

Can a Gray Seal (*Halichoerus grypus*) Generalize Call Classes?

Amanda L. Stansbury, Mafalda de Freitas, Gi-Mick Wu, and Vincent M. Janik
University of St. Andrews

Past researchers have found that gray seals (*Halichoerus grypus*) are capable of classifying vocal signals by call type using a trained set, but were unable to generalize to novel exemplars (Shapiro, Slater, & Janik, 2004). Given the importance of auditory categorization in communication, it would be surprising if the animals were unable to generalize acoustically similar calls into classes. Here, we trained a juvenile gray seal to discriminate novel calls into 2 classes, “growls” and “moans,” by vocally matching call types (i.e., the seal moaned when played a moan and growled when played a growl). Our method differed from the previous study as we trained the animal using a comparatively large set of exemplars with standardized durations, consisting of both the seal’s own calls and those of 2 other seals. The seal successfully discriminated growls and moans for both her own (94% correct choices) and the other seals’ (87% correct choices) calls. We used a generalized linear model (GLM) and found that the seal’s performance significantly improved across test sessions, and that accuracy was higher during the first presentation of a sound from her own repertoire but decreased after multiple exposures. This pattern was not found for calls from unknown seals. Factor analysis for mixed data (FAMD) identified acoustic parameters that could be used to discriminate between call types and individuals. Growls and moans differed in noise, duration and frequency parameters, whereas individuals differed only in frequency. These data suggest that the seal could have gained information about both call type and caller identity using frequency cues.

Keywords: classification, generalization, discrimination, gray seal, call types

Supplemental materials: <http://dx.doi.org/10.1037/a0039756.supp>

A fundamental ability in animal cognition is the capacity for classification and generalization of stimuli. Classifying objects reduces large numbers of stimuli into small categorical groups, allowing animals to cope with stimulus variability. Generalization places new stimuli into pre-existing categories, speeding up signal processing and response time.

Classification and generalization of stimuli have been well investigated in the visual domain for several species, such as bottlenose dolphins (e.g., Mercado, Killebrew, Pack, Mácha, & Herman, 2000), domestic dogs (Range, Aust, Steurer, & Huber, 2008), pigeons (e.g., Cook & Wasserman, 2006), parrots (e.g., Pepperberg, 1996, 1999) and nonhuman primates (e.g., Neiworth & Wright, 1994; Schrier & Brady, 1987). Many species are similarly capable of auditory categorization. For example, animals can

use sound categorization to classify objects, particularly echolocating species (e.g., von Helversen, 2004; Roitblat, Penner, & Nachtigall, 1990). Past research has mainly concentrated on the classification of biological sounds, such as bird song (e.g., Beecher, Campbell, & Burt, 1994) and predator-specific alarm calls (e.g., Cäsar, Zuberbühler, Young, & Byrne, 2013; Greene & Meagher, 1998; Seyfarth, Cheney, & Marler, 1980). Some species categorize conspecific call types by responding to another individual’s calls with a matching call type, referred to as antiphonal calling (e.g., Ghazanfar, Flombaum, Miller, & Hauser, 2001; Solitis, Leong, & Savage, 2005). Perhaps most importantly, auditory categorization helps animals to deal with individual variation in communication calls when different classes of signals have different meanings. Previous studies have used playback tests and discrimination training to find out how individuals categorize sounds, providing insight into how animals judge conspecific (e.g., Candiotti, Zuberbühler, & Lemasson, 2013; Lind, Dabelsteen, & McGregor, 1996; Maros et al., 2008; Weary & Krebs, 1992) and/or heterospecific sounds, including human speech and music (e.g., Bloomfield, Sturdy, Phillmore, & Weisman, 2003; Brooks & Cook, 2010; Hoeschele, Cook, Guillelte, Brooks, & Sturdy, 2012; Kuhl, 1981; Pepperberg, 1999; Porter & Neuringer, 1984). Several species are able to discriminate between hundreds of human speech sounds and generalize words produced by different speakers (e.g., dogs; Kaminski, Call, & Fischer, 2004; parrots; Pepperberg, 1999; and primates; Savage-Rumbaugh et al., 1993).

Given the importance of auditory categorization in establishing a common communication code, it would be surprising if animals were unable to generalize calls that are acoustically similar. Nev-

This article was published Online First October 12, 2015.

Amanda L. Stansbury, Mafalda de Freitas, Gi-Mick Wu, and Vincent M. Janik, Sea Mammal Research Unit, School of Biology, University of St. Andrews.

We thank Simon Moss, Matt Bivins, Ryan Milne, Alicia Widmer, and all of the student volunteers (particularly Sian Tarrant and Helen Hiley) at our facility for their help during this study. This study was funded by the Natural Environment Research Council of the United Kingdom, Grant NE/I024682/1. It was conducted under Home Office license number 60/3303.

Correspondence concerning this article should be addressed to Amanda L. Stansbury, Sea Mammal Research Unit, School of Biology, University of St. Andrews, Fife, KY16 8LB, United Kingdom. E-mail: as252@st-andrews.ac.uk

ertheless, such a result has been reported for gray seals. Shapiro, Slater, and Janik (2004) found that two juvenile gray seals named Kylie and Oscar were capable of discriminating between two call types, “growls” and “moans.” These call types are easily distinguishable by human observers. Growls are broadband, noisy calls with a bandwidth upper limit extending to frequencies up to 20 kHz, whereas moans are tonal, periodic calls with a harmonic structure and a bandwidth upper limit rarely exceeding 5 kHz (Shapiro et al., 2004; Figure 1). Kylie and Oscar were trained to vocalize upon the presentation of a sound: recordings of the seals’ own growls and moans. The seals were trained with a set of nine exemplars (five growls and four moans for Kylie, vice versa for Oscar), and were reinforced for responding with a matching call type (i.e., responded with a moan when played a moan and a growl when played a growl). Both seals successfully learned the task after several hundred trials (~800 trials for Oscar and ~1,650 for Kylie). However, neither seal accurately responded when presented with novel exemplars (new sets of their own growls and moans). One seal responded variably to the new calls, whereas the other consistently responded with a growl. Shapiro et al. (2004) concluded that although their seals were able to discriminate between call types, they were unable to generalize classification to novel call exemplars.

