, *Hyla chrysoscelis*, we used synthetic advertisement calls to examine the function of within-individual variation in two signal traits: pulse rate, a signal of species identity, and call rate, a signal of male quality. We examined relationships between male body condition and coefficients of variation within males for both signal traits. We then measured female preferences for within-individual variation. Because treefrogs communicate in noisy social aggregations, we repeated the experiment in quiet and at three amplitudes of chorus-shaped noise to evaluate how within-individual variation affects mate choice in natural settings. Within-individual variation in signal production was not predicted by male condition and likely acts as noise in this communication system. Females strongly discriminated against highly variable call rates, but not pulse rates, when the mean trait values of signals were taken into account. At its highest level, chorus noise abolished this effect, suggesting female preferences against within-individual variation are unlikely to be expressed in dense social aggregations.

**Key words:** animal communication, mate choice, noise, sexual selection, signaling, within-individual variation.

**INTRODUCTION**

Animal signals often consist of repeated elements, such that the rate of repetition of the whole display or its components may encode information that is important to receivers. In many animal systems, receivers exert selection on signalers and signal traits through mate choice (Andersson 1994), and temporal patterning of sexual displays is frequently a mate choice criterion (Ryan and Keddy-Hector 1992). For example, the rate of jumping displays is one of the best predictors of male mating success in Jackson’s widowbird (Andersson 1989). In the wolf spider *Schizocosa ocreata*, females attend to multiple male courtship traits, including the performance rate of components termed percussive strikes, cheliceral strikes, and double-taps (Delauney et al. 2007; Gibson and Uetz 2008). Females of several species in the firefly genus *Photinus* prefer males that emit bioluminescent flash signals with faster pulse rates (Demary et al. 2006; Lewis and Cratsley 2008).

Because many signals are repeated and consist of repeated elements, there is potential for within-individual variation in signal production. Significant within-individual variation in multi-component signals has been described in diverse taxa, including amphibians (Howard and Young 1998), reptiles (Crews 1975), and mammals (Mitani and Brandt 1994), and may be used to describe signal components along a continuum of “static” to “dynamic” (Gerhardt 1991; Shaw and Herlihy 2000). Static components vary little within individuals, while “dynamic” components may vary markedly within individuals, even over short periods of time, such as a single displaying bout. Despite documentation of within-individual variation in male signals, corresponding data about its role in communication systems are scarce (Gerhardt and Watson 1995). Experiments investigating mate choice, for example, generally reduce the dimensionality of multicomponent signals in order to determine the effect of a single variable, which is almost always presented as a sequence of invariant signals, that is, a series of displays in which every presentation is identical.

Eliminating within-individual variation in experimental stimuli has two potential consequences. First, it may cause us to overlook biologically relevant signals. Within-individual variation could

**Address correspondence to:** J.C. Tanner. E-mail: jessie.c.tanner@gmail.com.

Present Address: Centre for Evolutionary Biology, University of Western Australia, Crawley, WA 6009, Australia

© The Author(s) 2018. Published by Oxford University Press on behalf of the International Society for Behavioral Ecology. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

**Within-individual variation in sexual displays: signal or noise?**

Jessie C. Tannera,⁎ and Mark A. Beea,b,⁎

⁎Department of Ecology, Evolution, and Behavior, University of Minnesota, Twin Cities, 1479 Gortner Ave, St. Paul, MN 55108, USA and bGraduate Program in Neuroscience, University of Minnesota, Twin Cities, 321 Church Street SE, Minneapolis, MN 55455, USA

Received 26 June 2018; revised 8 November 2018; editorial decision 8 November 2018; accepted 13 November 2018; Advance Access publication 7 December 2018.
act as a signal of male quality, for instance, because consistently producing repeated elements accurately or rapidly may be energetically expensive or difficult to perform well (motor performance hypothesis; Podos 1997; Byers et al. 2010; but see Kroodsma 2017). Second, a failure to consider within-individual variation may cause us to chronically overestimate the strength of sexual selection on signals if within-individual variation simply acts as “noise” in the communication channel. In this sense, within-individual variation potentially presents a signal detection theory problem for receivers by introducing uncertainty that interferes with signal detection or discrimination (Shannon 1948; Brumm and Slabbekoorn 2005; Wiley 2015). By observing a signaler producing a sequence of displays, a receiver is essentially sampling from a distribution in order to estimate a signaler’s mean trait value. Higher within-individual variation signals are analogous to broader distributions. From the receiver’s perspective, higher within-individual variability in signal production could introduce uncertainty that obscures their estimate of trait means in temporally patterned signals. Consequently, the uncertainty introduced by within-individual variation could effectively shield male phenotypes from selection imposed by receivers. In this sense, within-individual variation would act like other kinds of noise to weaken sexual selection by preventing the expression of existing female mating preferences (Bee 2015; Reichert and Ronacher 2015).

