JASA ARTICLE

How humans discriminate acoustically among bottlenose dolphin signature whistles with and without masking by boat noise[a),b)]

Evan L. Morrison,[1,c)] Caroline M. DeLong,[1] and Kenneth Tyler Wilcox[2]
[1]*Department of Psychology, College of Liberal Arts, Rochester Institute of Technology, 18 Lomb Memorial Drive, Rochester, New York 14623, USA*
[2]*Department of Psychology, College of Arts and Letters, University of Notre Dame, 390 Corbett Family Hall, Notre Dame, Indiana 46556, USA*

ABSTRACT:
Anthropogenic noise in the world's oceans is known to impede many species' ability to perceive acoustic signals, but little research has addressed how this noise affects the perception of bioacoustic signals used for communication in marine mammals. Bottlenose dolphins (*Tursiops truncatus*) use signature whistles containing identification information. Past studies have used human participants to gain insight into dolphin perception, but most previous research investigated echolocation. In Experiment 1, human participants were tested on their ability to discriminate among signature whistles from three dolphins. Participants' performance was nearly errorless. In Experiment 2, participants identified signature whistles masked by five different samples of boat noise utilizing different signal-to-noise ratios. Lower signal-to-noise ratio and proximity in frequency between the whistle and noise both significantly decreased performance. Like dolphins, human participants primarily identified whistles using frequency contour. Participants reported greater use of amplitude in noise-present vs noise-absent trials, but otherwise did not vary cue usage. These findings can be used to generate hypotheses about dolphins' performance and auditory cue use for future research. This study may provide insight into how specific characteristics of boat noise affect dolphin whistle perception and may have implications for conservation and regulations.
https://doi.org/10.1121/10.0001450

(Received 9 December 2019; revised 10 May 2020; accepted 31 May 2020; published online 25 June 2020)
[Editor: Mark A. Bee] Pages: 4162–4174

I. INTRODUCTION

Bottlenose dolphins (*Tursiops truncatus*) and other similar species rely heavily on vocalizations for communication and social behavior (Harley, 2008; Sayigh *et al.*, 2007). There is evidence that boat traffic and noise disrupt dolphin behavior (Lusseau, 2003; Pirotta *et al.*, 2015), but it is unclear what the direct impact is of anthropogenic noise on dolphins' ability to communicate vocally. Current regulations prohibit disruption of the natural behavior of marine mammals (Marine Mammal Protection Act, 1972), but in order to properly understand what constitutes such a disruption, it is imperative to understand human impacts on the cognition, behavior, and perception of marine mammals.

Whistles, one of the many vocalizations produced by dolphins, seem to be used for a variety of functions, including group cohesion and coordination (Herzing, 1996; Quick and Janik, 2008). A subset of these vocalizations, called signature whistles, are used by both humans and dolphins to identify individual dolphins (Gridley *et al.*, 2014; Janik and Sayigh, 2013; Janik *et al.*, 2006; Janik and Slater, 1998; Quick and Janik, 2008, 2012; Sayigh *et al.*, 2007; Sayigh *et al.*, 1990; Sayigh *et al.*, 1999). The signature whistle hypothesis states that dolphins produce unique, stereotyped whistles, especially when isolated from conspecifics (Watwood *et al.*, 2004). Signature whistles are most often produced by the dolphin that they identify (Janik and Sayigh, 2013; Sayigh *et al.*, 2007). Signature whistles typically range in frequency from about 1 to 30 kHz (Sayigh and Janik, 2010), with durations of 0.1 to 4 s (Buckstaff, 2004).

Signature whistles vary on several dimensions, including timbre (also referred to as sound quality or voice cues), duration, amplitude, and frequency contour (Gridley *et al.*, 2014; Harley, 2008; Janik *et al.*, 1994; Janik and Sayigh, 2013; Janik *et al.*, 2006; Kershenbaum *et al.*, 2013). However, dolphins do not seem to consider timbre when discriminating signature whistles (Janik *et al.*, 2006; King and Janik, 2013; Sayigh *et al.*, 2017). Past studies indicate that frequency contour (i.e., the change in frequency over the duration of a sound) may be the most crucial identifying element in a signature whistle for dolphins (Harley, 2008;

[a)]Portions of this work were presented in "Human auditory discrimination of bottlenose dolphin signature whistles masked by noise: Investigating perceptual strategies for anthropogenic noise pollution," Meeting of the Acoustical Society of America, Louisville, KY, May 2019.
[b)]Portions of this work were presented in "Human auditory discrimination of bottlenose dolphin signature whistles masked by noise: Investigating perceptual strategies for anthropogenic noise pollution," International Conference on Comparative Cognition, Melbourne, FL, April 2019.
[c)]Electronic mail: elm3948@rit.edu, ORCID: 0000-0001-5097-6827.

Janik *et al.*, 1994; Janik *et al.*, 2013; Janik *et al.*, 2006; Sayigh *et al.*, 2007) and for humans discriminating among whistle-like sounds (Branstetter *et al.*, 2016b). Humans, both researchers and study participants, have successfully been able to distinguish individual dolphins' signature whistles visually using spectrograms (e.g., Caldwell and Caldwell, 1965; Gridley *et al.*, 2014; Harley, 2008; Janik, 1999; Janik *et al.*, 2013). However, no prior studies have explored how human participants can discriminate acoustically among recorded signature whistles in a laboratory context.

When researching dolphin auditory processing, humans may serve as a valuable model organism and subject for comparative research (Au and Martin, 1989; Branstetter *et al.*, 2016b; DeLong, 2017; DeLong *et al.*, 2007a; DeLong *et al.*, 2007b; DeLong *et al.*, 2014). Most previous studies using this technique have compared the ability of dolphins and humans to discriminate among recorded object echoes to investigate dolphin echolocation (e.g., Au and Martin, 1989; DeLong, 2017; DeLong *et al.*, 2007a; DeLong *et al.*, 2007b; DeLong *et al.*, 2014; Fish *et al.*, 1976). Many of these studies have also included questionnaires that have asked human participants what auditory features of the sounds they attended to when performing the discrimination tasks (Au and Martin, 1989; Branstetter *et al.*, 2016b; DeLong, 2017; DeLong *et al.*, 2007a; DeLong *et al.*, 2007b; DeLong *et al.*, 2014; Gorman and Sawatari, 1985). By analyzing human feedback and comparing performance between humans and dolphins (specifically error patterns), researchers have been able to generate informed hypotheses about which auditory features of echoes or whistles dolphins may use in their own auditory processing. For example, Branstetter and colleagues (2016b) investigated both species' ability to discriminate whistle-like sounds when those sounds were transformed in amplitude, duration, and frequency. Humans successfully identified sounds transformed in all three of these dimensions, while dolphins could not identify sounds transformed in frequency under some conditions (plus or minus half octaves). Human participants reported using frequency contour more than absolute frequency when identifying the sounds, and used both of these cues more than the other possible cues. Considering both comparative performance and human interview responses, the authors suggested that dolphins may use some cues in ways similar to humans, but may rely more heavily upon absolute frequency. This body of comparative research has shown that humans perform at least as well as dolphins at auditory discrimination tasks using echoes or whistle-like stimuli and that humans report attending to specific auditory features of the stimuli, depending on the task (Au and Martin, 1989; Branstetter *et al.*, 2016b; DeLong, 2017; DeLong *et al.*, 2007a; DeLong *et al.*, 2007b; DeLong *et al.*, 2014; Helweg *et al.*, 1995).

Comparative studies of dolphin and human auditory perception may be a useful tool for investigating the effects of masking noise on perception. Previous studies of masking noise and its impact on human auditory perception have primarily focused on perception of human speech and of pure tones masked by simple noise such as pure masking tones and broadband noise (e.g., Egan and Hake, 1950; Green, 1960; Shah *et al.*, 1999; Marshall and Jesteadt, 1986; Muller-Gass *et al.*, 2001; Nelken *et al.*, 1999). Masking noise increases the thresholds at which humans will perceive a signal, and noise closer in frequency to a signal will have greater impact on the perception of that signal (Egan and Hake, 1950). Evidence has suggested that, in humans, masking can vary depending on what is being masked (e.g., speech or non-speech sounds) and what sort of noise is masking it, with effects such as phonemic restoration occurring when speech is masked by certain types of noise (Warren, 1984; Warren and Obusek, 1971). More recent research has determined that human auditory cortices process simultaneous sources of background noise based on the frequency bands those noises occupy (Nelken *et al.*, 1999). This suggests that humans may process narrow-band noises better and supports the concept that masking noise will be more impactful when its frequency overlaps with the signal.

Anthropogenic noise may be a major source of masking for marine organisms. Marine anthropogenic noise has a broadly detrimental impact on perception, behavior, and, in extreme circumstances, well-being in a diverse array of marine animals (Würsig and Richardson, 2009). In dolphins, which are extremely hearing-reliant, noise has the potential to be particularly impactful. Shipping and other watercraft are the greatest source of marine anthropogenic noise, especially between 1 and 30 kHz, the frequency range of signature whistles (Albuquerque and Souto, 2013; Ross, 2005). Bottlenose dolphins' natural behaviors are significantly altered by the presence of boats across a wide range of vessel sizes and numbers, even after the vessels depart (Lusseau, 2003; Pirotta *et al.*, 2015). In addition, wild dolphins produce their signature whistles significantly more when approached by marine vessels (Buckstaff, 2004). These findings may be indicative of interference generated by the boats' noise or arousal due to the boats' presence.