There are several possible explanations for why the seals were unable to generalize novel exemplars, such as the training set size and parameter differences between training and testing stimuli. Training set size has been previously shown to impact generalization; for example, larger training sets allow for faster category formation and improved accuracy (e.g., Katz & Wright, 2006; Schrier & Brady, 1987; Wasserman & Bhatt, 1992; Wright, Cook, Rivera, Sands & Delius, 1988; Wright, Rivera, Katz, & Bachevalier, 2003). Shapiro et al. (2004) used a small, fixed set of stimuli (nine exemplars) to train classification before testing generalization with new stimuli. Although such small training sets have been successfully used for discrimination tasks in other species (e.g., Brooks & Cook, 2010; Watanabe, 2001), it is possible that the seals simply learned through trial and error to respond with the appropriate call type for each separate stimulus, as opposed to generalizing the calls into categories. This is consistent with results from other auditory classification studies which have been similarly unsuccessful at training call type categorization of novel stimuli in rhesus monkeys (*Macaca mulatta*: Le Prell, Hauser, & Moody, 2002) and a beluga whale (*Delphinapterus leucas*: Vergara, 2011). Both studies used small sets of training stimuli (eight exemplars of each type for the rhesus monkeys, three of each for the beluga) before testing generalization to novel sounds. Large, variable sets of stimuli have been used for other classification tasks. In a notable example, an African gray parrot (*Psittacus erithacus*), Alex, was successfully trained to classify objects in several ways, including color, size, and shape using dozens of training exemplars (Pepperberg, 1999). For example, during a size-discrimination task, Alex was required to identify which object in a pair was either larger or smaller (Pepperberg & Brezinsky, 1991). Initially, the training set consisted of 18 exemplars but was then expanded to include 32 different types of objects, with exemplars of each type varying in size and color. By using a larger exemplar set, it is unlikely that Alex learned specific responses to specific stimuli. Instead, it is more likely he generalized across exemplars to accomplish the task (Pepperberg & Brezinsky, 1991).

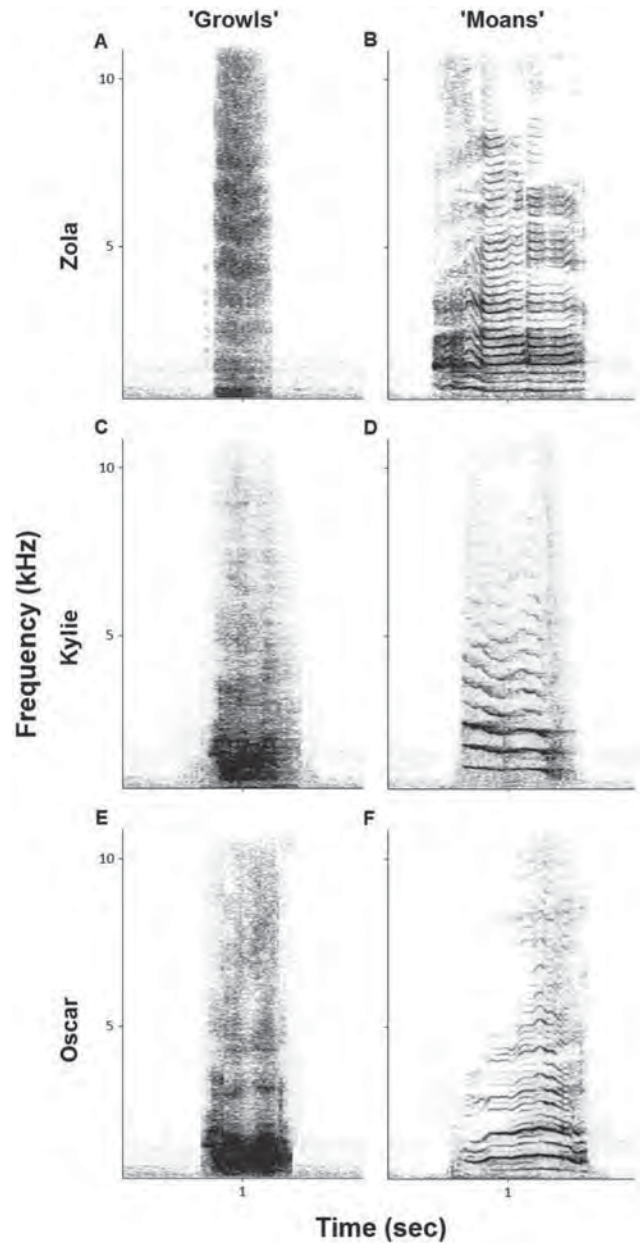


Figure 1. Spectrographic examples of call types from the three seals. Growls are broadband, noisy calls, and moans are tonal, periodic calls with a harmonic structure. Example growls are shown from (A) Zola, (C) Kylie, and (E) Oscar, and moans from (B) Zola, (D) Kylie, and (F) Oscar. Spectrograms created in Avisoft-SASlab Pro (FFT size 2048; frequency resolution, 46.87 Hz; time resolution, 10.66 ms; weighting function, hamming; window width, 100%).

Another possible explanation for Shapiro et al.’s (2004) result is that the duration of signals changed between the training set and novel stimuli. For example, during training, moans were played for Oscar with an average duration of 1.5 s (± 0.9), whereas growls had an average duration of 2.6 s (± 1.2). For the novel exemplars, the absolute duration of both call types changed (novel moans were played for an average of 0.4 s ± 0.1 , and growls for 0.7 s $\pm$

0.1). Although the relative duration remained consistent (moans were roughly half the duration of growls for both training and novel stimuli), if absolute duration was used as a cue to discriminate between call types, this would prevent the seals from successfully performing the task.

In addition, Shapiro et al. (2004) only presented the seals with their own sounds. Although this method does test the animal's ability to classify its own calls, the seal should also be capable of generalizing novel exemplars from conspecifics. For auditory categorization to be used for communication, animals must be able to classify both the calls they produce and the calls they perceive from others. Thus, generalization should be tested for both the seal's own calls and for those of other individuals.

For the present study, we tested a juvenile gray seal's ability to discriminate her own sounds in addition to those produced by other, unknown juvenile gray seals. We report that, when she was trained, the seal was capable of generalizing two vocalization classes, growls and moans, using large, continually changing sets of stimuli with standardized durations.

Method

Subject

A juvenile female gray seal, Zola, born November 7, 2011 on the Isle of May (Firth of Forth, Scotland), was the subject of this study. After weaning, Zola was transported to our laboratory (St. Andrews, Scotland) on December 6th, 2011. She was housed with four other juveniles in three enclosures that included a large rectangular pool (42 m × 6 m × 2.5 m) and two circular pools (3 m × 5 m × 2 m). The seals were fed a varied diet of several fish species (mostly consisting of herring, *Clupea harengus*, and sprat, *Clupea sprattus*). Training and testing occurred at the facility for 12 months, after which Zola was released into the wild.