Treefrogs are an excellent system in which to investigate how within-individual variation affects signaling for several reasons. The repeated production of advertisement calls in treefrogs is one of the most energetically expensive activities measured in vertebrates (Taigen and Wells 1985; Wells and Taigen 1986). Male body condition decreases with continued calling effort and chorus attendance (Murphy 1994). The number of nights that males attend a chorus is the best predictor of male mating success (Rikke and Semlitsch 1991; Sullivan and Hinshaw 1992), and there is some evidence that males whose body condition declines more slowly with chorus attendance have a higher probability of mating on any given night (Murphy 1994). Humfeld (2013) showed that males who lose body condition over short periods of time consequently produce relatively less-attractive calls and are more likely to adopt a satellite strategy instead of calling. Additionally, males who produce long call durations (one aspect of calling effort) can be of higher genetic quality, producing offspring with higher viability and faster growth rates than those of their short-calling conspecifics (Welch et al. 1998; Welch 2003). Taken together, these observations make within-individual variation in male call production a strong candidate for a signal of a male’s ability to consistently perform energetically expensive tasks repeatedly, an ability sometimes called male “vigor” (Darwin 1859; Darwin 1871; Sutton 1981; Sílves 1982; Byers et al. 1994, 2010). In addition, frogs exhibit indeterminate growth and likely breed for multiple years. We know little about the potential for behavioral or developmental plasticity to effect signal production in frogs, but there is some evidence that early acoustic experience mediates male signal production later in life (Dawson and Ryan 2009). Thus, plasticity in within-individual variation in advertisement calling across an individual’s lifetime might be reflected in a relationship between within-individual variation and measures of male body size.

Alternatively, it is possible that within-individual variation in anuran vocalizations simply obscures the between-individual variation that is the true target of mate choice. The advertisement calls produced by male treefrogs are well-known targets of sexual selection by female mate choice (Murphy et al. 2000; Schwartz et al. 2001; Gerhardt 2005; Gerhardt and Brooks 2009; Murphy 2012; Schröde et al. 2012; Tanner et al. 2017). The acoustic properties of male calls vary within individuals along the continuum between static and dynamic (Gerhardt 1991). Females commonly exhibit robust stabilizing preferences for static call properties and directional preferences for dynamic call properties (Gerhardt 1991; Ryan and Kelty-Hector 1992; Gerhardt and Huber 2002; Schul and Bush 2002; Gerhardt and Brooks 2009). Some differences between preferred and non-preferred call properties probably approach the minimum absolute differences that can be resolved by the female’s auditory system (Gerhardt and Bee 2007). Moreover, female treefrogs sample calling males over time periods that far exceed the much shorter time intervals associated with repeated calls and repeated elements within calls (Schwartz et al. 2004). Hence, there is considerable potential for within-individual variation to introduce uncertainty in a female’s estimate of the trait values on which mate choice preferences are based (Gerhardt and Watson 1995).

We investigated whether within-individual variation functions as a signal of male quality or acts as noise in Hyla chrysoscelis, Cope’s gray treefrog. Like most other anurans, H. chrysoscelis relies upon multicomponent acoustic signals as the primary means of mate attraction, discrimination, and localization (Gerhardt and Huber 2002; Tanner et al. 2017). A single call consists of a repeated series of pulses, which typically number about 30. Calls are repeated every several seconds over a period of minutes to form a calling bout. The patterns of trait variation within and between males and female preferences for male signal traits are well understood in H. chrysoscelis (Klump and Gerhardt 1987; Gerhardt 1991; Gerhardt et al. 1996; Ward et al. 2013b; Tanner et al. 2017). The repetition rate of pulses within a call (pulse rate) is a static property that varies little within individuals and signals species identity; the repetition rate of calls within a bout (call rate) is a dynamic property that is highly variable within individuals and potentially signals male quality. Within the range of natural variation for both traits, females show threshold-shaped preferences, in which individuals discriminate against lower-than-average trait values more strongly than they express preferences for higher-than-average trait values (Tanner et al. 2017). Treefrog receivers readily respond to synthetic signals, which allowed us to independently manipulate the coefficients of variation within males (CV$_m$: the standard deviation scaled by the mean) for both pulse rate and call rate while controlling other components of the signal.

In order to understand how within-individual variation functions in natural signaling contexts, we must also consider how the environment might affect the expression of female mating preferences. Cope’s gray treefrog breeds in a resource-based lek, and males must compete to be heard over a loud chorus of competing conspecifics and co-occurring heterospecifics that are also producing advertisement calls. The sound produced by anuran choruses limits the ability of receivers to discriminate among calls on the basis of biologically relevant traits and therefore also acts as noise (Bee 2008; Vélez and Bee 2011; Ward et al. 2013a; Lee et al. 2017). Realistically, complex signaling environments thus not only include variation within and between males but also background noise with spectral features that overlap those of natural male signals.

