

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

U.S. Navy Research

U.S. Department of Defense

2013

Exposure amplitude and repetition affect bottlenose dolphin behavioral responses to simulated mid-frequency sonar signals

Dorian S. Houser

National Marine Mammal Foundation, dorian.houser@nmmfoundation.org

Steve W. Martin

U.S. Navy Marine Mammal Program

James J. Finneran

U.S. Navy Marine Mammal Program

Follow this and additional works at: <https://digitalcommons.unl.edu/usnavyresearch>

Houser, Dorian S.; Martin, Steve W.; and Finneran, James J., "Exposure amplitude and repetition affect bottlenose dolphin behavioral responses to simulated mid-frequency sonar signals" (2013). *U.S. Navy Research*. 36.

<https://digitalcommons.unl.edu/usnavyresearch/36>

This Article is brought to you for free and open access by the U.S. Department of Defense at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in U.S. Navy Research by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



Exposure amplitude and repetition affect bottlenose dolphin behavioral responses to simulated mid-frequency sonar signals

Dorian S. Houser^{a,*}, Steve W. Martin^b, James J. Finneran^b

^a National Marine Mammal Foundation, 2240 Shelter Island Drive, Suite 200, San Diego, CA 92071, USA

^b U.S. Navy Marine Mammal Program, Space and Naval Warfare Systems Center Pacific, 53560 Hull Street, Code 71510, San Diego, CA 92152, USA

ARTICLE INFO

Article history:

Received 19 November 2012

Received in revised form 25 February 2013

Accepted 25 February 2013

Available online 25 March 2013

Keywords:

Acoustic exposure

Behavioral response

Bottlenose dolphin

Dose–response

Sound

ABSTRACT

Thirty bottlenose dolphins (*Tursiops truncatus*) participated in a controlled exposure study using an acoustic stimulus that consisted of a simulated tactical sonar signal (1-s duration, 3250–3450 Hz). Each dolphin was trained a behavior requiring it to swim across an enclosure, touch a paddle, and return to the starting location. Ten-trial control and exposure sessions were performed with each dolphin. Dolphins were placed into one of five groups, each of which received a 115, 130, 145, 160, 175 or 185 dB re 1 μ Pa (rms) sound pressure level (SPL) exposure while crossing the enclosure on exposure trials. A canonical correlation analysis was used to determine the set of behavioral responses most interrelated with the independent variables of exposure level, trial number, and age. Responses that significantly contributed to the canonical model were used to create dose–response functions based on the received SPL. Dose–response functions demonstrated a robust relationship between received SPL and the probability of response and indicated rapid habituation to repetitive exposures with received SPL \leq 160 dB. No habituation was observed at received SPL \geq 175 dB re 1 μ Pa and all dolphins refused to participate in trials when the received SPL = 185 dB re 1 μ Pa. Although bottlenose dolphins may rapidly habituate to sound exposures below a certain level, particularly if there is food motivation, abandonment of behaviors increases rapidly at received levels \geq 175 dB re 1 μ Pa.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

The overlap between the acoustic ecology of marine mammals and the intentional or unintentional introduction of anthropogenic noise in the ocean environment potentially impacts individuals and populations of marine mammals over largely varying temporal and spatial scales (Tyack, 2008). The consequences of anthropogenic sound exposure are most notable in dramatic changes that result from either the direct effect of sound on marine mammal tissues or through alterations in dive behavior that result in physical harm or stranding (Fernández et al., 2005; Jepson et al., 2003). Through masking of biological signals or disruption of normal behaviors, consequences might also be realized in missed breeding opportunities, reduced energy acquisition, habitat abandonment, and disruption of social organization (e.g. mother–calf separation). Determining how changes in behavior might produce negative consequences to an animal is a challenging task complicated by a number of factors. Among these are the disposition of the exposed animal (including species-wide sensitivities), its prior experience with the sound, the proximity of the sound source, the characteristics of the sound (e.g. level, frequency content, duration and duty cycle), and other behavioral contexts (e.g. breeding displays or foraging opportunities).