Acoustic Recordings

Acoustic recordings were obtained using a Sennheiser (Wedemark, Germany) MKH 416 P48 directional microphone (frequency response = 40–20,000 Hz ± 1 dB) and an Edirol FA-66 (Roland (UK) Ltd., Swansea, UK) external sound card (sampling rate = 96 kHz, 24-bit) with a laptop computer. Weather permitting, sessions were concurrently video recorded using a Sony HDR CX250E video camera. Stimuli were played from the laptop using the FA-66 through an external Skytec 170.170 active speaker (frequency response 32–22,000 Hz). Sounds were simultaneously played, recorded and spectrographically monitored in real time using the program Audacity 1.3 (sampling rate = 96 kHz, 24-bit, Audacity Team, 2012).

Training and Testing Procedure

Positive reinforcement behavioral training was conducted using the seal's normal daily diet as a reward. Husbandry training began in January 2012 and focused on general behaviors including exiting the water and stationing at the poolside for testing. Vocalization training began in March, 2012. Initially Zola was reinforced for any sound she produced while in the same enclosure as the other seals. This was done because the seals vocalized more

frequently while in a group, and were relatively silent when isolated. Once Zola regularly vocalized, she was separated from the others for individual sessions. At this time Zola produced the two distinct call types, growls and moans. Each call type was paired with a unique hand cue. The hand cues were then faded out to be replaced with the presentation of a sound (either a growl or moan) and a light. The LED light was used to indicate when the sound had finished and cue the seal's response. After the hand cues were replaced by sound, all training and testing sessions occurred with the experimenter waiting quietly out of sight.

Simultaneous to training the vocal responses, the trainer reinforced Zola for staying quiet when the vocal cues were not presented. This behavior was shaped until she remained quietly at station between trials for up to 15 min at a time. Although between most trials this stationing period was less than a min, in some cases the intertrial interval was extended to avoid an overlap with environmental noise. Training sessions were variable in length depending on the seal's performance; a set amount of food was used to reinforce the seal, and sessions continued until all the fish was used (mean number of trials per session = 61, $SD \pm 26$, mean session duration = 37 min, $SD \pm 9$).

Throughout training and testing, each trial was initiated by the seal leaving the water and stationing in front of the speaker. Once positioned, a sound (either a growl or a moan) was played. After the sound ended, the LED light illuminated and cued the seal's response. The animal was required to produce a vocalization of the same type. This response was then evaluated using a visual judgment of a real-time spectrogram display comparing the played stimuli to the seal's response in Audacity. If correct (produced same call type as the one played), the seal was reinforced with fish after it stopped vocalizing. If incorrect (produced different call type than the one played), a time out of 5–10 s was used before the seal could initiate the next trial. Training continued until the seal had reached the criterion of seven consecutive sessions with overall accuracy above 80%, at which point testing began. During testing, stimuli were presented in blocks of 50 trials. Depending on how quickly trials were completed (due to breaks for environmental noise), each test session would consist of either 50 or 100 trials (mean session duration = 47 min, $SD \pm 7$). If the seal's response overlapped part of the played stimulus, the seal was not reinforced, regardless of whether she responded correctly or not (this occurred in 136 of the 1,182 training and 72 of the 700 test trials). Because call type could still be determined, the seal's calls were still used when we measured its accuracy in the response.

Sound Stimuli and Acoustic Analysis

Training stimuli were composed of Zola's own previously recorded vocalizations in sets of 10 to 20 novel exemplars per session. The sounds were played randomly with one exception: If Zola responded incorrectly, the sound was replayed up to three times until she produced the correct response. If after three replays Zola had not responded correctly, we moved on to the next call. This self-correction procedure was only used during training trials. In total, 212 different calls (106 growls and 106 moans) were used throughout training.

Test stimuli consisted of 140 novel calls, 70 produced by Zola and 70 from the two juvenile seals studied by Shapiro et al. (2004), Kylie and Oscar, who were unknown to Zola. All calls were

between 0.7 and 1.2 s in duration. During each trial, one of the 140 novel-call stimuli was played. Stimuli were presented in blocks of 50 trials (five growls and five moans), presented five times per stimulus in random order assigned by a Gellerman series (Gellerman, 1933). Thus, Zola was tested with 14 test blocks (700 trials). Frequency and time parameters were measured using Avisoft-Saslab Pro 5.02.04 sonogram software (Glienicke, Germany). Parameters were chosen based on previous studies examining gray seal vocal repertoires (Asselin, Hammill, & Barrette, 1993; McCulloch, Pomeroy, & Slater, 1999). See Table 1 for a list of measured parameters and their definitions, including spectral richness (Miller & Murray, 1995). All statistical analyses were conducted using R 3.0.1.

Results

Call-Type Classification

Two independent human observers classified Zola's responses as growls or moans post hoc. At this time, the observers classified only the seal's response; they were unaware of which stimulus was played during each trial. To confirm that these calls were classified appropriately, Cohen's κ was used and demonstrated that interobserver reliability was high (99.7% agreement, Cohen's $\kappa = 0.9547$, $SE \pm 0.0109$). *K*-means cluster analysis ($k = 2$) was conducted using the frequency and time parameters measured to evaluate human classification. All calls grouped by k means were in agreement (100%) with one of the two human categorizations, which was also in complete agreement (100%) with the experimenter who reinforced the seal during testing. Thus, this grouping was deemed objective and used to score the seal's accuracy for the remaining statistical analysis.

Training

The seal's accuracy during training sessions, that is, all sessions until a criterion of seven consecutive sessions ($>80\%$ correct) was reached, was plotted as percent correct per session by call type (see Figure 2). Zola reached criterion after 1,182 trials across 18

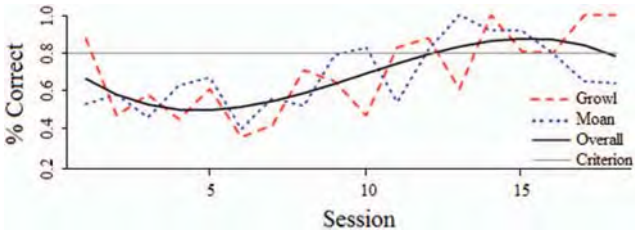


Figure 2. Performance accuracy by session for Zola's training period. Accuracy (percent correct) is shown by call type (i.e., growls and moans). A nonlinear least squares model was used to determine the best fit polynomial trend line (Order 3) for overall performance, weighted by the number of trials per session. Training continued until a criterion of seven consecutive sessions with $>80\%$ accuracy for growls and moans combined was met. See the online article for the color version of this figure.

sessions. This was comparable to Kylie's ($\sim 1,650$ trials, 47 sessions) and Oscar's (~ 800 trials, 23 sessions) performances (Shapiro et al., 2004). A nonlinear least squares (NLS) model was used to examine how the subject's overall accuracy changed over the learning period by date and call type. As session length throughout training was dependent upon the seal's performance, the number of trials per session varied and the model was weighted by number of trials per day. The seal's overall accuracy was plotted with a polynomial trendline fitted for the best fit model, determined using the lowest corrected Akaike information criterion (AICc) score tested for all orders up to seven (see Figure 2).