Studies of dolphins' ability to echolocate in masking noise suggest that, like in humans, the effects of masking noise vary with both the frequency and amplitude of both the noise and the signal (e.g., Au *et al.*, 1988; Au and Penner, 1981; Branstetter *et al.*, 2016a; Branstetter *et al.*, 2013a). However, masking may affect dolphin perception differently at the lower frequencies dolphins typically use for communicative vocalizations. Dolphins have a constant Q auditory filter system below 40–60 kHz and above 100–120 kHz, with a constant-bandwidth auditory filter system between 40 to 60 kHz and 100–120 kHz, believed to be specialized for echolocation (Lemonds *et al.*, 2011; Lemonds *et al.*, 2012). Humans and dolphins share similar critical ratios within the human range of best hearing, with dolphin critical ratios ranging from about 15 to 35 dB between 0 and 35 kHz (Lemonds *et al.*, 2011; Johnson, 1968). Given the specialized nature of dolphins' auditory filter shape, it is unsurprising that some types of dolphin vocalizations are affected differently by different masking noises (Branstetter *et al.*, 2016a). This disparity in masking patterns across different noise types emphasizes the need for further study of specific types of signals and noises, especially common signals such as signature whistles. Due to the fact that many populations of wild dolphins frequent areas that are exposed to anthropogenic noise with the potential to mask their communication (Albuquerque and Souto, 2013; David, 2006; Kaplan and Mooney, 2015), it is important to understand how such noise might affect their ability to communicate.

The current study utilized human listeners to gain insight into auditory perception in dolphins, similar to previous research using this technique (e.g., Au and Martin, 1989; Branstetter *et al.*, 2016b; DeLong, 2017; DeLong *et al.*, 2007a; DeLong *et al.*, 2007b; DeLong *et al.*, 2014). Though there are many extant species of dolphins, this research focused on the Atlantic bottlenose dolphin, a widely ranging species with well-studied perceptual and cognitive abilities (e.g., Au, 1993; Mercado and DeLong, 2010). This study specifically used signature whistles, as their repeated, stereotyped nature made it easy to acquire multiple instances of each whistle and because they are particularly well studied (e.g., Janik *et al.*, 1994; Quick and Janik, 2012; Sayigh and Janik, 2010; Sayigh *et al.*, 1999; Sayigh *et al.*, 2017). In addition, signature whistles make up approximately half of all bottlenose dolphin whistles in the wild (Buckstaff, 2004; Cook *et al.*, 2004). The current study was the first to investigate human auditory perception of recorded signature whistles in a laboratory setting. The study also endeavored to determine to what extent anthropogenic noise from coastal watercraft may impede human auditory discrimination of dolphin signature whistles. One advantage of the current study was the use of sample boat noises from an area where dolphins are relatively common: the U.S. Virgin Islands (Kaplan and Mooney, 2015).

We hypothesized that humans would be able to discriminate auditorily between recorded signature whistles with high accuracy, since humans have previously discriminated successfully among whistle-like sounds under similar conditions (Branstetter *et al.*, 2016b). We hypothesized that low signal-to-noise ratio would result in poorer performance (e.g., Egan and Hake, 1950). We also predicted that boat noise with frequencies closer to the frequencies of the signature whistles would most severely impair discrimination, as previous work suggests that the frequency of masking noise plays a role in its impact on hearing thresholds (Egan and Hake, 1950; Nelken *et al.*, 1999).

II. EXPERIMENT 1

The first experiment investigated to what extent human participants were able to auditorily discriminate among signature whistles. Participants discriminated among signature whistles from three dolphins without boat noise. Participants then were interviewed about which auditory cues they attended to during the listening task.

A. Method

1. Participants

Fourteen participants (seven male and seven female) participated in Experiment 1, approximately the same number as in previous studies (e.g., Au and Martin, 1989; Branstetter *et al.*, 2016b; DeLong, 2017). Participants ranged in age from 18 to 23 yr ($M = 20$ yr). Participants were recruited using the Rochester Institute of Technology (RIT) online participant recruitment system and paper flyers, and were compensated for their time with their choice of course credit or $10.

Only individuals with normal hearing were able to participate, as this study involved discriminating among auditory stimuli. Participants were screened using a standardized hearing test (Home Audiometer,) to ensure adequate hearing ability at 125, 250, 500, 750, 1000, 1500, 2000, 3000, 4000, 6000, and 8000 Hz in both ears. All participants successfully passed this test. Of the 14 participants, one reported having had a previous hearing problem (preferred not to give further information), none reported having any current hearing-related difficulties, and none reported a family history of hearing loss.

2. Materials

This study utilized Audacity® 2.22 () for the editing and assembly of auditory stimuli. Participants used Bose® On-Ear headphones (Bose Corporation, Framingham, MA) to listen to stimuli. Qualtrics® was used as the software platform for presenting the training and testing phases on a 14-inch IdeaPad Y700 laptop computer (Lenovo Group Ltd., Beijing, China). Sound pressure levels were standardized using Audacity® 2.22 and checked using a digital sound level meter (Model SLM01, Tacklife, Levittown, NY).

The auditory stimuli consisted of signature whistles from three individual adult male dolphins, which were referred to throughout the study as "Dolphin A," "Dolphin B," and "Dolphin C." All three dolphins resided in the same facility in Florida when the whistles recordings were made. Dolphins A and B shared a father. Dolphin A was 14 yr old, Dolphin B was 17 yr old, and Dolphin C was 27 yr old at the time their signature whistles were recorded. There were six exemplars of signature whistles from each dolphin (one for training, five for testing), for a total of 18 signature whistle files (see Fig. 1 for exemplars from each dolphin). Signature whistles were recorded using a hydrophone array with a sampling rate of at least 62.5 kHz. Signature whistles were primarily determined by inspecting spectrograms of recordings and visually identifying the most common whistle each dolphin produced in isolation, similar to many previous studies (e.g., Janik and Sayigh, 2013). Some whistles were not recorded in isolation, but could be identified as matching a dolphin's signature whistle by the whistle's frequency contour.

An entry survey was used to collect demographic information about participants, including information about their

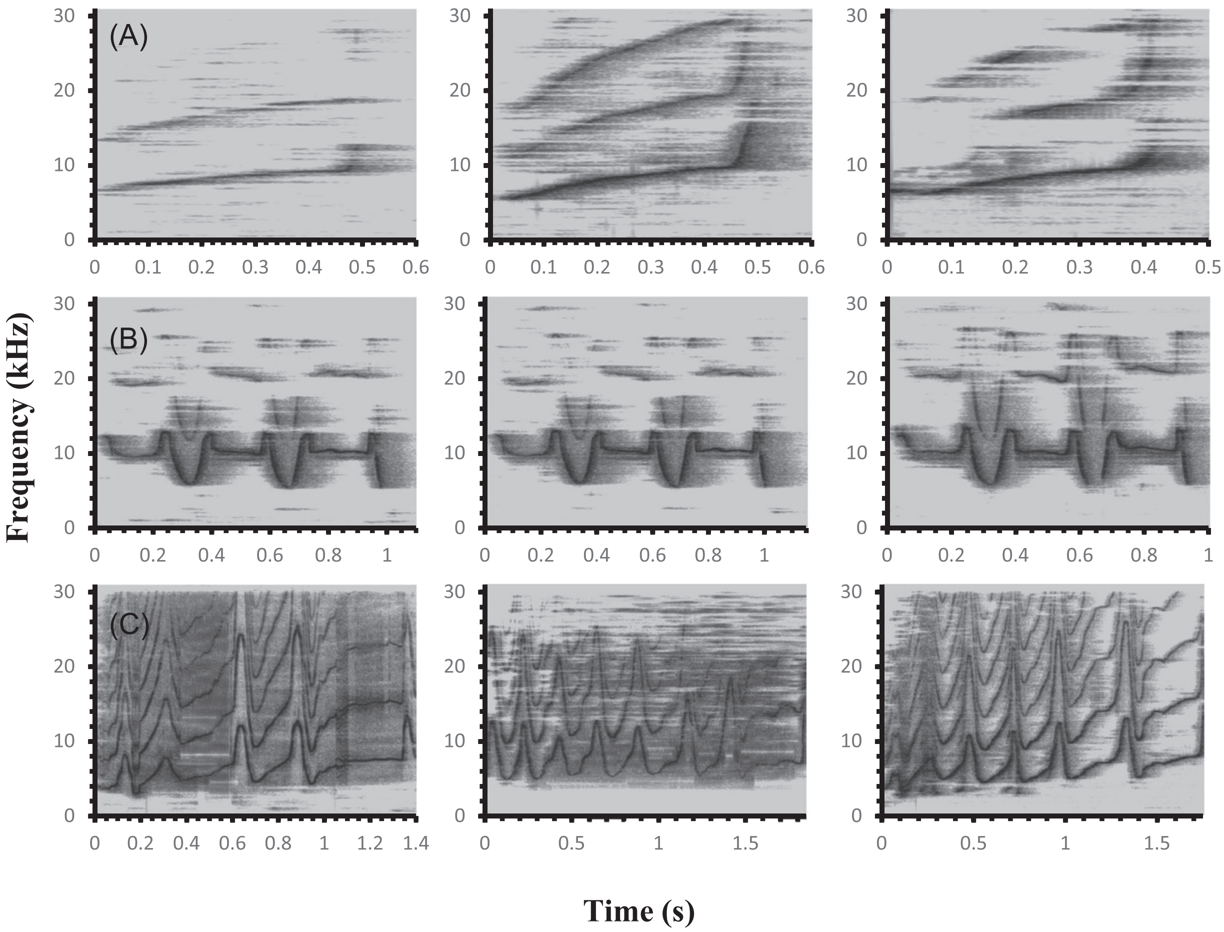

FIG. 1. Three exemplars of Dolphin A, B, and C's signature whistles used during the experiments (six total exemplars were used for each dolphin). Note that the duration of shown whistles varied. All FFT sizes equaled 1024 samples. (A) Dolphin A's signature whistle consisted of a segment with relatively flat frequency followed by an upsweep. (B) Dolphin B's signature whistle contained segments of relatively constant frequency interspersed with frequency-modulated segments. (C) Dolphin C's whistle contained repeated sequences (loops) that were all similar in frequency contour and duration, but the overall number of these loops were not the same for every sample of Dolphin C's signature whistle, so the samples of this signature whistle were not always of the same duration.

prior and current hearing status and family history of hearing loss. During the training phase, participants were given a printed handout with written definitions and images that described several terms relevant to auditory perception: amplitude, duration, frequency, frequency contour, and timbre. All terms were accompanied with graphics depicting two sounds, which differed in regards to the acoustic term. Each illustration was paired with matching audio samples, typically simple tones, which participants listened to in tandem with the experimenter's verbal explanation of each term. The experimenter encouraged participants to also attend to cues that were not necessarily on their vocabulary sheet and expressed that participants were encouraged to discuss any such additional cues during the interview. Interviews given to the participant after the listening task were comprised of questions that pertained to which auditory cues were used by the participants, which two dolphins they most often confused, and a rating of how often they used each cue overall on a Likert scale from 1 to 7 (1 = they never used the cue, 4 = they used the cue in about half of the trials, and 7 = they used the cue in every trial).