Here, we examined two alternative, but not mutually exclusive, hypotheses: within-individual variation in male sexual displays 1) signals male quality or 2) obscures biologically meaningful information conveyed by pulse rate and call rate. First, we examined the signals and signalers to test the prediction that, if within-individual variation signals male ability to consistently produce high-effort displays, it should be negatively correlated with male body condition, such that individuals in better condition produce less variable sexual displays.
We also examined relationships between body size and within-individual variation to determine whether larger males are better able to produce consistent displays. Second, we manipulated the amount of within-individual variation in sequences of synthetic acoustic signals and measured female preferences across the range of natural variation. We predicted that receivers should discriminate on the basis of within-individual variation even when accounting for differences in the mean trait values of two alternatives. Importantly, receivers might be expected to discriminate against highly variable call traits regardless of whether or not within-individual variation functions as a signal because within-individual variation makes trait values harder to estimate. Finally, to determine the extent to which within-individual variation is likely to affect the outcome of mating decisions in natural communication contexts, we repeated our experiment in quiet and at three playback amplitudes of background noise designed to simulate a treefrog chorus. We predicted that, if receivers discriminated on the basis of within-individual variation, the expression of this preference would be impaired in high levels of background noise.

**METHODS**

All experiments described here met the Animal Behavior Society/Association for the Study of Animal Behaviour guidelines for the ethical treatment of animals. The protocol was approved by the University of Minnesota Institutional Animal Care and Use Committee under protocol No. 1701-34456A.

**Measuring coefficients of variation within males, body condition, and body size**

We used a published dataset of 1000 calls temperature-corrected to 20 °C, to which each of 50 male *H. chrysoscelis* contributed a sequence of 20 consecutive calls. Details of how these acoustic recordings were made and analyzed are provided by Ward et al. (2013b). The coefficient of variation within males (CV$_w$) with respect to pulse rate and call rate was calculated for each individual male (Ward et al. 2013b). For a subset of 44 of these 50 males (880 calls total), snout-vent lengths (SVLs) were measured to the nearest 0.1 mm using dial calipers, and body masses were measured to the nearest 0.1 g using Pesola spring scales. The mean (±SD) SVL and mass were, respectively, 38.8 ± 2.4 mm (N = 44, range: 33.6–43.5 mm) and 4.3 ± 0.9 g (N = 44, range: 2.8–5.9 g (Ward et al. 2013b). For this subset of males, we calculated a body condition index for each male by dividing the residuals from a linear regression of the cube root of mass on SVL by SVL (Baker 1992; Howard et al. 1997; Howard and Young 1998; Bee et al. 2000). The goal of calculating body condition indices is generally to control for absolute size in order to compare individuals on the basis of relative mass (Jakob et al. 1996). Individuals said to be in good condition are presumed to have larger amounts of metabolizable tissue and water than individuals that are relatively lighter (Schulte-Hostedde et al. 2001). We calculated Pearson’s correlation coefficients to describe the relationship between SVL and mass of calling males, as well as the relationships between coefficients of variation and raw body condition scores, mass, and SVL. All statistical tests were performed using a significance criterion of 0.050.

**Measuring female preference functions**

**Subjects**

Subjects for phonotaxis tests were 80 gravid female *H. chrysoscelis* of the western genetic lineage (Placek et al. 1994) captured in amplexus at night (2200–0100 h) between 4 and 15 June 2017. Female treefrogs caught in amplexus are no less discriminating than gravid females caught immediately prior to mate choice (Murphy and Gerhardt 1996). Capture sites in east-central Minnesota, USA, included Carver Park Reserve (44.52490, −93.43031; Carver County), Lake Maria State Park (45.32012, −93.94389; Wright County), Hyland Lake Park Reserve (44.84214, −93.37148; Hennepin County), and Crow-Hassan Park Reserve (45.19471, −93.65368; Hennepin County).

Amplexant pairs were housed in plastic containers and taken to the laboratory, where they were placed in aged tap water and maintained at approximately 2 °C for up to 48 h to prevent egg deposition. Prior to testing, treefrogs were placed in room-temperature water and housed in an incubator for at least 30 min, until they reached a body temperature of 20 ± 1 °C. Temperature control is important for studies of communication in frogs because both male call properties and female preferences are temperature-dependent (Gerhardt and Huber 2002). Between behavioral trials, we housed females with their original mates in the incubator for a break that lasted a minimum of 3 min. All treefrogs were tested and returned to their capture locations within 3 days.

**Experimental design and acoustic stimuli**

We used a within-subjects, 4 × 4 factorial design in which each female was tested in up to 16 behavioral trials (n$_{obs} = 80$; n$_{sex}$ = 1103) to generate preference functions, which describe the strength of preferences across a continuum of male phenotypes. We measured preference functions for pulse rate and call rate separately in different subjects (N = 40 each). Levels of within-individual variation for alternative signals (“low,” “mean,” or “high”; Figure 1) were performed using a significance criterion of 0.050.
were chosen to approximate the minimum, mean, and maximum coefficients of variation within males (CV\textsubscript{w}) measured for call rate and pulse rate by Ward et al. (2013b).