Testing

For test sessions, the seal's accuracy was plotted as percent correct by blocks of 50 trials and compared between her own and the calls of the two unknown seals (see Figure 3). Overall Zola responded with 93.7% accuracy to her own calls and with 87.2% accuracy to Kylie's and Oscar's calls. This performance was significantly above chance, $\chi^2(1, 700) = 426.05$, $p < .001$.

Table 1
Definitions for Measured Parameters

Parameter	Definition
Duration	Time from the beginning to end of the call, within 35 dB of the spectrum's maximum peak.
Peak frequency	The frequency with the highest amplitude measured at the beginning, middle, end, and maximum (i.e., the frequency with the highest amplitude) of the whole call.
Fundamental frequency	Measured only for tonal, harmonic calls. Measured at the centre of the call as the lowest integer multiple of corresponding amplitude peaks.
Maximum frequency	The highest frequency reached at any time point within 20 dB of the spectrum's maximum peak.
Wiener entropy	The ratio of the geometric mean to the arithmetic mean of the spectrum. This parameter is closer to zero for pure-tone signals and closer to one for random noise. The signal was measured at the centre of the call, and compared to background noise around the call.
Harmonic to noise ratio (HNR)	The dB ratio between the harmonic and non-harmonic energy. The ratio is averaged for each frequency bin (46.87 Hz), for a 10.66 ms section measured at the centre of the call.
Spectral richness (SR)	A measure of broadband noise and harmonic spacing as defined by Miller and Murray (1995). Calls are rated on a scale of 0 to 3 based on separation between harmonic bands (H) and the width of the fundamental frequency band (F). F is measured by subtracting the lowest from the highest frequency of the fundamental at the band's widest point. H is measured by subtracting the highest frequency of the fundamental from the lowest frequency of the second harmonic at the centre of the call. A score of 0 is applied to broadband signals, 1 if $H > F$, 2 if $H = F$, and 3 if $H < F$.

Note. All parameters were measured from spectrograms (FFT size = 2048; frequency resolution = 46.87 Hz; time resolution = 10.66 ms).

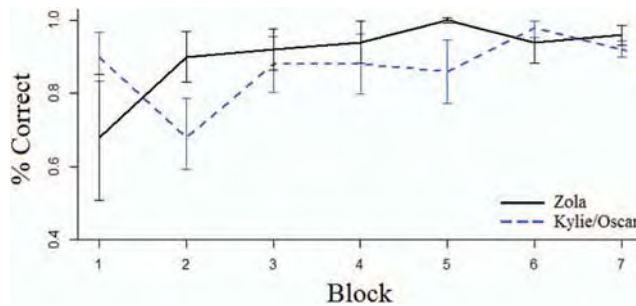


Figure 3. Test-session accuracy by blocks of 50 trials. Accuracy (percent correct) and standard deviation are shown, separated by test–stimuli source (Solid line: Zola; dashed line: Kylie and Oscar, whose calls were taken from Shapiro et al., 2004). See the online article for the color version of this figure.

To determine if alternative parameters might have affected the seal's test performance, we used a generalized linear model (GLM). Accuracy was predicted using a binomial distribution and logit link function with the following covariates: date, trial number, signal exposure (the number of times the seal had been exposed to stimuli; each test signal was played five times), call source (Zola's, Kylie's and Oscar's calls), and call type (growl vs. moan). All possible combinations were tested including interactions between date, call source, and call type. The best models were selected using the AICc score; all models within 10 AICc points of the best model were used for model averaging (the models selected for averaging are shown in Supplemental Material 1). The best models were averaged using modelavg (in R package MuMIn, Version 1.9.13; Barton, 2013) to determine the final model. The model indicated that Zola's accuracy throughout testing significantly improved across sessions (see Table 2). However, her accuracy decreased with signal exposure across sessions. Each test signal was played five times, and Zola had significantly higher accuracy during the first exposures, and lower accuracy for the last exposure. This significant decrease in accuracy across exposures occurred when hearing her own calls, but not for Kylie and Oscar's calls (Table 2 and supplemental material 2). However, the degree to which these variables affect the model is very small. The covariates only explain a small proportion (7.5%) of the variance (Nagelkerke's pseudo $r^2 = 0.075$).

Table 2
Final Generalized Linear Model Selected Using Model Averaging

Variable	Model estimate e^a	95% CI	p
Intercept	.421	[-.028, .950]	.848
Call source	.736	[-.220, .965]	.381
Date	.643	[-.517, .751]	.026
Signal exposure	.500	[-.299, .701]	.999
Date $\times$ Signal Exposure	.470	[-.447, .493]	.009
Call Source $\times$ Signal Exposure	.630	[-.502, .741]	.047

Note. The models used for averaging are shown in Supplementary Material 1. Log-adjusted model weights, 95% confidence intervals (CI), and p values are shown.

Factor Analysis for Mixed Data

Descriptive statistics (mean and standard deviation) for measured acoustic parameters by call type are shown in Table 3. Factor analysis for mixed data (FAMD) was used to further examine the variability between calls and identify possible cues the seal could have used to distinguish between call types (package FactoMineR, Version 1.26; Husson, Josse, Le, & Mazet, 2014). FAMD was chosen because it allowed both continuous and categorical parameters to be used to group calls similarly to k -means cluster analysis (Ding & He, 2004). Calls were analyzed both by call type (growl and moan) and by individual (Zola, Kylie, and Oscar). These methods allowed us to determine what parameters explained the most variation between call types across individuals, and those that explained differences between individuals.