3. Procedure

A single experimenter tested participants in a sound-treated chamber (Controlled Acoustical Environments, Industrial Acoustics Company, Inc., Bronx, NY) on the RIT campus. For the duration of the experiment, the experimenter sat opposite the participant, and the computer faced toward the experimenter and away from the participant. The experimenter remained present throughout the experiment, which took approximately 60 min per participant. Participants all heard the same verbal instructions from the experimenter throughout the training, testing, and interview phases.

a. Training phase. The experimenter played one sample whistle from each dolphin. These samples were a sixth whistle from each dolphin, not used during the testing phase. Each of these sample whistles was played three times for every participant, and the participants were encouraged to ask the experimenter to play any of the whistles additional times until the participant felt confident in their ability to discriminate among the three whistle stimuli. In Experiment 1, participants listened 3.2 times per whistle, on average. The participants were encouraged to note which acoustic features allowed them to more easily identify the whistles.

b. Testing phase. For the testing phase, the experimenter played each whistle stimulus, recorded the participant's response, informed the participant whether they were correct or incorrect, then immediately began the next trial. The experimenter played each whistle only once per trial. Participants were presented with all five test exemplar whistles for each of the three dolphins, with each whistle occurring 16 times during the testing phase. In total, participants underwent 240 trials in the testing phase. There was a total of 16 consecutive blocks, each made up of 15 trials with one sample of each exemplar from each dolphin, and the order of trials within each block of 15 trials were randomized. The order in which the trials were presented was randomized separately for each participant. After completing 120 of the 240 trials, participants were given a two-minute break during which they were asked to take off their headphones, in order to avoid possible effects of fatigue.

c. Interview. After the completion of the testing phase, the experimenter administered a brief verbal interview. First, the experimenter played one whistle from each dolphin, the same exemplar as was used during the training session. The experimenter then led a structured interview in which the participant answered questions about the acoustic features of the whistles they utilized (see Sec. II A 2).

4. Data analyses

A multilevel logistic regression model (Raudenbush and Bryk, 2002) was used to analyze the accuracy of participant responses with an experimental type I error level of $\alpha = 0.05$. Since we used a within-subjects design, intraindividual correlation was accounted for by including random intercepts. *Post hoc* analyses were conducted to further investigate significant effects using Tukey's (1949) HSD for multiple comparisons. In order to test the hypothesis that participants would be able to discriminate auditorily among dolphin whistles, the model-based marginal mean was used to test the overall accuracy against chance (approximately 33.33%).

B. Results

1. Performance accuracy

Participants in Experiment 1 achieved near-perfect performance on the signature whistle discrimination task ($M = 98.8\%$, $SE = 0.2\%$). Participant performance was significantly better than chance ($Z = 12.69$, $p < 0.001$). There were significant differences in participant accuracy among dolphins ($\chi^2(2) = 11.19$, $p = 0.004$). Performance was significantly better when participants were identifying Dolphin A ($M = 99.8\%$, *95% CI:* [99.2%, 100.0%]) compared to Dolphin B ($M = 99.2\%$, *95% CI:* [97.9%, 99.7%]) and Dolphin C ($M = 98.7\%$, *95% CI:* [96.4%, 99.4%]; both $p < 0.05$), and performance did not differ significantly between Dolphin B and Dolphin C ($p = 0.369$). Due to near-zero variance in performance for some of the blocks, it was impossible to test the main effect of block on performance or the interaction effect between block and dolphin. Though participant performance in these conditions could not be compared using inferential statistics, the mean performances did increase after the first test block. Average performance in block 1 was 91.4%, and average performance on blocks 2 through 16 ranged from 97.6% to 100%.

2. Interview responses

All participants reported hearing a difference between the three dolphins' whistles. Of the 14 participants, all 14 reported a difference in frequency contour, 13 reported a difference in timbre, 12 reported a difference in duration, 11 reported a difference in frequency, 7 reported a difference in amplitude, and 4 reported some other difference. One reported "other" difference, for example, was the perceived harshness of the whistle, expressing a degree of emotional valence to at least one of the whistles. The majority of participants in Experiment 1 (12 of 14) reported the whistles of Dolphins B and C as being the most similar and easily confused. Participants were most likely to report frequency contour either as the primary source of confusion or one of two equally confusing factors (11 of 14 participants). Participants' ratings of how frequently they used each auditory cue during the Experiment 1 listening task are given in Fig. 2(A). Participants generally claimed to use frequency contour most frequently.

C. Discussion

One of the primary objectives of Experiment 1 was to test how well human participants could auditorily discriminate among the recorded signature whistles of the three dolphins. In this experiment, there was no masking noise present, and whistle amplitude was high (approximately 55 dB). As predicted, participants were able to discriminate among the signature whistles. Participants under these conditions exhibited near-perfect performance on the signature whistle discrimination task, achieving an average accuracy of almost 99%.

Participants' responses in the interview seemed to indicate that, in this experiment, frequency contour was consistently the most helpful cue. This is consistent with Branstetter and colleagues' (2016b) study on human and dolphin perception of whistle-like sounds, which showed that humans were capable of discriminating among whistle-like stimuli with near-perfect accuracy and cited frequency contour as the most useful cue during the task. Evidence from studies using artificial neural networks have indicated that frequency contour is key for identifying signature whistles (e.g., Janik, 1999). Finally, humans have previously been shown to be able to discriminate among signature whistles visually, using spectrograms, even when frequency

JASA https://doi.org/10.1121/10.0001450

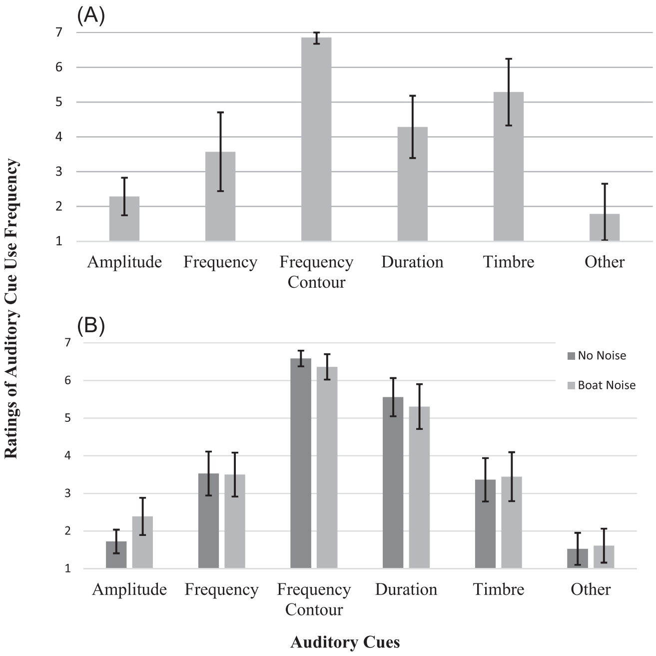

FIG. 2. Participants' ratings of how often they used each auditory cue to help identify the whistle in a given trial in both experiments. Experiment 1 is shown in (A) and Experiment 2 is shown in (B). For each cue, participants were asked to give a number from one to seven, where one would indicate that they never used the cue in question, four would indicate that they used it in about half of the trials, and seven would indicate that they used it for every trial. Error bars indicate 95% confidence intervals.

contour is the only cue visualized in the spectrograms (e.g., Sayigh *et al.*, 2007). Therefore, it is unsurprising to find that participants in Experiment 1 cited frequency contour as being the cue that they used most frequently to help identify the dolphin whistles during the task.

III. EXPERIMENT 2

The second experiment investigated how the presence of boat noise affected participants' ability to auditorily discriminate among signature whistles. Participants discriminated among the same signature whistles used in Experiment 1, with eight signal-to-noise ratios. Participants were then interviewed about which auditory cues they attended to during the test phase, in both the presence and absence of boat noise.

A. Method

1. Participants

Thirty-six participants (21 male, 14 female, 1 genderqueer) participated in this study. Participants ranged from 18 to 43 yr in age ($M = 21$ yr). Participants were recruited from the same population as those in Experiment 1 using the same recruitment techniques, with no participants partaking in both experiments. Participants underwent the same screening procedure and received the same compensation. All participants passed the hearing test and screening questionnaire. Of the 36 participants, none reported having had a previous hearing problem, one reported having difficulty hearing a specific type of sound (quiet, low-frequency voices; this participant was not removed because they passed the hearing test), and none reported any family history of hearing loss.