For each level of within-individual variation, subjects completed a two alternative choice test between an invariant, “standard” call and an “alternative” that was variable with respect to the focal trait of either pulse rate or call rate. Each acoustic stimulus was designed to simulate a male producing a sequence of calls that lasted the duration of a phonotaxis test. Each individual female was tested with a set of three unique alternative stimuli, one for each level of within-individual variation ("low", "mean," "high"); Figure 1; that is, the total number of alternative stimuli used in phonotaxis trials was equal to three times the sample size (n\textsubscript{subjects} = 80; n\textsubscript{stimuli} = 240). Importantly, the two stimuli in a phonotaxis test always had the same nominal mean values of the specified focal trait and differed only in the level of within-individual variation in the focal trait. An additional control condition gave each female the choice between two identical, standard calls. In this condition, one signal was randomly and arbitrarily labeled the “alternative” call prior to testing so that a choice could be scored. These four tests (control plus three levels of CV\textsubscript{w}) were replicated for each female at each of four levels of chorus-shaped noise.

Synthetic calls similar to those described by Tanner et al. (2017) were generated de novo using custom scripts in MATLAB version 2016a (The MathWorks, Natick, MA, USA). All call traits not under manipulation were fixed at the population mean values published by Ward et al. (2013b). Standard stimuli consisted of sequences of calls that had invariant, mean trait values at every presentation. The standard pulse rate was 48.8 pulses per second (pps). The standard call rate of 11.15 calls per minute (cpm) was output through a MOTU model 16A 16-channel sound card (MOTU, Inc., Cambridge, MA, USA) and amplified with Crown XLS1000 High-Density Power Amplifiers (HARMAN Professional, Northridge, CA, USA). Two Mod 1 Orb speakers (Orb Audio, New York, NY, USA) were used to broadcast stimuli.

We calibrated synthetic signals to a maximum sound pressure level (SPL) of 85 dB (re 20 \textmu P, fast RMS, C-weighted) at 1 m from the playback speaker and the approximate height (3 cm) of a female’s head before release. Chorus-shaped noise was co-located with signal presentation. Each of the two speakers played the same chorus-shaped noise replicate at the same time. Noise because they approximated the mean amplitude generated by a natural chorus, plus and minus one standard deviation.

**Experimental apparatus and protocol**

Stimuli were broadcast using Adobe Audition 3.0 (Adobe Systems, Inc., San Jose, CA, USA) and a Dell Optiplex 980 PC (Dell Computer Corporation, Round Rock, TX, USA). Audio output was through a MOTU model 16A 16-channel sound card (MOTU, Inc., Cambridge, MA, USA) and amplified with Crown XLS1000 High-Density Power Amplifiers (HARMAN Professional, Northridge, CA, USA). Two Mod 1 Orb speakers (Orb Audio, New York, NY, USA) were used to broadcast stimuli.

We conducted phonotaxis trials at 20 ± 1 °C in a custom-built, temperature-controlled, semi-anechoic chamber (2.8 m × 2.3 m × 2.1 m, length × width × height; IAC Acoustics, Winchester, UK). The ceiling and walls of the sound chamber were acoustically insulated and covered in dark gray, perforated metal (IAC Planarchoic™ panel system). The chamber floor was covered with dark gray, low-pile carpet. Within the chamber, we constructed a circular testing arena (2 m × 0.6 m, diameter × height) from hardware cloth covered in black fabric. In the center of the arena, we placed an acoustically transparent, circular release cage (9 cm × 2 cm, diameter × height) on the floor. The distance between the center of the release cage and each speaker was 1 m. The lid of the release cage could be operated via a rope and pulley system by an...
observer outside the sound chamber. Speakers were located on the floor of the chamber outside the circular arena wall and positioned 90° apart. We conducted all phonotaxis trials under infrared light (Tracksys, Ltd., Nottingham, UK) and observed the subject from outside the sound chamber by means of an infrared-sensitive video camera (Panasonic WV-BP334; Panasonic Corporation of North America, Secaucus, NJ, USA) mounted from the ceiling, directly above the release cage, and a closed-circuit television system. Trials were scored in real time.

Between subjects, we randomized the trial order; the speaker assignments for the standard and alternative signals, and whether the standard or alternative signal was broadcasted first. At the beginning of each trial, the subject was separated from her mate and placed in the acoustically transparent release cage, where she was permitted to acclimate in quiet for 60 s. At the end of the acclimation period, we began playback. The standard and alternative signals were initially presented with an alternating and non-overlapping temporal arrangement but were subsequently timed relative to each other according to their respective call rates, such that calls continued to be non-overlapping when call rate was invariant for both signals.

When both the standard and alternative stimuli had been presented to the subject at least twice, the lid of the release cage was lifted, and the female was allowed to move freely within the arena. In all tests, the trials ended when one of the following conditions was met: 1) the subject indicated a choice by entering a response zone described by a 10-cm radius semi-circle in front of a speaker, 2) the subject’s first contact with the arena wall occurred in the quadrant opposite the quadrant separating the two playback speakers, 3) at the end of 3 min, the subject had not left the release cage, or 4) at the end of 5 min, the subject did not indicate a choice. For each trial, we recorded a female’s binary choice and her latency to respond, measured as the time that elapsed between release and crossing into the response zone. If a choice was not indicated during the trial, the subject was given a score of “no response,” returned to the incubator for at least 3 min and then retested. Subjects that scored no response twice for the same test (16 out of 90) were not tested further. The tests that these individuals had previously completed were included in analyses. For the 16 subjects that completed fewer than all 16 tests, the mean number of tests completed was 7 (range: 1–15 tests; median: 6.5 tests).