The Kaiser–Meyer–Olkin (KMO) measure verified that sample size was adequate with a KMO = 0.78. Bartlett's test of sphericity, $\chi^2(903) = 89,533.15$, $p < .001$ indicated that correlations between variables were sufficiently large to conduct FAMD. Dimensions were plotted against each other to examine clustering by call type and individual. For call type, growls and moans were clearly clustered for the first three dimensions, which explained 64.16% of the data's variance (Figure 4a). Calls were clustered by individual for the first two dimensions, which explained 53.29% (growls) and 40.9% (moans) of the data's variance (Figure 4b and 4c).

To identify what parameters explained the variance between call types, factor loadings for the first three dimensions were examined (Supplemental Material 3). Based on the loading weights these related to noise content, frequency and duration parameters. Call clustering suggested that growls scored high on noise content and frequency but low on duration dimensions. Moans were calls that scored low on noise content and frequency but high on duration dimensions. The first dimension was also heavily weighted by categorical parameters; growls had no fundamental frequency, and scored low on signal to noise categories, and moans had a fundamental frequency and scored high on signal to noise categories. When examining parameters that explained variance between individuals, factor loadings for the first 2 dimensions were examined (Supplemental Material 3). Based on the loading weights, these components related to frequency parameters. Call clustering by individual suggested that for growls, individuals differed in peak frequency parameters. For moans, both fundamental and peak frequency differed between individuals, however fundamental frequency explained more of the individual variation (Supplemental Material 3).

Table 3
Means and Standard Deviations for Measured Acoustic Parameters by Call Type

Parameter	Growls	Moans
Duration (sec)	.85 $\pm$.12	.92 $\pm$.12
Fundamental frequency (Hz)	N/A	357 $\pm$ 141
Peak frequency (Hz)	522 $\pm$ 358	874 $\pm$ 300
Maximum frequency (Hz)	16512 $\pm$ 3221	12989 $\pm$ 2431
Weiner entropy	.45 $\pm$.18	.28 $\pm$.09
Harmonic to noise ratio	16 $\pm$ 5.7	23.5 $\pm$ 5.5

Note. Values for growls and moans averaged across all playback stimuli.

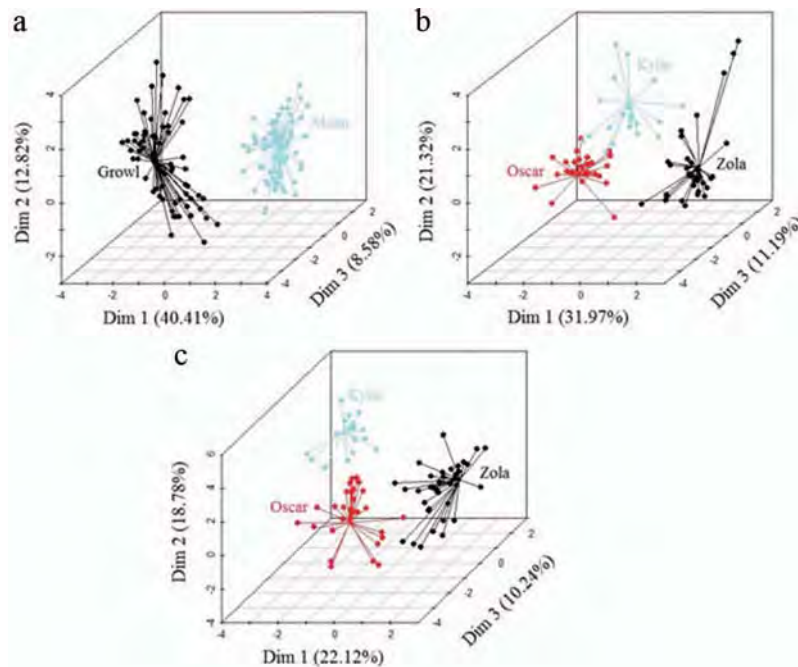


Figure 4. Three-dimensional scatterplots of Components 1 through 3 from factor analysis for mixed data (FAMD). The proportions of the data's variance explained by each component are shown in parentheses. (4a) Calls from all three seals are clustered by call type, (4b) growls are clustered by individual, and (4c) moans are clustered by individual. See Supplemental Material 3 for a table of factor loadings showing how variables loaded onto the dimensions. See the online article for the color version of this figure.

Discussion

This study demonstrated that a gray seal was clearly capable of categorizing and generalizing between call classes, contrasting with Shapiro et al.'s (2004) findings that gray seals were unable to generalize novel stimuli. This disparity may have been due to methodological differences. Shapiro et al.'s (2004) seals were exposed to a small set of training stimuli and then tested with novel stimuli. As the seals were trained with such a small set they may have learned to respond to each sound as an individual cue, rather than to generalize to new stimuli. Other studies using small training sets have been similarly unsuccessful at training call-type categorization of novel stimuli in other species (Le Prell et al., 2002; Vergara, 2011). In generalization tasks using other modalities, training-set size has been found to impact learning speed and accuracy (e.g., Katz & Wright, 2006; Schrier & Brady, 1987; Wasserman & Bhatt, 1992; Wright et al., 1988; Wright et al., 2003). Thus, Kylie and Oscar's inability to generalize may be attributed to the training with a small number of exemplars; through our use of a variety of stimuli Zola was required to generalize from the start of training, and was thereby successful when tested with novel stimuli.

The duration of stimuli used by Shapiro et al. (2004) also varied between training and testing sets, which may have prevented the seal's successful classification of test stimuli if absolute duration was used as a cue. Zola was tested with calls of standardized duration, which may have facilitated her success. Interestingly, although the duration of all calls fell within a small range (0.7 and 1.2 s), call types could still be separated by duration. In addition,

Zola was 2 months older than the animals studied by Shapiro et al. (2004), but it seems unlikely that this would have made a difference, as she appeared to learn the task at comparable rates to Shapiro et al.'s (2004) animals.

Zola performed significantly better during the first exposures to individual signals in the tests than later on, supporting further that the seal generalized signal type rather than learned items individually. The decrease in accuracy for later signal exposures may be attributed to fatigue; the later presentations of each signal would occur near the end of each block and long sessions may have impacted accuracy. Interestingly, the average duration of sessions consisting of both 50 and 100 trials were comparable (overall mean duration = 47 min, ± 7). The similar overall session length for sessions consisting of either one or two blocks may be attributed to variable intertrial intervals; during some sessions, background noise was low and several trials were quickly completed, allowing for two blocks to be finished within one session. However, on days with varying background noise the intertrial interval was sometimes longer as trials did not resume until background noise was reduced, only allowing for one block to be completed. Future studies could address effects of fatigue and trial variability by using shorter test blocks with fixed intertrial intervals.