Protecting marine mammals from the potential impact of anthropogenic sound exposure is difficult because little is known about the consequences of sound exposure to marine mammals and evidence suggests considerable variation in individual and species tolerances and reactions to introduced sound (Diaz López and Mariño, 2011; Finneran et al., 2003; Holst et al., 2011; Holt et al., 2011; Johnston, 2002; Kastelein et al., 2012; McCarthy et al., 2011; Miksis-Olds and Wagner, 2011; Miksis-Olds et al., 2007; Miller et al., 2009; Morton and Symonds, 2002; Niu et al., 2012; Nowacek et al., 2004; Olesiuk et al., 2002; Parks et al., 2007; Tyack et al., 2011). Ultimately, it is the mapping of changes in natural behaviors as a result of anthropogenic sound exposure to potential fitness consequences that is of interest (National Research Council (NRC), 2005). The ability to achieve this goal is hindered at the outset by an inability to accurately relate the numerous types of possible sound exposures to repeatable and predictable changes in the suite of behaviors exhibited across species, populations, and individuals. The development of dose–response functions has been recommended as one means by which the relationship between behavioral variability and sound exposure might be explored (Southall et al., 2007).

Dose–response functions generally relate the dose of a chemical agent or sensory stimulus to a pre-defined end state. The approach is widely used in the pharmaceutical industry and in toxicological studies to determine the relationship between chemical exposure and physiological impacts, behavioral alterations, or death. Although a recommended approach for improving the assessment of the impact of sound on wildlife

* Corresponding author. Tel.: +1 877 360 5527x112; fax: +1 877 773 3153.

E-mail address: dorian.houser@nmmfoundation.org (D.S. Houser).

(Pater et al., 2009; Tyack, 2009), relatively few dose–response studies relating sound exposure to behavioral responses of wild species have been performed. Several studies of the impact of noise from military and industrial activities have been conducted in bird species of concern (Delaney et al., 1999, 2011; Goudie and Jones, 2004). Similarly, dose–response studies of noise exposure in laboratory animals are relatively rare (Wöhr et al., 2005). Little information on dose–response relationships between sound exposure and marine mammal behavior exists. Nevertheless, because of the conceptual sensibility underlying the dose–response relationship, the application of dose–response functions in predicting behavioral impacts to sound-exposed marine mammals has established itself in the realm of marine mammal environmental compliance (Department of the Navy (DoN), 2008a,b,c).

There are few observations of marine mammal responses to anthropogenic sound that have sufficient information or control over the exposure level or context of the exposure to define a dose–response relationship with species and context specificity. The purpose of this study was to perform a controlled exposure study using a simulated tactical sonar signal in a large sample of bottlenose dolphins. The choice of a tactical sonar signal as a stimulus was based on the current concern for the impact of sonar systems on marine life, particularly marine mammals. The study was designed to impose the same context across exposure conditions and control the exposure levels such that a dose–response relationship between behavioral alterations and the received level of the simulated sonar signal could be determined. The design also permitted the influence of repetitive exposure and age of the animal on the dose–response relationship to be explored. The resultant dose–response functions are relevant to environmental compliance issues and facilitate our understanding of the proximate consequences of anthropogenic sound exposure to marine mammals.

2. Materials and methods

Thirty bottlenose dolphins (*Tursiops truncatus*) participated in the controlled exposure study. All procedures of the controlled exposure study were approved by the Institutional Animal Care and Use Committee of the Biosciences Division, Space and Naval Warfare Systems Center (SSC) Pacific, in San Diego, CA, USA, and the Department of the Navy Bureau of Medicine. The study followed all applicable USA Department of Defense guidelines for the care of laboratory animals.

All dolphins were maintained by the US Navy Marine Mammal Program (MMP) and were maintained in open-water, netted enclosures, 9 × 9 to 12 × 24 m, located within San Diego Bay. Each dolphin received a daily diet of fish and squid required to maintain ideal weights as established by the veterinary staff of the MMP. Water and air temperatures to which the dolphins were exposed were dictated by the natural environment of San Diego Bay. Dolphins ranged in age from 6.5 to 45.5 years of age at the time of the study (Table 1). Due to the gender distribution of dolphins available for the study, there was a bias toward participation by male dolphins (21 males/9 females). Although many of the subjects had previously participated in cooperative psychophysical research tasks, none of the subjects had previously participated in any research related to noise-induced threshold shifts (changes in hearing sensitivity) or had, to our knowledge, been exposed to high-intensity sonar signals.