**Statistical analyses**

In phonotaxis tests, the trait value of each call in an alternative stimulus sequence was drawn at random from a normal distribution, and therefore, the realized coefficient of variation experienced during a trial was variable and dependent upon the response latency of the subject. Trials with longer response latencies were more likely to match the population mean trait value and nominal latency of the subject. Trials with longer response latencies were also not predicted by male body condition (r = −0.047, P = 0.761; Figure 2D), or SVL (r = 0.077, P = 0.619; Figure 2C). Coefficients of variation in call rate were also not predicted by male body condition (r = −0.047, P = 0.761; Figure 2D), or SVL (r = 0.026, P = 0.866; Figure 2F). Body condition is by definition not correlated with SVL because it is a size-corrected mass. Male body mass was positively correlated with SVL (r = 0.865, P < 0.001; N = 44).

**RESULTS**

**Male body condition and within-individual variability in signal production**

Coefficients of variation in pulse rate were not predicted by measurements of male body condition (r = −0.224, P = 0.144; Figure 2A), mass (r = −0.046, P = 0.768; Figure 2B), or SVL (r = 0.077, P = 0.619; Figure 2C). Coefficients of variation in pulse rate were not predicted by male body condition (r = −0.047, P = 0.761; Figure 2D), or SVL (r = 0.026, P = 0.866; Figure 2F). Body condition is by definition not correlated with SVL because it is a size-corrected mass. Male body mass was positively correlated with SVL (r = 0.865, P < 0.001; N = 44).

**Female preferences for pulse rate variability**

The realized within-individual variation (CVw) in the variable pulse rate of the alternative stimulus did not predict the probability of choosing the alternative stimulus over the invariant standard stimulus (Table 1; Figure 3A) and had only a marginally significant effect on the latency to respond (β = 638.34, t = 1.78, P = 0.075; Table 1; Figure 3B). Within the range of natural variation, there was no consistent effect of the realized mean pulse rate of the alternative on the probability of response (Figure 3C). There was a small range of variation in realized mean alternative pulse rate, thus the chaotic visual trends in the loess-smoothed data in Figure 3C are explained by data points on the tails of the distribution. However, there was a quadratic effect of the realized mean pulse rate of the alternative on response latency (β = −47.82, t = −3.02, P = 0.003; Figure 3D), with the maximum response latency occurring when the mean alternative pulse rate was slightly higher than the standard pulse rate of 48.8 pps (Figure 3D). That is, females spent the most time making decisions when the realized mean pulse rate of the variable alternative stimulus was, on average, slightly higher than that of the invariant, standard stimulus, and less time when the standard and alternative stimuli had more different pulse rates. Chorus-shaped noise had no significant effect on the probability of choosing the alternative at any level (Table 1). However, at the highest level of chorus-shaped noise, females had longer response latencies (β = 18.12, t = 2.26, P = 0.024; Figure 3B, D), suggesting that the noise of a natural chorus on nights of particularly high
Table 1

Output from two GLMM examining the outcome of phonotaxis experiments in which we manipulated within-individual variation in pulse rates (532 observations of 40 individuals). Significant model terms (α = 0.05) are shown in bold.