The factor analysis indicated that call types could be classified using noise content, frequency, and duration parameters whereas individuals could be separated using frequency parameters. Growls had a higher maximum call frequency than moans, but lower peak frequency (see Table 3). Within growls, individuals were separated by peak frequency, whereas for moans although peak frequency

contributed, fundamental frequency explained more of the individual variation (Figure 5 and Supplemental Material 3). Thus, the seal could have used different frequency parameters to provide information about both call type and caller identity. These results are also comparable to McCulloch et al.'s (1999) findings that wild gray seal pup calls could be individually identified using duration and frequency parameters. Interestingly, although the calls were found to be individually distinctive, mothers did not differentially respond to playbacks of their own versus unknown pups in Scotland (McCulloch et al., 1999). This was not the case at a different seal colony, Sable Island, where gray seal mothers were found to respond to playback of their own pups (McCulloch & Boness, 2000). Although gray seal vocalizations appear to provide caller information, the biological significance is unclear.

We found it interesting that one of Shapiro et al.'s (2004) seals tended to reply to any novel stimulus with a growl. It is possible that the seal responded to the novel stimuli as if presented with an intruding conspecific. Zola similarly responded to growls by growling even in the first session of training, suggesting that initially her response may also be a natural reply to the growl of a conspecific rather than a learned response. Later in the sessions, her growling response deteriorated, most likely due to a habituation effect in the playback context, before she successfully replied to growls by growling again after a learning period (see Figure 2). The initial growling in our and Shapiro et al.'s (2004) study suggests that growls may be aggressive or defensive signals when challenged by a conspecific. The change in behavior throughout the experiments further confirms Shapiro et al.'s (2004) finding of contextual learning in the vocal domain (Janik & Slater, 2000) in gray seals. Although our study only examined calls produced in an artificial training context and can thus only hypothesize as to the function of these call types, future studies could examine the biological significance of these calls in wild animals using playback studies.

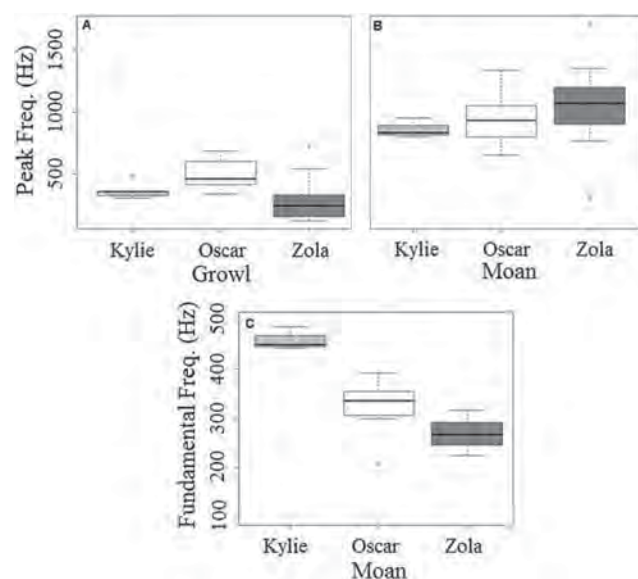


Figure 5. Tukey's box plots for test stimuli by caller showing the peak frequency for (A) growls and (B) moans, and fundamental frequency for (C) moans. Fundamental frequency for growls is not shown, as growls are broadband calls without a periodic structure.

It is unknown what information Zola used for discriminating between call types. She may have used a single parameter to classify the calls, or a combination of multiple parameters. Frequency parameters between call types and individuals may be more reliable in combination with other call features. Although parameters salient to human researchers were measured, it is also possible other unconsidered cues were used for discrimination by the seal. Features relevant to humans may or may not also be relevant to nonhuman species, which is a commonly acknowledged problem when examining animal communication (e.g., Harley, 2008). However, the fact that Zola performed successfully in these trials demonstrated that generalization across different versions of the same call type is possible in this species. At this time, it is unclear whether Zola used an exemplar based strategy for this, where features that best predicted group membership across multiple learned examples were used to classify novel calls, or prototype based discrimination, which is based on similarity to a 'best' example of each call type. Future research could investigate these issues by altering training stimuli systematically.

Further studies on learning, production, and perception capabilities and limitations of vocal behavior would be valuable for pinnipeds, given their interesting vocal learning skills (Janik, 2006) and their dynamic social structures (Ruddell, Twiss, & Pomeroy, 2007).

References

- Asselin, S., Hammill, M. O., & Barrette, C. (1993). Underwater vocalizations of ice breeding grey seals. *Canadian Journal of Zoology*, 71, 2211–2219. <http://dx.doi.org/10.1139/z93-310>
- Audacity Team. (2012). Audacity (Version 1.3) [Computer Software]. Retrieved from <http://audacity.sourceforge.net/>
- Barton, K. (2013). MuMIn: Multi-model inference [Computer software R]. Retrieved from <http://CRAN.R-project.org/package=MuMIn>
- Beecher, M. D., Campbell, S. E., & Burt, J. M. (1994). Song perception in the song sparrow: Birds classify by song type but not by singer. *Animal Behaviour*, 47, 1343–1351. <http://dx.doi.org/10.1006/anbe.1994.1182>
- Bloomfield, L. L., Sturdy, C. B., Phillmore, L. S., & Weisman, R. G. (2003). Open-ended categorization of chick-a-dee calls by black-capped chickadees (*Parus atricapilla*). *Journal of Comparative Psychology*, 117, 290–301. <http://dx.doi.org/10.1037/0735-7036.117.3.290>
- Brooks, D. I., & Cook, R. G. (2010). Chord discrimination by pigeons. *Music Perception: An Interdisciplinary Journal*, 27, 183–196. <http://dx.doi.org/10.1525/mp.2010.27.3.183>
- Candiotti, A., Zuberbühler, K., & Lemasson, A. (2013). Voice discrimination in four primates. *Behavioural Processes*, 99, 67–72. <http://dx.doi.org/10.1016/j.beproc.2013.06.010>
- Cäsar, C., Zuberbühler, K., Young, R. J., & Byrne, R. W. (2013). Titi monkey call sequences vary with predator location and type. *Biology Letters*, 9, 20130535. <http://dx.doi.org/10.1098/rsbl.2013.0535>
- Ding, C., & He, X. (2004). *K-means clustering via principle component analysis*. Proceedings of the Twenty-First International Conference on Machine Learning (p. 29). New York, NY: ACM.
- Gellerman, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *The Pedagogical Seminary and Journal of Genetic Psychology*, 42, 206–208. <http://dx.doi.org/10.1080/08856559.1933.10534237>
- Ghazanfar, A. A., Flombaum, J. I., Miller, C. T., & Hauser, M. D. (2001). The units of perception in the antiphonal calling behavior of cotton-top tamarins (*Saguinus oedipus*): Playback experiments with long calls. *Journal of Comparative Physiology: A. Neuroethology, Sensory, Neural*