2.1. Behavioral task and study design

Each bottlenose dolphin was trained to complete a task in which it left one side of an enclosure (station A), traveled to the other side of the enclosure to touch a target paddle (station B) with its rostrum, and then returned to the starting station A. This process was referred to as the ABA behavior. An animal trainer located at station A initiated each trial by giving a visual signal to the dolphin. The dolphin then had 30 s to complete the ABA behavior. A second trainer, located at station B, was responsible for verifying that the dolphin touched the

Table 1

Subject ID, and age and gender distribution of dolphins participating in the controlled exposure study.

Subject ID	Age (years)	Gender
D1	6.6	M
D2	8.5	M
D3	9.5	M
D4	7.5	M
D5	39.5	M
D6	29.5	M
D7	11.5	M
D8	29.5	M
D9	39.5	M
D10	31.5	F
D11	9	F
D12	18.3	M
D13	29.5	M
D14	27.5	M
D15	28.8	M
D16	29.5	M
D17	32.5	F
D18	17.7	M
D19	35.5	F
D20	10	M
D21	21	F
D22	45.5	M
D23	17.7	M
D24	25.5	M
D25	27.5	F
D26	26	F
D27	27.5	F
D28	25.5	F
D29	28.5	M
D30	28.5	M

target paddle. If the dolphin completed the trial and returned to the trainer at station A before the trial was over, it was rewarded with a fixed amount of Icelandic capelin (*Mallotus villosus*) equal to 1% of its daily allotment of capelin. The next trial began immediately after the end of the prior 30-s trial and 10 trials were conducted in sequence to complete a session. Thus, a session, whether control or experimental, had a duration of 5 min. Each dolphin was trained on the task until it performed 10 repetitions of the task (i.e. one session) without error. The dolphin was tested within several days of reaching this criterion.

Control and experimental sessions were conducted on the same day and in the same order. Control sessions were always conducted first and each dolphin was given a short rest of several minutes between the control session and the experimental session. Exposure sessions were performed exactly as the control sessions except that the dolphin was exposed to a playback of a simulated mid-frequency sonar signal as it crossed the midpoint of the enclosure. Playbacks occurred once per trial and only if the dolphin swam past the midpoint of the enclosure; i.e. if the dolphin refused to participate on a trial and did not cross the pen, no playback occurred for that trial. If a dolphin refused to participate on a trial, it was given an opportunity to participate on the following trial through a trainer recall, which consisted of a hand slap of the water surface. The recall was given 25 s into the trial for which the dolphin refused to participate, and then again at 5 and 10 s into the next trial. If the dolphin returned to the trainer before 10 s into the subsequent trial, it was given the cue to perform the task. If the dolphin returned after 10 s into the trial, it was kept on station by the trainer until the next trial.

Stimulus received levels were consistent for each individual (i.e. all trials within a session were at the same received level), but varied from individual to individual in order to provide a range of received levels. The received (rms) sound pressure levels (SPL) were 115, 130, 145, 160, 175, or 185 dB re 1 μ Pa (hereafter denoted as “dB SPL”). Five dolphins were tested at each SPL such that there were six groups of five dolphins each and one group for each received SPL. Prior to

performing the study, each dolphin was assigned an identifier from D1 to D30 so that its age, history, and prior experience could not be identified by anyone except the primary investigator (PI). This was the “study ID.” A co-investigator (CI), who was blind to which dolphin a particular study ID was assigned, randomly assigned an acoustic exposure level to each study ID. On the day of testing for a given dolphin, the PI requested the acoustic exposure level for that dolphin’s study ID from the CI. Thus, only the PI knew the exposure level for the dolphin and the level was not known until the day that the control and experimental sessions took place.

Dolphins were tested as their schedules permitted their participation. However, three dolphins lived at the floating facility where testing occurred. These dolphins were tested first so as to avoid habituation or sensitization to the playbacks. Each of these dolphins was tested while the other two dolphins were working or training in the open ocean away from the test site.