<table>
<thead>
<tr>
<th>Response variable</th>
<th>Term</th>
<th>β</th>
<th>Std error</th>
<th>df</th>
<th>T</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(Alternative)</td>
<td>Intercept</td>
<td>0.55</td>
<td>0.06</td>
<td>482</td>
<td>8.98</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Alternative mean</td>
<td>0.17</td>
<td>0.28</td>
<td>482</td>
<td>0.58</td>
<td>0.561</td>
</tr>
<tr>
<td></td>
<td>Alternative mean quadratic</td>
<td>0.09</td>
<td>0.15</td>
<td>482</td>
<td>0.59</td>
<td>0.558</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$</td>
<td>-0.25</td>
<td>3.39</td>
<td>482</td>
<td>-0.07</td>
<td>0.946</td>
</tr>
<tr>
<td></td>
<td>Chorus-shaped noise at 60 dB</td>
<td>0.07</td>
<td>0.08</td>
<td>482</td>
<td>0.86</td>
<td>0.391</td>
</tr>
<tr>
<td></td>
<td>Chorus-shaped noise at 70 dB</td>
<td>-0.02</td>
<td>0.08</td>
<td>482</td>
<td>-0.21</td>
<td>0.835</td>
</tr>
<tr>
<td></td>
<td>Chorus-shaped noise at 80 dB</td>
<td>0.08</td>
<td>0.08</td>
<td>482</td>
<td>0.95</td>
<td>0.341</td>
</tr>
<tr>
<td></td>
<td>Alternative mean × Alternative $CV_w$</td>
<td>-2.90</td>
<td>11.46</td>
<td>482</td>
<td>-0.25</td>
<td>0.800</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$ × Chorus-shaped noise at 60 dB</td>
<td>-4.82</td>
<td>5.04</td>
<td>482</td>
<td>-0.96</td>
<td>0.340</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$ × Chorus-shaped noise at 70 dB</td>
<td>1.53</td>
<td>5.07</td>
<td>482</td>
<td>0.30</td>
<td>0.763</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$ × Chorus-shaped noise at 80 dB</td>
<td>-3.76</td>
<td>4.90</td>
<td>482</td>
<td>-0.75</td>
<td>0.451</td>
</tr>
<tr>
<td>Latency</td>
<td>Intercept</td>
<td>76.77</td>
<td>7.39</td>
<td>482</td>
<td>10.39</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Alternative mean</td>
<td>-2.28</td>
<td>29.46</td>
<td>482</td>
<td>-0.08</td>
<td>0.938</td>
</tr>
<tr>
<td></td>
<td>Alternative mean quadratic</td>
<td>-47.82</td>
<td>15.84</td>
<td>482</td>
<td>-3.02</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$</td>
<td>638.34</td>
<td>538.17</td>
<td>482</td>
<td>1.78</td>
<td>0.075</td>
</tr>
<tr>
<td></td>
<td>Chorus-shaped noise at 60 dB</td>
<td>-0.29</td>
<td>7.99</td>
<td>482</td>
<td>-0.04</td>
<td>0.972</td>
</tr>
<tr>
<td></td>
<td>Chorus-Shaped Noise at 70 dB</td>
<td>-1.51</td>
<td>7.97</td>
<td>482</td>
<td>-0.19</td>
<td>0.850</td>
</tr>
<tr>
<td></td>
<td>Chorus-shaped noise at 80 dB</td>
<td>18.12</td>
<td>8.02</td>
<td>482</td>
<td>2.26</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>Alternative mean × Alternative $CV_w$</td>
<td>138.47</td>
<td>1154.00</td>
<td>482</td>
<td>0.12</td>
<td>0.907</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$ × Chorus-shaped noise at 60 dB</td>
<td>241.52</td>
<td>502.12</td>
<td>482</td>
<td>0.48</td>
<td>0.631</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$ × Chorus-Shaped Noise at 70 dB</td>
<td>-63.10</td>
<td>504.82</td>
<td>482</td>
<td>-0.12</td>
<td>0.901</td>
</tr>
<tr>
<td></td>
<td>Alternative $CV_w$ × Chorus-shaped noise at 80 dB</td>
<td>-90.11</td>
<td>496.84</td>
<td>482</td>
<td>-0.18</td>
<td>0.856</td>
</tr>
</tbody>
</table>

Figure 2

Coefficients of variation in (A–C) pulse rate and (D–F) call rate were not predicted by individual differences in body condition or two measures of body size (mass and SVL).

Table 1

Output from two GLMM examining the outcome of phonotaxis experiments in which we manipulated within-individual variation in pulse rates (532 observations of 40 individuals). Significant model terms (α = 0.05) are shown in bold.
calling activity may increase the amount of time females spend listening to or localizing their chosen mates.

**Female preferences for call rate variability**

The realized within-individual variation (CV\(_w\)) in the variable call rate of the alternative stimulus was the best predictor of the probability that a subject would choose the alternative over the invariant standard stimulus: females discriminated strongly against alternative stimuli with highly variable call rates (\(\beta = -0.38, t = -4.41, P < 0.001;\) Table 2; Figure 4A). When the realized alternative CV\(_w\) was high, response latencies were significantly longer (\(\beta = 66.00, t = 3.20, P = 0.002\)). The presence of chorus-shaped noise at 70 dB decreased the likelihood that females chose the variable alternative (\(\beta = -0.16, t = -1.99, P = 0.047\)). The realized CV\(_w\) and

---

**Figure 3**

Results from phonotaxis tests in which pulse rate was variable. The top row of plots shows the effect of within-individual variability (CV\(_w\)) on (A) the probability that a female chose the variable alternative stimulus [P(Alternative)] over the invariant standard stimulus and (B) her latency to respond. In plots A and B, the distributions of realized coefficients of variation are plotted within levels of within-individual variability (“low”, “mean,” and “high”) on the second (right-side) y-axis as solid black lines. The bottom row shows the effect of the centered realized mean alternative pulse rate on (C) the probability that a female chose the alternative stimulus [P(Alternative)] and (D) her latency to respond. In all plots, the thicker, blue lines show data from the quiet condition and grayscale lines show data from each level of chorus-shaped background noise.
the highest level of chorus-shaped noise interacted to affect the probability of choosing the variable alternative ($\beta = 0.64, t = 2.32, P = 0.021$) such that the preference against highly variable call rates was effectively abolished in the presence of 80 dB chorus-shaped noise (Figure 4A). Females expressed a directional preference for higher mean alternative call rates ($\beta = 0.16, t = 2.22, P = 0.027$; Figure 4C). There were no significant linear or non-linear effects of the realized mean call rate of the alternative stimulus on the latency to respond (Figure 4D).