2. Materials

This experiment used the same materials as Experiment 1, with the exception of the auditory stimuli used during the testing phase and additional questions added to the interview in regards to the trials with masked whistles. This study utilized noise samples to mask three quarters of the auditory stimuli. Boat noise samples were recorded by Kaplan and Mooney (2015) in the U.S. Virgin Islands National Park. Sounds were collected from three reefs and were sampled for about one minute every two hours, providing samples from a variety of times of day. Recordings were made using autonomous underwater recording devices with hydrophones positioned approximately 0.3 m above the sea floor. Boat noises were then detected within sample recordings using visual identification from spectrograms and auditory confirmation. Six samples of boat noise, all of which lasted more than two seconds (long enough to fully mask all signature whistle exemplars), were selected for use in Experiment 2 (Table I). See Fig. 3(B) for spectrograms of three of the boat noise samples.

Whistle stimuli were presented both with and without masking noise from boats. Audio files were created that contained simultaneous samples of boat noise and dolphin whistles (Fig. 3). The amplitudes of the individual tracks containing either boat noise or signature whistles were adjusted to create different ratios of signal-to-noise, based on the maximum amplitude of each sample. There were two levels of amplitude for the signature whistles: low amplitude (50 dB) and medium amplitude (60 dB). There were four levels of amplitude for the noise samples: no noise, low amplitude (60 dB), medium amplitude (65 dB), and high amplitude (70 dB), for a total of eight signal-to-noise ratios. These signal-to-noise ratios were chosen based on pilot testing to elicit a range of performance from chance to near-perfect accuracy.

The interview for Experiment 2 included sets of questions about which auditory cues participants used during the listening task, which were the same as those used in the Experiment 1 interview. However, these same questions were asked in regard to which cues they used in the absence of noise for the first part of the interview, and in the presence of noise for the second part. Before part one of the interview, participants were played each dolphin's training whistle without background noise, and before part two, participants were played the same whistles with background noise.

3. Procedure

Experiment 2 used the same procedure as Experiment 1, with exceptions noted below. Experiment 2 typically took a total of about an hour and 15 min to complete.

a. Training phase. The training procedure was exactly the same as in Experiment 1. Participants listened to each sample whistle an average of 3.5 times during training.

TABLE I. Description of the samples of boat noise used. The local time where recordings were collected was 4 hours behind Coordinated Universal Time (UTC) time. Peak amplitude of peak frequency is the maximum amplitude in decibels at the frequency that achieves the greatest amplitude over the duration of the noise sample.

Sample	Date and Time of Recording (UTC)	Peak Amplitude (dB)	Peak Frequency (Hz)	Peak Amplitude of Peak Frequency (dB)
N0	July 7, 2016 17:30:02	−20.310	170	−48.0
N1	May 30, 2016 21:20:02	−18.771	63	−44.4
N2	June 27, 2016 19:30:02	−18.913	186	−34.0
N3	July 3, 2016 15:20:02	−16.250	233	−42.7
N4	July 7, 2016 16:60:02	−11.995	132	−33.8
N5	July 11, 2016 17:50:02	−10.867	166	−36.8

b. Testing phase. The testing phase was the same as in Experiment 1, with a few exceptions. Participants were informed that some trials would contain other sounds as well as a signature whistle, and that every trial would contain one signature whistle for them to identify. The testing phase was divided into two blocks, with a break between them. Each block contained one trial for each whistle exemplar-boat noise pairing, at eight levels of signal and noise, for a total of 120 trials per block (240 trials total).

c. Interview. The experimenter led a structured interview in which the participant answered questions about the acoustic features they utilized during the testing phase.

4. Data analyses

To investigate the effects of different trial conditions on performance, we fit a multilevel logistic regression model with two factors, signal-to-noise ratio and test block, and two continuous variables, boat noise peak frequency and

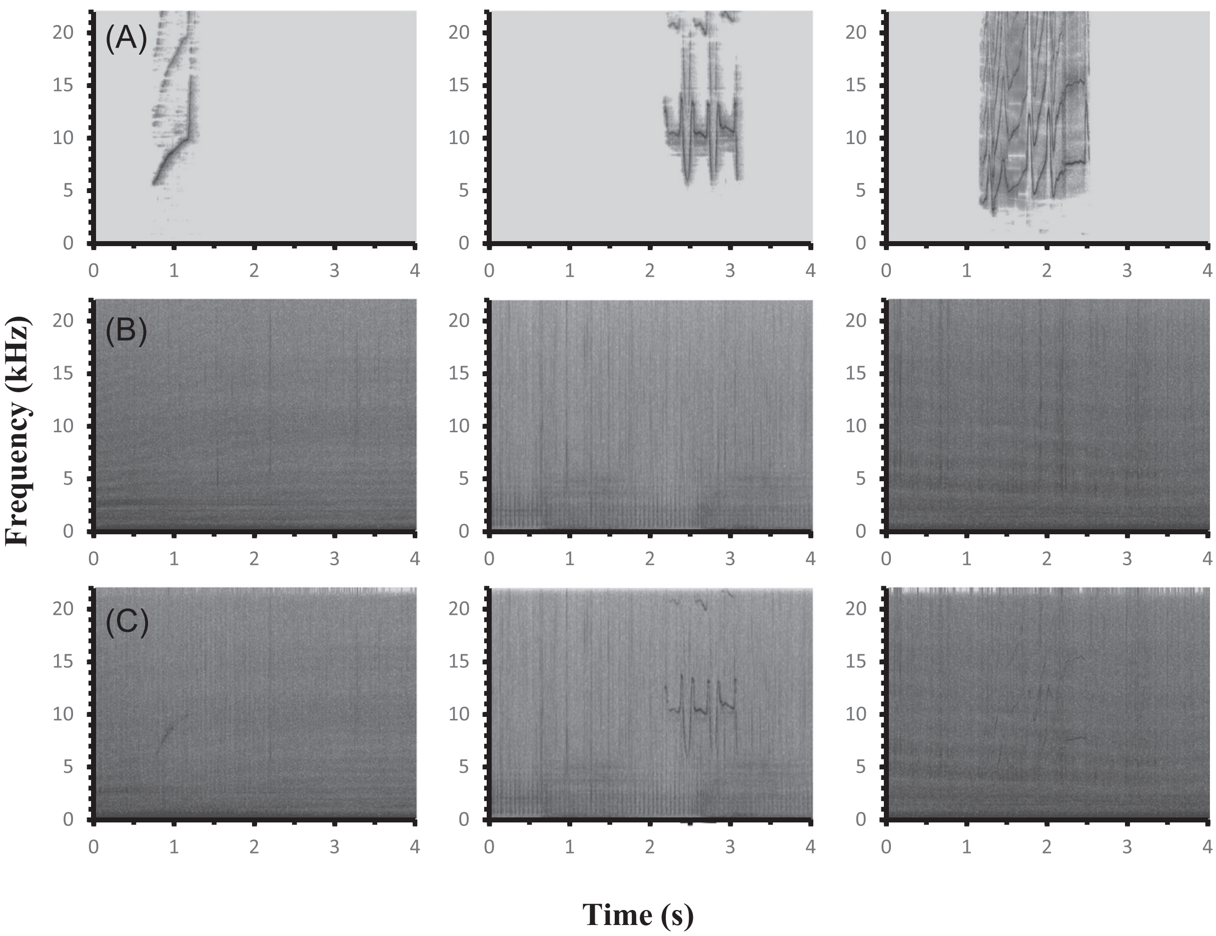

FIG. 3. Sample signature whistles (A), samples of boat noise (B), and then the two above samples merged into single sound files (C), with both the whistles and the boat noises at their original amplitudes. All FFT sizes equaled 1024 samples. Noises were generally broad-spectrum, continuous sounds. Some boat noises also had intermittent or repeated broadband sounds in quick succession. See Table I for acoustic characteristics of the boat noises.

dolphin whistle peak frequency. Since we used a within-subjects design, intraindividual correlation was accounted for by including random intercepts. We tested whether whistle and boat noise frequency and signal-to-noise ratio affected performance while controlling for test block. First-order (i.e., main and linear effects) and two-way interactions among signal frequency, noise frequency, and signal-to-noise ratio were included, but test block differences were not tested. P-values were adjusted for multiple tests using Holm's (1979) procedure. *Post hoc* analyses were conducted to further investigate significant effects using Tukey's (1949) HSD for multiple comparisons.

To test the prediction that participants would change their use of auditory cues in the presence and absence of boat noise, participants' ratings of how frequently they used each auditory cue in either noise condition during Experiment 2 were compared using a 2 (absence or presence of boat noise) × 6 (auditory cue: amplitude, frequency, frequency contour, duration, timbre, other) repeated measures analysis of variance (ANOVA). Degrees of freedom were corrected using the Greenhouse-Geisser adjustment.

B. Results

1. Performance accuracy

Participants in Experiment 2 performed well overall ($M = 80.7\%$, $SE = 0.4\%$). The three tested interactions were all statistically significant. First, there was a significant interaction effect between signal-to-noise ratio and boat noise peak frequency ($\chi^2(5) = 53.31$, $p < 0.001$). The relationship between peak frequencies of boat noises and performance was tested within each of the six levels of signal-to-noise ratio where noise was present. As shown in Fig. 4, performance decreased significantly as the frequency of boat noise increased in low signal–medium noise, low signal–high noise, and medium signal–high noise conditions. Performance did not change significantly as a function of noise frequency in the low signal–low noise and medium signa –medium noise conditions, and performance increased significantly as noise frequency increased in the medium signal–low noise condition.