2.2. Experimental setup

The study was performed in a 9.1×18.3 m netted enclosure located within San Diego Bay (Fig. 1). The enclosure was part of a floating pier complex and was physically separated from the facilities where most of the MMP dolphins are housed. Station A, the location at which a dolphin started a trial, was located to the east side of the enclosure pointing into San Diego Bay (Figs. 1 and 2). Station B, was opposite of station A and closest to the shore. A target paddle was mounted to the side of the enclosure at station B and extended just below the water surface. A sound source for generating the acoustic stimulus was located 1 m underwater and 1 m behind the station B target paddle.

A large steel arch, with a maximum height of 6.4 m above the water surface, spanned the middle of the enclosure (Fig. 1). At the apex of the arch, two gigabit ethernet (GigE) cameras (Prosilica Inc., Burnaby, British Columbia) were placed and oriented so they looked down on the enclosure. The cameras were approximately 6.1 m above the water surface.

One camera operated at 120 frames-per-second (fps) and recorded at a 640×480 resolution (Model# GE680C). This camera was fit with a 120° field-of-view (FOV) lens (Model# SY110M; Theia Technologies, Wilsonville, Oregon) with minimal edge (fish-eye) distortion. The other camera (Model# GE 1050C), mounted directly beside the first, operated at 35 fps and recorded at a 1024×1024 resolution. It was fitted with a 125° FOV lens (Model# MY125M; Edmund Optics, Barrington, New Jersey) that also had minimal edge distortion. Video from the cameras was captured utilizing StreamPix (v.4.20) digital video recorder software (Norpix Inc., Montreal, Canada).

Two “bullet” cameras (CVC-320WP) were mounted to the north side of the enclosure and oriented so that they viewed the test enclosure from the middle of the pen to the edge of the pen. One bullet camera was mounted on the overhead arch so that it viewed the side of the enclosure where station B was located. The other bullet camera was mounted to the control hut and oriented so that it viewed the side of the enclosure where station A was located. The bullet cameras were connected to a digital video recorder and a video viewing screen so that sessions could be viewed in real time and recorded for later analysis.

A hydrophone (Reson TC-4013) was placed at a depth of 1 m approximately midway between the sound source and the target paddle at station B, but slightly offset to the side. The hydrophone was used to monitor underwater phonations of the dolphins as well as any other underwater sound present during the sessions. The hydrophone was amplified by 32 dB with a Reson VP1000 (Model# EC6061) prior to being run through a stereo mixer (Ashly MX-206).

Both the trainers and PI, who served as the session controller for acoustic playbacks, wore headset intercoms (XO-1; Telex, Burnsville, Minnesota) in order to communicate with one another during a session. All communications were relayed to a single unit with the output attached to the audio mixer. The monitor hydrophone recordings and intercom audio were mixed with the stereo mixer and the mixer output used as the audio input to the StreamPix recording software and the digital video recorder that recorded inputs from the bullet cameras. The audio sampling rate within the StreamPix software was 48 kHz.



Fig. 1. Picture of the test enclosure with station locations designated by yellow letters. The arch over the enclosure held the overhead video cameras. Dolphins initiated each trial sequence on the middle of the right side of the pen (station A), swam across the pen and touched a paddle on the opposite side (station B), and then returned to the starting location. The sound source was 1 m behind station B and 1 m below the surface. During exposure trials, dolphins were exposed to the simulated sonar signal at the midpoint of the enclosure.