**DISCUSSION**

We examined the idea that within-individual variation could be a signal of a male's physical condition or body size. According to the motor performance hypothesis, we should expect condition to influence a male's ability to perform energetically expensive acts repeatedly and consistently (Podos 1997; Podos et al. 2004; Byers et al. 2010). To the extent that males in better condition are more likely to have enough energy to produce consistent displays, body condition might predict within-individual variation in call traits. However, we found no evidence that within-individual variation in either pulse rate or call rate is related to male body condition. Kroodsma (2017) recently reviewed the evidence for the motor performance hypothesis in songbirds and concluded that a missing piece of support is direct evidence that “high-performance” male vocalizations honestly signal male quality. Similarly, we found no evidence that larger males produce less variable advertisement calls. While we cannot conclude from our dataset that within-individual variation is not a signal of some aspect of male quality in treefrogs, it does not appear to signal male body condition, as reflected by measures of residual body mass, or overall size. To the extent that size may reflect age in amphibians with indeterminate growth, we also did not find evidence of a relationship between within-individual variation in signal production and male age.

Nevertheless, within-individual variation in signal production affected receiver behavior. Within-individual variation in male call rate, but not pulse rate, affected female choice between alternative signals, even when we statistically accounted for the mean trait value receivers experienced during a trial. Indeed, females discriminated against highly variable call rates much more strongly than they expressed their frequently documented, relatively strong preferences for faster call rates (Ward et al. 2013b; Tanner et al. 2017). This finding is in contrast with that of Gerhardt and Watson (1995), who manipulated two dynamic properties (pulse number and call rate) simultaneously and found in the eastern gray treefrog, *Hyla versicolor* (the polyplody sister species of *H. chrysoscelis*), that receivers did not discriminate on the basis of within-individual variation in dynamic call properties as long as the two alternatives were of approximately equivalent mean trait values. As noted originally by Gerhardt and Watson (1995), one limitation of their study was that some of the mean temporal properties of signals were confounded with within-individual variation, such that one documented preference ostensibly against within-individual variation may have been driven instead by differences in the number of pulses contained in two alternative stimuli. Because we presented each subject with a set of unique realizations of alternative stimuli and calculated and statistically accounted for differences in realized mean call traits, we have demonstrated unequivocally that receivers attend to within-individual variation in call rate. In gray treefrogs, females can gain genetic benefits for their offspring by mating with males that produce high call efforts (Welch et al. 1998; Welch 2003). Discriminating against males with highly variable call rates could be adaptive because failure to discriminate on the basis of this call trait carries potential fitness costs, and thus, uncertainty around a male's mean trait value injects risk into mating decisions.

That females should discriminate against highly variable call rates but not pulse rates is interesting in light of the species recognition variant of the multiple messages hypothesis for multi-component signaling, which proposes that some call components...
encode species identity while others signal male quality (Møller and Pomiankowski 1993; Candolin 2003; Hebets and Papaj 2005). Pulse rate is a static call property that is considered a good candidate for a signal of species identity because, controlling for temperature, it changes little within bouts of calling, breeding seasons, or an individual’s lifetime (Gerhardt 1991). It is potentially very costly for receivers to make discrimination errors when signals carry information about species identity because hybridization is usually maladaptive (Servedio and Noor 2003). Cope’s gray treefrog breeds syntopically across much of its range with *H. versicolor*; relies primarily on pulse rate to distinguish between conspecific and heterospecific calls (Schul and Bush 2002); and produces largely sterile, triploid offspring when hybridization occurs (Ptacek et al. 1994; Tucker and Gerhardt 2012). The risk of hybridization with *H. versicolor* therefore imposes a high cost for females that make pulse rate discrimination errors.

Figure 4
Results from phonotaxis tests in which call rate was variable. The top row of plots shows the effect of within-individual variability (CV\textsubscript{w}) on (A) the probability that a female chose the variable alternative stimulus [P(Alternative)] over the invariant standard stimulus and (B) her latency to respond. In plots A and B, the distributions of realized coefficients of variation are plotted within levels of within-individual variability (“low,” “mean,” and “high”) on the second (right-side) y-axis as solid black lines. The bottom row shows the effect of the centered realized mean alternative pulse rate on (C) the probability that a female chose the alternative stimulus [P(Alternative)] and (D) her latency to respond. In all plots, the thicker, red lines show data from the quiet condition and grayscale lines show data from each level of chorus-shaped background noise.
By contrast, discrimination errors regarding other aspects of male quality, potentially encoded by dynamic properties, are usually less costly but still carry fitness consequences when high-quality males impart genetic benefits (Norris 1993; Sheldon et al. 1997; Pennig 2000; Head et al. 2003). Thus, we expected that females would be more willing to accept males with highly variable call rates than highly variable pulse rates, based on the inherent risk of misidentifying a male as a conspecific based on his pulse rate. We found the opposite pattern, in which females discriminated against stimuli with highly variable call rates, but not pulse rates. Failure to detect female preferences with regard to within-individual variation in pulse rate may be a consequence of our experimental approach, which prioritized introducing realistic levels of within-individual variation. Calling males exhibit lower within-individual variation in pulse rate than call rate, and so did our experimental stimuli. One interpretation of the data is that we only introduced within-individual variation within the range of female tolerance for variable pulse rates. It is possible that receivers might express preferences against within-individual variation in pulse rate if presented with stimuli that exceeded the natural range of variation. Such a finding might indicate that past selection imposed by receivers had acted to limit the within-individual variation in pulse rates.