Second, there was a significant interaction effect between boat noise peak frequency and dolphin whistle peak frequency ($Z = 4.79$, $b = 1.67$, $SE = 0.35$, $p < 0.001$). Performance decreased significantly as the frequency of boat noise increased for two of the three dolphin whistle peak frequencies, and the rate of decline depended on the whistle peak frequency. The three rates of decline were all significantly different from each other, all $p < 0.001$. Performance (on the log-odds scale) declined fastest ($b = -6.65$, $SE = 0.78$, *95% CI:* [−8.18, −5.13]) as boat noise frequency increased for Dolphin C (peak whistle frequency of 7.07 kHz). Performance declined significantly, but more moderately ($b = -3.97$, $SE = 0.68$, *95% CI:* [−5.30, −2.64]), for Dolphin A (peak whistle frequency of 8.68 kHz). Performance did not change

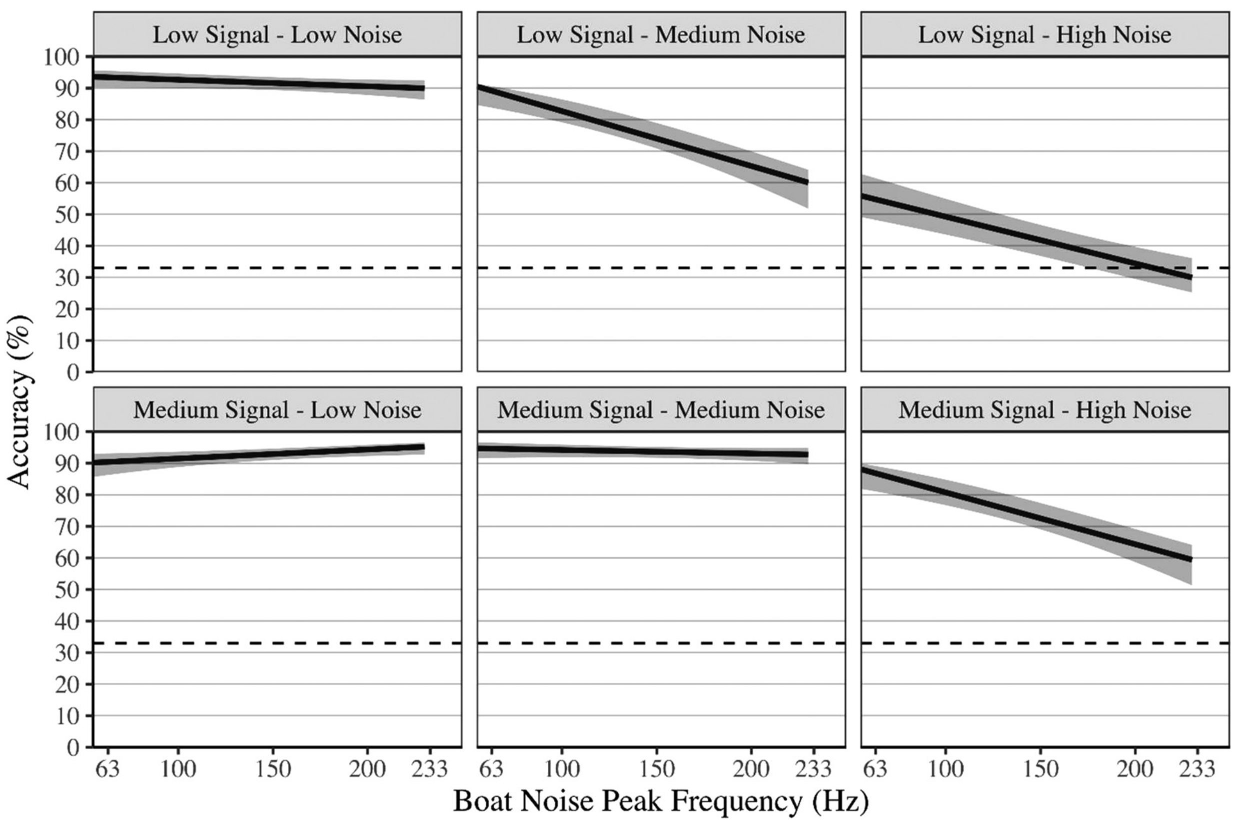

FIG. 4. Relationship between participants' performance accuracy and boat noise frequency at different levels of signal and noise in Experiment 2. Chance performance had 33% accuracy. Gray bands denote 95% confidence regions. Neither condition with no noise was included in this analysis because there was no noise frequency to analyze.

significantly ($b = -1.69$, $SE = 0.91$, *95% CI:* [−3.47, 0.10]) with increasing boat noise frequency for Dolphin B (peak whistle frequency of 10.05 kHz).

Third, there was a significant interaction between signal-to-noise ratio and dolphin whistle peak frequency ($\chi^2(7) = 154.00$, $p < 0.001$). The relationship between dolphin whistle peak frequencies and performance was tested within each level of signal-to-noise ratio. As shown in Fig. 5, performance did not change significantly as the frequency of the dolphin whistle increased for the no-noise conditions. For all noise-present conditions except low signal–high noise, performance increased significantly as the frequency of the dolphin whistle increased.

2. Interview responses

In Experiment 2, 100% of participants reported hearing a difference between the three dolphins' whistles. Thirty-five of 36 reported hearing a difference in frequency contour, 35 reported a difference in duration, 31 reported a difference in timbre, 27 reported a difference in frequency, 20 reported a difference in amplitude, and 8 reported some other difference, such as attributing personality or emotional characteristics to the dolphins' whistles (e.g., Dolphin C was described as more "excited"). Participants in Experiment 2 most commonly reported confusing Dolphins B and C, and participants were most likely to report that frequency contour was one of the most confusing similarities between the whistles. Participants' ratings of how frequently they used each auditory cue to identify the dolphin whistles with and without boat noise present during Experiment 2 are shown in Fig. 2(B). Participants used frequency contour, followed by duration, as their most frequent listening cues.

In the ANOVA, there was a significant main effect for auditory cue, $F(3.57, 125.10) = 65.71$, $p < 0.001$, but no significant effect for the presence or absence of boat noise, $F(1, 35) = 0.42$, $p > 0.05$. The main effect of auditory cue could not be interpreted because there was also a significant interaction effect between the presence of boat noise and auditory cue, $F(3.16, 110.43) = 2.72$, $p = 0.045$. In both the presence and the absence of noise, frequency contour and duration were the most-used cues for whistle discrimination. There was a small but significant increase in participants' reported attendance to the whistles' amplitudes in the presence of noise, from 1.7 to 2.4 on a Likert scale ranging from 1 to 7.

C. Discussion

Experiment 2 added the challenges of masking noise and lower-amplitude signals to the signature whistle identification task. Ratio of signal-to-noise played a major role in participants' accuracies, with performance increasing with higher ratios of signal-to-noise (Figs. 4 and 5). Performance neared chance under conditions with low signal and high noise. Though dolphin performance in this type of task

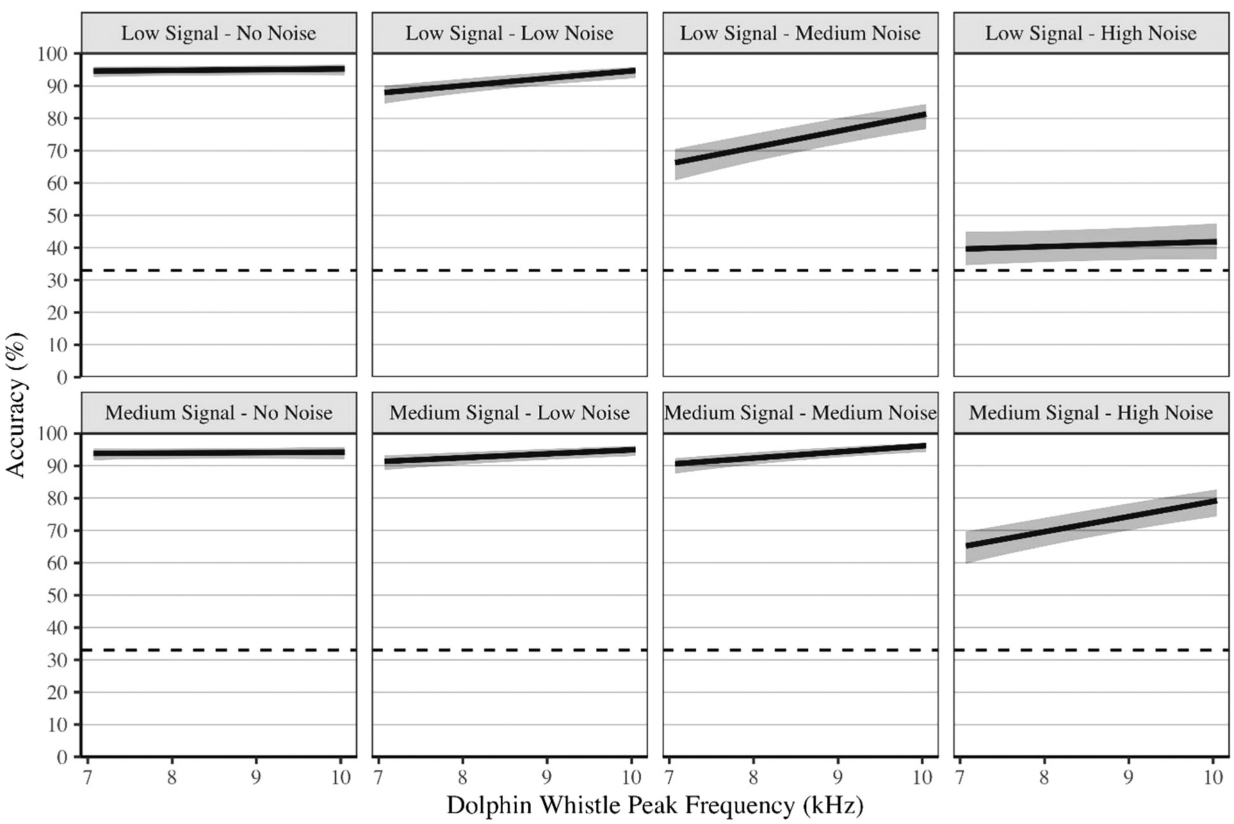

FIG. 5. Relationship between participants' performance accuracy and dolphin whistle frequency at different levels of signal and noise in Experiment 2. Chance performance had 33% accuracy. Gray bands denote 95% confidence regions.

might vary from that of humans, these results provide evidence that low ratios of signal-to-noise impaired signature whistle identification, which is also likely to occur for dolphins. It is also worth noting that human participants in this experiment knew that a signature whistle would be played in each trial, and were actively listening for one whenever a sound clip was played. In the ocean, dolphins may be more likely to miss a whistle outright under some of these challenging conditions, thus losing even the opportunity to successfully identify it.