- and Behavioral Physiology, 187, 27–35. <http://dx.doi.org/10.1007/s003590000173>
- Greene, E., & Meagher, T. (1998). Red squirrels, *Tamiasciurus hudsonicus*, produce predator-class specific alarm calls. *Animal Behaviour*, 55, 511–518. <http://dx.doi.org/10.1006/anbe.1997.0620>
- Harley, H. E. (2008). Whistle discrimination and categorization by the Atlantic bottlenose dolphin (*Tursiops truncatus*): A review of the signature whistle framework and a perceptual test. *Behavioural Processes*, 77, 243–268. <http://dx.doi.org/10.1016/j.beproc.2007.11.002>
- Hoeschele, M., Cook, R. G., Guille, L. M., Brooks, D. I., & Sturdy, C. B. (2012). Black-capped chickadee (*Parus atricapillus*) and human (*Homo sapiens*) chord discrimination. *Journal of Comparative Psychology*, 126, 57–67. <http://dx.doi.org/10.1037/a0024627>
- Husson, F., Josse, J., Le, S., & Mazet, J. (2014). FactoMineR: Multivariate exploratory data analysis and data mining with R [Computer software]. Retrieved from <http://CRAN.R-project.org/package=FactoMineR>
- Janik, V. M. (2006). Communication in marine mammals. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (2nd ed., Vol. 2, pp. 646–654). Oxford, UK: Elsevier. <http://dx.doi.org/10.1016/B0-08-044854-2/00834-8>
- Janik, V. M., & Slater, P. J. B. (2000). The different roles of social learning in vocal communication. *Animal Behaviour*, 60, 1–11. <http://dx.doi.org/10.1006/anbe.2000.1410>
- Kaminski, J., Call, J., & Fischer, J. (2004). Word learning in a domestic dog: Evidence for “fast mapping.” *Science*, 304, 1682–1683. <http://dx.doi.org/10.1126/science.1097859>
- Katz, J. S., & Wright, A. A. (2006). Same/different abstract-concept learning by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 32, 80–86. <http://dx.doi.org/10.1037/0097-7403.32.1.80>
- Kuhl, P. K. (1981). Discrimination of speech by nonhuman animals: Basic auditory sensitivities conducive to the perception of speech-sound categories. *The Journal of the Acoustical Society of America*, 70, 340–349. <http://dx.doi.org/10.1121/1.386782>
- Le Prell, C. G., Hauser, M. D., & Moody, D. B. (2002). Discrete or graded variation within rhesus monkey screams? Psychophysical experiments on classification. *Animal Behaviour*, 63, 47–62. <http://dx.doi.org/10.1006/anbe.2001.1888>
- Lind, H., Dabelsteen, T., & McGregor, P. K. (1996). Female great tits can identify mates by song. *Animal Behaviour*, 52, 667–671. <http://dx.doi.org/10.1006/anbe.1996.0211>
- Maros, K., Pongrácz, P., Bárdos, G., Molnár, C., Faragó, T., & Miklósi, A. (2008). Dogs can discriminate barks from different situations. *Applied Animal Behaviour Science*, 114, 159–167. <http://dx.doi.org/10.1016/j.applanim.2008.01.022>
- McCulloch, S., & Boness, D. J. (2000). Mother–pup vocal recognition in grey seals (*Halichoerus grypus*) of Sable Island, Nova Scotia, Canada. *Journal of Zoology*, 251, 449–455.
- McCulloch, S., Pomeroy, P. P., & Slater, P. J. B. (1999). Individually distinctive pup vocalizations fail to prevent allo-suckling in grey seals. *Canadian Journal of Zoology*, 77, 716–723. <http://dx.doi.org/10.1139/z99-023>
- Mercado, E., III, Killebrew, D. A., Pack, A. A., Mácha, I. V. B., IV, & Herman, L. M. (2000). Generalization of ‘same-different’ classification abilities in bottlenosed dolphins. *Behavioural Processes*, 50, 79–94. [http://dx.doi.org/10.1016/S0376-6357\(00\)00100-5](http://dx.doi.org/10.1016/S0376-6357(00)00100-5)
- Miller, E. H., & Murray, A. V. (1995). Structure, complexity, and organization of vocalizations in harp seal (*Phoca groenlandica*) pups. In P. E. Nachtigall, J. A. Thomas, & R. A. Kastelein (Eds.), *Sensory systems of aquatic mammals* (pp. 237–264). Woerden, the Netherlands: De Spil.
- Neiwirth, J. J., & Wright, A. A. (1994). Monkeys (*Macaca mulatta*) learn category matching in a nonidentical same-different task. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 429–435. <http://dx.doi.org/10.1037/0097-7403.20.4.429>
- Pepperberg, I. M. (1996). Categorical class formation by an African grey parrot (*Psittacus erithacus*). In T. R. Zentall & P. M. Smeets (Eds.), *Advances in psychology: Stimulus class formation in humans and animals* (pp. 71–90). Amsterdam, the Netherlands: Elsevier. [http://dx.doi.org/10.1016/S0166-4115\(06\)80104-4](http://dx.doi.org/10.1016/S0166-4115(06)80104-4)
- Pepperberg, I. M. (1999). *The Alex studies: Cognitive and communicative abilities of grey parrots* (p. 434). Cambridge, MA: Harvard University Press.
- Pepperberg, I. M., & Brezinsky, M. V. (1991). Acquisition of a relative class concept by an African gray parrot (*Psittacus erithacus*): Discriminations based on relative size. *Journal of Comparative Psychology*, 105, 286–294. <http://dx.doi.org/10.1037/0735-7036.105.3.286>
- Porter, D., & Neuringer, A. (1984). Music discriminations by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, 10, 138–148. <http://dx.doi.org/10.1037/0097-7403.10.2.138>
- Range, F., Aust, U., Steurer, M., & Huber, L. (2008). Visual categorization of natural stimuli by domestic dogs. *Animal Cognition*, 11, 339–347. <http://dx.doi.org/10.1007/s10071-007-0123-2>
- Roitblat, H. L., Penner, R. H., & Nachtigall, P. E. (1990). Matching-to-sample by an echolocating dolphin (*Tursiops truncatus*). *Journal of Experimental Psychology: Animal Behavior Processes*, 16, 85–95. <http://dx.doi.org/10.1037/0097-7403.16.1.85>
- Ruddell, S. J. S., Twiss, S. D., & Pomeroy, P. P. (2007). Measuring opportunity for sociality: Quantifying social stability in a colonially breeding phocid. *Animal Behaviour*, 74, 1357–1368. <http://dx.doi.org/10.1016/j.anbehav.2007.01.024>
- Savage-Rumbaugh, E. S., Murphy, J., Sevcik, R. A., Brakke, K. E., Williams, S. L., Rumbaugh, D. M., & Bates, E. (1993). Language comprehension in ape and child. *Monographs of the Society for Research in Child Development*, 58, 1–252. <http://dx.doi.org/10.2307/1166068>
- Schrier, A. M., & Brady, P. M. (1987). Categorization of natural stimuli by monkeys (*Macaca mulatta*): Effects of stimulus set size and modification of exemplars. *Journal of Experimental Psychology: Animal Behavior Processes*, 13, 136–143. <http://dx.doi.org/10.1037/0097-7403.13.2.136>
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: Evidence of predator classification and semantic communication. *Science*, 210, 801–803. <http://dx.doi.org/10.1126/science.7433999>
- Shapiro, A. D., Slater, P. J. B., & Janik, V. M. (2004). Call usage learning in gray seals (*Halichoerus grypus*). *Journal of Comparative Psychology*, 118, 447–454. <http://dx.doi.org/10.1037/0735-7036.118.4.447>
- Soltis, J., Leong, K., & Savage, A. (2005). African elephant vocal communication: I. Antiphonal calling behaviour among affiliated females. *Animal Behaviour*, 70, 579–587. <http://dx.doi.org/10.1016/j.anbehav.2004.11.015>
- Vergara, V. (2011). Acoustic communication and vocal learning in belugas (*Delphinapterus leucas*) (PhD thesis) The University of British Columbia, British Columbia, Canada.
- von Helversen, D. (2004). Object classification by echolocation in nectar feeding bats: Size-independent generalization of shape. *Journal of Comparative Physiology: A. Neuroethology, Sensory, Neural, and Behavioral Physiology*, 190, 515–521. <http://dx.doi.org/10.1007/s00359-004-0492-9>
- Wasserman, E. A., & Bhatt, R. S. (1992). Conceptualization of natural and artificial stimuli by pigeons. In W. K. Honig & J. G. Fetterman (Eds.), *Cognitive aspects of stimulus control* (pp. 203–223). Hillsdale, NJ: Erlbaum.
- Cook, R. G., and Wasserman, E. A. (2006). Relational discrimination learning in pigeons. In E. A. Wasserman & T. R. Zentall (Eds.), *Comparative cognition* (pp. 307–324). New York, NY: Oxford University Press. <http://dx.doi.org/10.1093/oxfordhb/9780195392661.013.0028>