In acoustically communicating animals, latency to respond may be used as the sole indication of female preferences in single-stimulus, forced choice experiments, in which faster responses indicate stronger preferences for a signal (Wagner 1998). In our assay, response latencies were affected by the realized mean alternative pulse rate, with females spending longer on average to make mating decisions when the two alternative stimuli had similar pulse rates. There was also a marginally significant tendency for females to spend longer making mating decisions at higher levels of within-individual variation in pulse rate. Additionally, females had substantially longer response latencies when alternative call rates were highly variable. Taken together, these results suggest females may spend more time sampling male signals when they are uncertain about the signalers’ mean trait values. A receiver can be thought to be sampling from a distribution of male trait values by hearing successive calls. She faces a signal detection theory problem in which the wider the distribution is, the more difficult it becomes to accurately identify a mean value with a small number of observations (Wiley 1994, 2006, 2013; Bradbury and Vehrencamp 2011). Listening for a longer period of time, analogous to drawing a larger sample from the distribution, would reduce uncertainty around the male’s mean trait value and potentially mitigate the risk of signal discrimination errors. This observation is consistent with the hypothesis that within-individual variation in call traits acts as noise to obscure trait means in *H. chrysoscelis*.

We presented chorus-shaped noise at three levels based on its measured intensity in natural choruses. While noisy settings can impair intelligibility of speech sounds in humans (the cocktail party problem; Cherry 1953; Bee and Michely 2008) and signal detection and discrimination in non-human animals (Bee and Schwartz 2009; Schmidt and Romer 2011), receivers may improve their performance in noisy settings by gaining spatial separation from a source of noise (Nityananda and Bee 2012; Ward et al. 2013a). Because we presented signals and chorus-shaped noise in a co-located condition, from the same speaker, we likely measured the upper limit of the effect of chorus-shaped noise. Within the range of natural variation we explored, however, chorus-shaped noise had relatively little impact on female choice or response latencies. Few of the estimated effects of chorus-shaped noise attained statistical significance, suggesting that treefrog receivers were not strongly affected by its presence in this assay. However, at its highest level, chorus-shaped noise increased response latencies for females making pulse rate discriminations and also interacted with CV in call rate to abolish the otherwise strong female preference for more stereotyped (less variable) stimuli.

While considerable attention has been paid in recent years to how biotic and abiotic noise affect the learning, production, transmission, and reception of acoustic signals (Brumm 2004; Brumm and Slabbekoorn 2005; Bee and Swanson 2007; Nowack et al. 2007; Love and Bee 2010; Romer 2013; Zuk et al. 2017; Gurule-Small and Tinghitella 2018), comparatively little attention has been paid to the consequences of that noise for the shape of female preference functions (Ward et al. 2013b; Reichert and Ronacher 2015). When noise prohibits the expression of existing female preferences in nature, we may chronically overestimate the strength of sexual selection acting in wild populations because our testing conditions eliminate relevant environmental conditions. Further research is needed to elucidate how sexual selection acts on signals under conditions of realistic complexity, including in the presence of background noise.

Eliminating the within-individual variation present in natural signals when measuring preference functions may cause researchers to overlook biologically relevant signals or to systematically overestimate the strength of sexual selection on signals. Here, we show that within-individual variation does not function as a signal of male body condition or size but that receivers nevertheless discriminate on the basis of within-individual variation in at least some components of multicomponent sexual displays. This behavior may protect receivers from signal discrimination errors and related fitness consequences. Subjects discriminated against highly variable call rates more strongly than they expressed well-documented preferences for differences in alternative call rates, suggesting that within-individual variation in signal production acts as noise to obscure signals of male quality. A valuable avenue for future research would be to examine female preference functions across a range of trait values when alternative stimuli are equally variable. Such data are necessary in order to understand the potential impact of within-individual variation in natural signaling contexts.

**FUNDING**

This work was supported by the National Science Foundation Graduate Research Fellowship under Grant No. 00039202, a Ford Foundation Pre-Doctoral Fellowship, and a Joyce Davenport Fellowship in Natural History through the Bell Museum of Natural History to JCT and by a National Science Foundation Grant (IOS-1452831) to M.A.B.

We are grateful to all members of the Bee lab 2017 for assistance with collecting treefrogs and running trials; H. Brumm, H.C. Gerhardt, G. Klump, and J. Schul for helpful discussions about this work; R.G. Shaw and M. Zuk for feedback on earlier versions of the manuscript; M.S. Caldwell for measuring the amplitude of Minnesota treefrog choruses; and the Three Rivers Park District and Minnesota State Parks for access to field sites.

Data accessibility: Analyses reported in this article can be reproduced using the data provided by Tanner and Bee (2018).

**Handling Editor:** Dr Amanda Ridley
REFERENCES