The underlying cause for different performance with different boat noise samples may be explained by the proximity in frequency between the noise and the signal, as hypothesized. There were significant interaction effects between boat noise peak frequency and signal-to-noise ratio (Fig. 4), whistle frequency and signal-to-noise ratio (Fig. 5), and boat noise peak frequency and whistle peak frequency. *Post hoc* analyses revealed that, in most conditions, performance decreased as boat noise peak frequency increased and as whistle frequency decreased. Given that all of the dolphin whistles were higher in frequency than boat noises, this would indicate that boat noises of more similar frequencies to the whistles would be more negatively impactful on performance. The increase in performance relative to the increase in boat noise frequency in the medium signal–low noise condition was small enough that this result could have resulted from type I error or from other characteristics of the boat noises used, rather than frequency. There was also no relationship between whistle frequency and performance in the low signal–high noise condition; performance in that condition was near chance at all whistle frequencies tested. The analysis of the interaction between boat noise peak frequency and whistle peak frequency confirmed that higher-frequency boat noises paired with lower-frequency whistles both led to poorer performance. In addition, the rate at which increasing boat noise frequency quelled performance was greater for lower-frequency whistles. This supports the hypothesis that proximity in frequency between the whistles and the boat noises would impair whistle identification. Though it was impossible to test for a three-way interaction due to the low amount of variance in certain conditions, it is likely that the interaction between boat noise frequency and whistle frequency was dependent upon signal-to-noise ratio, which would constitute a three-way interaction. In keeping with the existing literature on masking noise and auditory perception (e.g., Egan and Hake, 1950), lower signal-to-noise ratios led to greater difficulty and poorer performance. There appeared to be variation in how impactful different boat noise samples were on whistle identification. Qualitatively, noises that were more intermittent yielded higher performance in more challenging conditions. The present study was conducted on humans, but follows patterns of masking seen across different animal taxa, including fish (Vasconcelos *et al.*, 2007), cetaceans (e.g., Branstetter *et al.*, 2013b), and humans (e.g., Egan and Hake, 1950).

In the interview, participants almost universally cited frequency contour as the most important cue for identifying signature whistles. Participants reported that similarity of frequency contour between two dolphins was the most common cause of mistaken whistle identity. Participants reportedly used frequency contour almost every trial, and relied nearly as heavily on duration. Participants used overall frequency and timbre about the same: slightly less than half of the time. Finally, participants reported that they never or almost never used amplitude or other cues. The lack of usefulness of amplitude might be deceptive, however, as the artificial manipulation of signal amplitude might have left only change in amplitude over time as an available cue. It is unlikely that this would be different *in situ*, however, given that perceived amplitude will vary with distance as well as other factors. In this study, most participants seemed to agree that the frequency contours of Dolphin B and C's signature whistles were qualitatively similar, and those were the dolphins they most often confused. Overall, these reports indicate that frequency contour is consistently important for whistle identification, in accordance with previous research (e.g., Branstetter *et al.*, 2016a; Branstetter *et al.*, 2016b).

IV. GENERAL DISCUSSION

This study is consistent with previous studies that showed humans can discriminate visually among signature whistles (e.g., Sayigh *et al.*, 2007) and auditorily among whistle-like sounds (Branstetter *et al.*, 2016b). Furthermore, this study is consistent with previous studies that suggest signature whistles are largely defined by, and therefore identified using, frequency contour (Harley, 2008; Janik *et al.*, 1994; Janik *et al.*, 2013; Janik *et al.*, 2006; Sayigh *et al.*, 2007). Previous research has shown that humans and dolphins can discriminate among whistle-like stimuli even when frequency contour was designed to be the only differentiating cue among stimuli (Branstetter *et al.*, 2016b), and that frequency contour is critical in visual identification of signature whistles using spectrograms (e.g., Janik, 1999; Janik and Slater, 1998; Kershenbaum *et al.*, 2013; Sayigh *et al.*, 2007; Watwood *et al.*, 2004). In the present study, frequency contour was used most of all auditory cues, and participants reported that their mistakes largely were the result of instances when frequency contour was rendered less useful either due to the timing and severity of masking noise, or due to different whistles having similar frequency contours.

In this study, humans were tested in a laboratory setting to model possible trends in dolphin performance, and to give insight into possible auditory cues used by dolphins. Human and dolphin hearing differ in that dolphins are able to detect sounds of higher frequencies than humans (Herman and Arbeit, 1972; Thompson and Herman, 1975; Wier *et al.*, 1977). However, human and dolphin hearing have many similarities, including a similar threshold of amplitude discrimination (Au, 1993; Green, 1993) and similar frequency discrimination abilities for tonal stimuli (Herman and Arbeit, 1972; Thompson and Herman, 1975; Wier *et al.*, 1977), such as the stimuli in the current study. Dolphins and humans also have similar auditory filter systems and critical ratios within the majority of the human hearing range, including the range of the stimuli used in this study (Lemonds *et al.*, 2011; Lemonds *et al.*, 2012; Johnson, 1968). Though dolphins and humans both experience sound and masking noise frequently, dolphins should be more accustomed to hearing signature whistles than our human participants and may pick up more readily on salient features. The ability to locate the origin of sounds and noise has been shown to contribute to many species' abilities to separate target signals from sources of masking noise, thus allowing or improving perception of the signal (e.g., Arbogast *et al.*, 2002; Bee, 2007; Bee and Riemersma, 2008; Darwin, 1997; Holt and Schusterman, 2007; Kidd *et al.*, 2016; Popov *et al.*, 2020). The absence of localization cues in this study's stimuli may have increased the task's difficulty. In order to compare human performance with that of dolphins in a future study, this factor should be considered. Another aspect of the experiment to consider is the use of self-report, which can be subject to biases, demand characteristics, and other issues (e.g., McCambridge *et al.*, 2012; Speltz and Bernstein, 1976). We attempted to mitigate this by clearly defining all vocabulary and ordering interview questions to reduce order-driven biases, but we are aware people may not always consistently report the cues they are using for discrimination.

This study used a sample of human subjects, who were much more accessible and inexpensive to test than dolphins. Testing dolphins residing in aquaria, zoos, or research facilities is difficult due to the limited number of dolphins able to perform long-term research (and the willingness of zoos or aquaria to undertake such research). Training dolphins for new behavioral research procedures can take months to years, and collecting experimental data can take several more months or years. Because humans are faster and less expensive to test than dolphins, and can be tested in greater numbers, humans may be used as a model organism to efficiently test hypotheses, which can then be verified using dolphin subjects. In addition, human participants can give feedback about which auditory cues they use, allowing follow-up studies with dolphins to begin with clear hypotheses about which cues dolphins, themselves, may be using. To follow up on the current study, a version of this methodology should be repeated with trained dolphins in managed care. For dolphins, all of the variables from this current study could be included, as well as the new variable of familiar versus unfamiliar signature whistles. For example, a mother dolphin might be more capable of identifying her own signature whistle or that of her calf than an unfamiliar dolphin due to familiarity and practice, both with and without masking noise. Another follow-up study could be performed by restricting the available cues to see whether subtracting a cue has an impact on performance. For example, in order to test the usefulness of duration, signature whistles for all the samples would have the same duration. In these future studies, researchers should consider that dolphins may experience some degree of perceptual restoration when listening to masked dolphin vocalizations, much like humans experience phonemic restoration when listening to masked human speech (Kashino, 2006; Warren, 1984; Warren and Obusek, 1971). Prior research has found that some non-human species experience this sensory illusion and others do not (e.g., Petkov *et al.*, 2003; Seeba and Klump, 2009; Seeba *et al.*, 2010; Sugita, 1997), and similar research in dolphins could improve our understanding of how noise affects their perception of masked vocalizations.

This study used a limited number of boat noise and dolphin whistle samples, which are unlikely to represent the diverse range of dolphin vocalizations and anthropogenic noises present in the marine soundscape. The use of specifically signature whistles may limit the extent to which these findings generalize to wild dolphins, which will often need to discriminate among a wider range of whistle types. In order to expand the practical applications of this research, future studies should investigate an even wider variety of noises, including marine construction, sonar from watercraft, shipping noise, and environmental noises such as wind. These studies can also test a wider variety of signals, potentially comparing other types of vocalizations (such as burst pulses or whistles that are not signature whistles), or vocalizations from other species.