- Watanabe, S. (2001). Van Gogh, Chagall and pigeons: Picture discrimination in pigeons and humans. *Animal Cognition*, 4, 147–151. <http://dx.doi.org/10.1007/s100710100112>
- Weary, D. M., & Krebs, J. R. (1992). Great tits classify songs by individual voice characteristics. *Animal Behaviour*, 43, 283–287. [http://dx.doi.org/10.1016/S0003-3472\(05\)80223-4](http://dx.doi.org/10.1016/S0003-3472(05)80223-4)
- Wright, A. A., Cook, R. G., Rivera, J. J., Sands, S. F., & Delius, J. D. (1988). Concept learning by pigeons: Matching to sample with trial-unique video picture stimuli. *Animal Learning & Behavior*, 16, 436–444. <http://dx.doi.org/10.3758/BF03209384>
- Wright, A. A., Rivera, J. J., Katz, J. S., & Bachevalier, J. (2003). Abstract-concept learning and list-memory processing by capuchin and rhesus monkeys. *Journal of Experimental Psychology: Animal Behavior Processes*, 29, 184–198. <http://dx.doi.org/10.1037/0097-7403.29.3.184>

Received December 16, 2014

Revision received July 20, 2015

Accepted July 24, 2015 ■

Call for Nominations

The Publications and Communications (P&C) Board of the American Psychological Association has opened nominations for the editorships of *Emotion*; *Experimental and Clinical Psychopharmacology*; *Journal of Comparative Psychology*; *Journal of Experimental Psychology: Human Perception and Performance*; *Journal of Experimental Psychology: Applied*; *Journal of Abnormal Psychology*; *Journal of Personality and Social Psychology: Attitudes and Social Cognition*; *Journal of Counseling Psychology* and *Rehabilitation Psychology* for the years 2018–2023. David DeSteno, PhD; Suzette Evans, PhD; Josep Call, PhD; James T. Enns, PhD; Neil Brewer, PhD; Sherry Goodman, PhD; Eliot Smith, PhD; Terence Tracey, PhD and Stephen Wegener, PhD respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2017 to prepare for issues published in 2018. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

Search chairs have been appointed as follows:

- *Emotion*, Co-chairs: Pamela T. Reid, PhD and Jennifer Crocker, PhD
- *Experimental and Clinical Psychopharmacology*, Chair: Mark Sobell, PhD
- *Journal of Comparative Psychology*, Co-Chairs: Stephen Rao, PhD and Gary VandenBos, PhD
- *Journal of Experimental Psychology: Human Perception and Performance*, Co-chairs: Wendy A. Rogers, PhD and Gary VandenBos, PhD
- *Journal of Experimental Psychology: Applied*, Chair: Neal Schmitt, PhD
- *Journal of Abnormal Psychology*, Chair: Annette La Greca, PhD
- *Journal of Personality and Social Psychology: Attitudes and Social Cognition*, Chair: David Dunning, PhD
- *Journal of Counseling Psychology*, Chair: Kate F. Hays, PhD
- *Rehabilitation Psychology*, Co-chairs: James C. Quick, PhD and Gary VandenBos, PhD

Candidates should be nominated by accessing APA's *EditorQuest* site on the Web. Using your Google Chrome Web browser, go to <http://editorquest.apa.org>. On the Home menu on the left, find "Guests." Next, click on the link "Submit a Nomination," enter your nominee's information, and click "Submit."

Prepared statements of one page or less in support of a nominee can also be submitted by e-mail to Ieshia Haynie, P&C Board Search Liaison, at ilhaynie@apa.org.

Deadline for accepting nominations is Friday, January 29, 2016, after which phase one vetting will begin.