The current study provides evidence that anthropogenic masking noise impairs human perception of dolphin signature whistles. If these findings generalize to dolphins, then even slight disruptions of dolphin communication could have population-scale effects on fitness. This might be a cause for concern in regions where anthropogenic noise and dolphins are both present. This study indicates that boat noise may be more likely to negatively impact acoustic communication when the signal and noise are of more proximal frequencies. This principle could be used to generate hypotheses about which species of cetaceans are more likely to be hurt by the masking noises present in their regions. Based on these findings, it seems likely that boat noise would disturb the natural social behaviors of dolphins, which would clearly be in violation of the Marine Mammal Protection Act (1972). Conservationists and policymakers must determine whether this constitutes a significant threat to marine organisms and, if so, create clear guidelines regarding anthropogenic noise. One measure might be for boat and ship designers to prioritize making quieter vessels and machinery. A potentially beneficial direction for future research would be to determine at what distance anthropogenic noise from any individual boat is likely to disturb the behavior of marine mammals. Current guidelines for boat distance are based on risk of injury to the animals, not based on other impacts of boat noise (e.g., Marine Mammal Protection Act, 1972). Such research could help regulate the distance vessels must stay from marine mammals based on the noise they produce.

In conclusion, this was the first study to quantitatively assess how well human listeners can auditorily discriminate among recorded signature whistles from bottlenose dolphins with and without masking by boat noise. Performance was near perfect without masking, and when whistles were masked, performance was impacted by signal-to-noise ratio, and proximity of the whistle and noise frequencies. Listeners primarily relied upon frequency contour when identifying whistles with and without boat noise. Human participants are a valuable model organism through which to investigate perceptual processes in dolphins. Human participants are abundant and inexpensive, can be tested quickly, and can verbally answer questions about which cues they utilized in the task. Results of these studies with human listeners can be used to generate hypotheses for studies done with dolphins. These comparative studies, when conducted on both humans and dolphins using the same stimuli and similar procedures, may expedite and augment perceptual research and may provide additional data to improve conservation efforts for dolphins.

ACKNOWLEDGMENTS

The authors would like to thank the faculty at RIT who generously offered their time and knowledge toward the successful completion of this project. Thank you to Dr. Kirsten Condry for your many insights, especially in the study's methodological design and pilot testing. Thank you to Dr. Stephanie Godleski for your instrumental advice when it came to developing effective means by which to address the study's hypotheses, both methodologically and mathematically. This research was supported by the RIT College of Liberal Arts Student Research Fund and the RIT Department of Psychology. This study was approved by the RIT Institutional Review Board. The authors would also like to thank Dr. Heidi Harley and Moriah Deimeke (New College of Florida) for the dolphin signature whistle recordings that served as the stimuli. Thanks also to Dr. Max Kaplan, Dr. T. Aran Mooney, and Jason Dinh (Woods Hole Oceanographic Institution) for providing samples of boat noise. Kaplan and colleagues' work was supported by a grant to A. Mooney from the National Science Foundation's Division of Ocean Sciences (OCE Grant No. 1536782).

Albuquerque, N. d. S., and Souto, A. (**2013**). "Motorboat noise can potentially mask the whistle sound of estuarine dolphins (*Sotalia guianensis*)," Ethnobiol. Conserv. **2**, 1–15.

Arbogast, T. L., Mason, C. R., and Kidd, G. (**2002**). "The effect of spatial separation on informational and energetic masking of speech," J. Acoust. Soc. Am. **112**(5), 2086–2098.

Au, W. W. L. (**1993**). *The Sonar of Dolphins* (Springer-Verlag, Berlin).

Au, W. W. L., and Martin, D. W. (**1989**). "Insights into dolphin sonar discrimination capabilities from human listening experiments," J. Acoust. Soc. Am. **86**(5), 1662–1670.

Au, W. W. L., Moore, P. W. B., and Pawloski, D. A. (**1988**). "Detection of complex echoes in noise by an echolocating dolphin," J. Acoust. Soc. Am. **83**(2), 662–668.

Au, W. W. L., and Penner, R. H. (**1981**). "Target detection in noise by echolocating Atlantic bottlenose dolphins," J. Acoust. Soc. Am. **70**(3), 687–693.

Bee, M. A. (**2007**). "Sound source segregation in grey treefrogs: Spatial release from masking by the sound of a chorus," Anim. Behav. **74**, 549–558.

Bee, M. A., and Riemersma, K. K. (**2008**). "Does common spatial origin promote the auditory grouping of temporally separated signal elements in grey treefrogs?," Anim. Behav. **76**(3), 831–843.

Branstetter, B. K., Bakhtiari, K., Black, A., Trickey, J. S., Finneran, J. J., and Aihara, H. (**2016a**). "Energetic and informational masking of complex sounds by a bottlenose dolphin (*Tursiops truncatus*)," J. Acoust. Soc. Am. **140**(3), 1904–1917.

Branstetter, B. K., DeLong, C. M., Dziedzic, B., Black, A., and Bakhtiari, K. (**2016b**). "Recognition of frequency modulated whistle-like sounds by a bottlenose dolphin (*Tursiops truncatus*) and humans with transformations in amplitude, duration and frequency," PloS One **11**(2), e0147512.

Branstetter, B. K., Trickey, J. S., Aihara, H., Finneran, J. J., and Liberman, T. R. (**2013a**). "Time and frequency metrics related to auditory masking of a 10 kHz tone in bottlenose dolphins (*Tursiops truncatus*)," J. Acoust. Soc. Am. **134**(6), 4556–4565.

Branstetter, B. K., Trickey, J. S., Bakhtiari, K., Black, A., Aihara, H., and Finneran, J. J. (**2013b**). "Auditory masking patterns in bottlenose dolphins (*Tursiops truncatus*) with natural, anthropogenic, and synthesized noise," J. Acoust. Soc. Am. **133**(3), 1811–1818.

Buckstaff, K. C. (**2004**). "Effects of watercraft noise on the acoustic behavior of bottlenose dolphins, *Tursiops truncatus*, in Sarasota Bay, Florida," Mar. Mammal Sci. **20**(4), 709–725.

Caldwell, M., and Caldwell, D. (**1965**). "Individualized whistle contours in bottle-nosed dolphins (*Tursiops trancatus*)," Nature **207**, 434–435.

Cook, M. L. H., Sayigh, L. S., Blum, J. E., and Wells, R. S. (**2004**). "Signature-whistle production in undisturbed free-ranging bottlenose dolphins (*Tursiops truncatus*)," Proc. Royal Soc. B **271**, 1043–1049.

Darwin, C. J. (**1997**). "Auditory grouping," Trends Cog. Sci. **1**(9), 327–333.

David, J. A. (**2006**). "Likely sensitivity of bottlenose dolphins to pile-driving noise," Water Environ. J. **20**(1), 48–54.

DeLong, C. M. (**2017**). "Human listening experiments provide insight into cetacean auditory perception," Proc. Mtgs. Acoust. **29**(1), 010001.

DeLong, C. M., Au, W. W. L., Harley, H. E., Roitblat, H. L., and Pytka, L. (**2007a**). "Human listeners provide insights into echo features used by dolphins (*Tursiops truncatus*) to discriminate among objects," J. Comp. Psychol. **121**(3), 306–319.

DeLong, C. M., Au, W. W. L., and Stamper, S. A. (**2007b**). "Echo features used by human listeners to discriminate among objects that vary in material or wall thickness: Implications for echolocating dolphins," J. Acoust. Soc. Am. **121**(1), 605–617.

DeLong, C. M., Heberle, A. L., Wisniewski, M. G., and Mercado, E. (**2014**). "The ability to recognize objects from bottlenose dolphin (*Tursiops truncatus*) echoes generalizes across multiple orientations in humans and neural networks," Anim. Cogn. **17**(3), 543–557.

Egan, J. P., and Hake, H. W. (**1950**). "On the masking pattern of a simple auditory stimulus," J. Acoust. Soc. Am. **22**(5), 622–630.

Fish, J. F., Johnson, C. S., and Ljungblad, D. K. (**1976**). "Sonar target discrimination by instrumented human divers," J. Acoust. Soc. Am. **59**(3), 602–606.

Gorman, R. P., and Sawatari, T. (**1985**). "The use of multidimensional perceptual models in the selection of sonar echo features," J. Acoust. Soc. Am. **77**(3), 1178–1184.

Green, D. M. (**1960**). "Psychoacoustics and detection theory," J. Acoust. Soc. Am. **32**(10), 1189–1203.

Green, D. M. (**1993**). "Auditory intensity discrimination," in *Human Psychophysics*, edited by W. A. Yost, A. N. Popper, and R. R. Fay (Springer, New York), pp. 13–55.

Gridley, T., Cockcroft, V., Hawkins, E. R., Lemon Blewitt, M., Morisaka, T., and Janik, V. M. (**2014**). "Signature whistles in free ranging populations of Indo-Pacific bottlenose dolphins, *Tursiops aduncus*," Mar. Mammal Sci. **30**(2), 512–527.

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," Behav. Process. **77**(2), 243–268.

Helweg, D. A., Roitblat, H. L., Nachtigall, P. E., Au, W. W. L., and Irwin, R. J. (**1995**). "Discrimination of echoes from aspect-dependent targets by a bottlenose dolphin and human listeners," in *Sensory Systems of Aquatic Mammals*, edited by R. A. Kastelein, J. A. Thomas, and P. E. Nachtigall (De Spil Publishers, Woerden), pp. 129–136.

Herman, L. M., and Arbeit, W. R. (**1972**). "Frequency difference limens in the bottlenose dolphin: 1–70 kc/s," J. Aud. Res. **2**, 109–120.

Herzing, D. L. (**1996**). "Vocalizations and associated underwater behavior of free-ranging Atlantic spotted dolphins, *Stenella frontalis* and bottlenose dolphins, *Tursiops truncatus*," Aquat. Mamm. **22**, 61–80.

Holm, S. (**1979**). "A simple sequentially rejective multiple test procedure," Scand. J. Stat. **6**(2), 65–70.
Holt, M. M., and Schusterman, R. J. (**2007**). "Spatial release from masking of aerial tones in pinnipeds," J. Acoust. Soc. Am. **121**(2), 1219–1225.
Janik, V. M. (**1999**). "Pitfalls in the categorization of behaviour: A comparison of dolphin whistle classification methods," Anim. Behav. **57**(1), 133–143.
Janik, V. M., Dehnhardt, G., and Todt, D. (**1994**). "Signature whistle variations in a bottlenosed dolphin, *Tursiops truncatus*," Behav. Ecol. Sociobiol. **35**, 243–248.
Janik, V. M., King, S. L., Sayigh, L. S., and Wells, R. S. (**2013**). "Identifying signature whistles from recordings of groups of unrestrained bottlenose dolphins (Tursiops truncatus)," Mar. Mammal Sci. **29**(1), 109–122.
Janik, V. M., and Sayigh, L. S. (**2013**). "Communication in bottlenose dolphins: 50 years of signature whistle research," J. Compar. Physl. A **199**, 479–489.
Janik, V. M., Sayigh, L. S., and Wells, R. S. (**2006**). "Signature whistle shape conveys identity information to bottlenose dolphins," P. Natl. A. Sci. USA **103**(21), 8293–8297.
Janik, V. M., and Slater, P. J. (**1998**). "Context-specific use suggests that bottlenose dolphin signature whistles are cohesion calls," Anim. Behav. **56**(4), 829–838.
Johnson, C. S. (**1968**). "Masked tonal thresholds in the bottlenosed porpoise," J. Acoust. Soc. Am. **44**(4), 965–967.
Kaplan, M. B., and Mooney, T. A. (**2015**). "Ambient noise and temporal patterns of boat activity in the US Virgin Islands National Park," Mar. Pollut. Bull. **98**(1), 221–228.
Kashino, M. (**2006**). "Phonemic restoration: The brain creates missing speech sounds," Acoust. Sci Technol. **27**(6), 318–321.
Kershenbaum, A., Sayigh, L. S., and Janik, V. M. (**2013**). "The encoding of individual identity in dolphin signature whistles: How much information is needed?," PloS One **8**(10), e77671.
Kidd, G., Mason, C. R., Swaminathan, J., Roverud, E., Clayton, K. K., and Best, V. (**2016**). "Determining the energetic and informational components of speech-on-speech masking," J. Acoust. Soc. Am. **140**(1), 132–144.
King, S. L., and Janik, V. M. (**2013**). "Bottlenose dolphins can use learned vocal labels to address each other," P. Natl. A. Sci. USA **110**(32), 13216–13221.
Lemonds, D. W., Au, W. W. L., Vlachos, S. A., and Nachtigall, P. E. (**2012**). "High-frequency auditory filter shape for the Atlantic bottlenose dolphin," J. Acoust. Soc. Am. **132**(2), 1222–1228.
Lemonds, D. W., Kloepper, L. N., Nachtigal. P. E., Au, W. W. L., Vlachos, S. A., and Branstetter, B. K. (**2011**). "A re-evaluation of auditory filter shape in delphinid odontocetes: Evidence of constant-bandwidth filters," J. Acoust. Soc. Am. **130**(5), 3107–3114.
Lusseau, D. (**2003**). "Effects of tour boats on the behavior of bottlenose dolphins: Using Markov chains to model anthropogenic impacts," Conserv. Biol. **17**(6), 1785–1793.
Marshall, L., and Jesteadt, W. (**1986**). "Comparison of pure-tone audibility thresholds obtained with audiological and two-interval forced-choice procedures," J. Speech Lang. Hear. R. **29**(1), 82–91.
Marine Mammal Protection Act, 16 USC §§ 1361 *et seq*. (**1972**).
McCambridge, J., de Bruin, M., and Witton, J. (**2012**). "The effects of demand characteristics on research participant behaviours in non-laboratory settings: A systematic review," PloS One **7**(6), e39116.
Mercado, E. III., and DeLong, C. M. (**2010**). "Dolphin cognition: Representations and processes in perception and memory," Int. J. Comp. Psychol. **23**, 344–378.
Muller-Gass, A., Marcoux, A., Logan, J., and Campbell, K. B. (**2001**). "The intensity of masking noise affects the mismatch negativity to speech sounds in human subjects," Neurosci. Lett. **299**(3), 197–200.
Nelken, I., Rotman, Y., and Yosef, O. B. (**1999**). "Responses of auditory-cortex neurons to structural features of natural sounds," Nature **397**, 154–157.
Petkov, C. I., O'Connor, K. N., and Sutter, M. L. (**2003**). "Illusory sound perception in macaque monkeys," J. Neurosci. **23**(27), 9155–9161.
Pirotta, E., Merchant, N. D., Thompson, P. M., Barton, T. R., and Lusseau, D. (**2015**). "Quantifying the effect of boat disturbance on bottlenose dolphin foraging activity," Biol. Conserv. **181**, 82–89.
Popov, V. V., Supin, A. Ya., Gvozdeva, A. P., Nechaev, D. I., Tarakanov, M. B., and Sysueva, E. V. (**2020**). "Spatial release from masking in a bottlenose dolphin *Tursiops truncatus*," J. Acoust. Soc. Am. **147**(3), 1719–1726.
Quick, N. J., and Janik, V. M. (**2008**). "Whistle rates of wild bottlenose dolphins (*Tursiops truncatus*): Influences of group size and behavior," J. Comp. Psychol. **122**(3), 305–311.
Quick, N. J., and Janik, V. M. (**2012**). "Bottlenose dolphins exchange signature whistles when meeting at sea," P. Roy. Soc. B-Biol. Sci. **279**(1738), 2539–2545.
Raudenbush, S. W., and Bryk, A. S. (**2002**). *Hierarchical Linear Models: Applications and Data Analysis Methods* (Sage, California).
Ross, D. (**2005**). "Ship sources of ambient noise," IEEE J. Oceanic Eng. **30**(2), 257–261.
Sayigh, L. S., Esch, H. C., Wells, R. S., and Janik, V. M. (**2007**). "Facts about signature whistles of bottlenose dolphins, *Tursiops truncatus*," Anim. Behav. **74**(6), 1631–1642.
Sayigh, L. S., and Janik, V. M. (**2010**). "Dolphin signature whistles," in *Encyclopedia of Animal Behavior*, edited by M. D. Breed and J. Moore (Academic Press, Oxford), pp. 553–561.
Sayigh, L. S., Tyack, P. L., Wells, R. S., and Scott, M. D. (**1990**). "Signature whistles of free-ranging bottlenose dolphins, *Tursiops truncatus*: Stability and mother-offspring comparisons," Behav. Ecol. Sociobiol. **26**(4), 247–260.
Sayigh, L. S., Tyack, P. L., Wells, R. S., Solow, A. R., Scott, M. D., and Irvine, A. B. (**1999**). "Individual recognition in wild bottlenose dolphins: A field test using playback experiments," Anim. Behav. **57**(1), 41–50.
Sayigh, L. S., Wells, R. S., and Janik, V. M. (**2017**). "What's in a voice? Dolphins do not use voice cues for individual recognition," Anim. Cogn. **20**(6), 1067–1079.
Seeba, F., and Klump, G. M. (**2009**). "Stimulus familiarity affects perceptual restoration in the European starling (Sturnus vulgaris)," PloS One **4**(6), e5974.
Seeba, F., Schwartz, J. J., and Bee, M. A. (**2010**). "Testing an auditory illusion in frogs: Perceptual restoration or sensory bias?," Anim. Behav. **79**, 1317–1328.
Shah, N. J., Jäncke, L., Grosse-Ruyken, M. L., and Müller-Gärtner, H. W. (**1999**). "Influence of Acoustic masking noise in fMRI of the auditory cortex during phonetic discrimination," J. Magn. Reson. Imaging **9**(1), 19–25.
Speltz, M. L., and Bernstein, D. A. (**1976**). "Sex differences in fearfulness: Verbal report, overt avoidance and demand characteristics," J. Behav. Ther. Exp. Psy. **7**(2), 117–122.
Sugita, Y. (**1997**). "Neuronal correlates of auditory induction in the cat cortex," Neuroreport **8**(5), 1155–1159.
Thompson, R. K. R., and Herman, L. M. (**1975**). "Underwater frequency discrimination in the bottlenosed dolphin (1–140 kHz) and the human (1–8 kHz)," J. Acoust. Soc. Am. **57**(4), 943–948.
Tukey, J. (**1949**). "Comparing individual means in the analysis of variance," Biometrics **5**(2), 99–114.
Vasconcelos, R. O., Amorim, M. C. P., and Ladich, F. (**2007**). "Effects of ship noise on the detectability of communication signals in the Lusitanian toadfish," J. Exp. Biol. **210**(12), 2104–2112.
Warren, R. M. (**1984**). "Perceptual restoration of obliterated sounds," Psychol. Rev. **96**, 2016–2022.
Warren, R. M., and Obusek, C. J. (**1971**). "Speech perception and phonemic restoration," Percept. Psychophys. **9**, 358–362.
Watwood, S. L., Tyack, P. L., and Wells, R. S. (**2004**). "Whistle sharing in paired male bottlenose dolphins, *Tursiops truncatus*," Behav. Ecol. Sociobiol. **55**(6), 531–543.
Wier, C. C., Jesteadt, W., and Green, D. M. (**1977**). "Frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **61**(1), 178–184.
Würsig, B., and Richardson, W. J. (**2009**). "Noise, Effects of," in *Encyclopedia of Marine Mammals*, 2nd ed. (Academic Press, London), pp. 765–773.