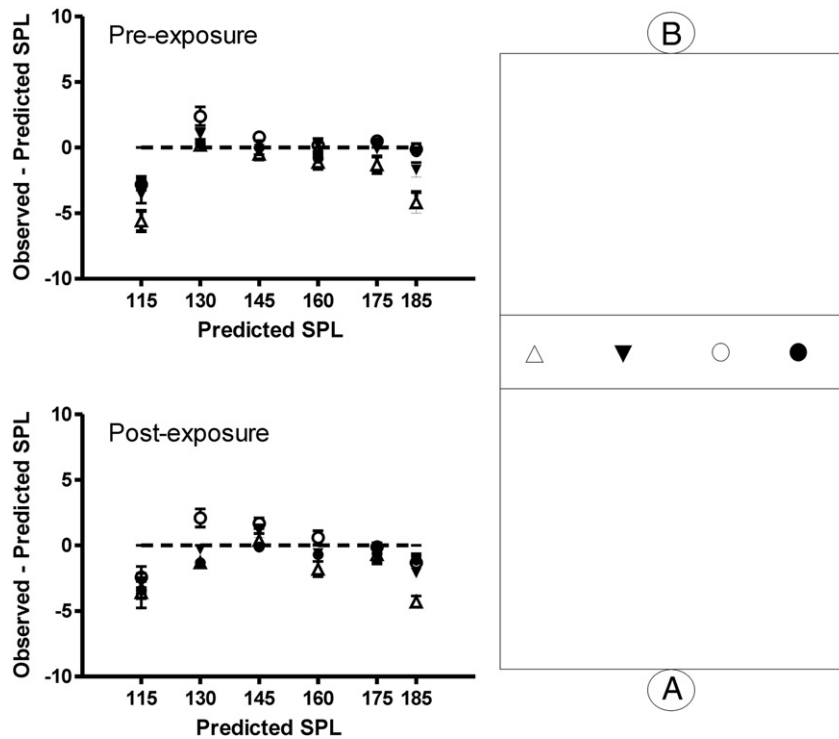


Fig. 2. Locations of the calibration hydrophones relative to stations A and B. The right side of the panel is a simplified, overhead schematic of the test enclosure. The letters A and B correspond to the location of stations A and B as shown in Fig. 1. The hydrophone locations are designated by the small circles and triangles distributed across the center of the enclosure. The smaller rectangle, which bisects the enclosure and contains the symbols for the hydrophones, designates the region where the sound exposure occurred during exposure trials. The figures on the left show the mean (\pm SD) difference between the predicted and measured SPL at each recording location and for each predicted SPL for both pre- and post-exposure calibrations.

2.3. Sound source and acoustic stimulus

Stimuli were generated using a piezoelectric cylinder located approximately 1 m behind station B and placed at 1 m of depth. The stimulus was played back from a laptop computer, amplified (AE Techtron LVC 5050), and input to the sound projector. The maximum source level of the projector was \sim 202 dB SPL, which equated to a maximum received level at the location of the sound exposure of \sim 185 dB SPL. The waveform of the acoustic stimulus was based on that of mid-frequency active sonar waveforms commonly used by the US Navy for anti-submarine warfare. It was not an exact replicate, but contained many of the salient features of the signal waveform. This signal was chosen because of the relevance to broad-scale Navy activities and its putative relationship to prior stranding events (D'Amico et al., 2009; Fernández et al., 2005; Jepson et al., 2003). The exposure waveform consisted of a 0.5-s upward FM sweep (center frequency \sim 3250 Hz) with a 50 ms rise time immediately followed by a 0.5-s CW (\sim 3450 Hz). Most subjects were believed to be naïve to the particular waveform since tactical signals are not generated within the shallows of San Diego Bay. However, it cannot be ruled out that some of the dolphins could hear the signal when played back at its highest level, even though they lived at different parts of San Diego Bay.

Acoustic stimuli were calibrated at the midpoint of the enclosure prior to and following each control and experimental session for a dolphin. Instantaneous sound pressure levels were measured with four hydrophones (Reson TC-4013) stretched across the width of the enclosure at approximately equal increments (Fig. 2). The hydrophone signals were high pass filtered (100 Hz) and amplified (Reson VP1000) prior to being digitized (National Instruments USB-6259). All signals were calibrated from the pressure waveform except the lowest level signals (\sim 115 dB SPL). The received levels of these signals were determined from the frequency domain because the signal was imperceptible in the time domain, i.e. it was indistinguishable from the ambient noise in the time-waveform.

2.4. Behavioral response severity scoring

Prior to performing the controlled exposure study, a list of the potential behavioral reactions expected from the dolphins was created with written descriptions of what each of the reactions would entail (see Supplemental File 1). Several of the responses had a quantitative basis for determination (see 2.5 Session Analysis). Twenty-four marine mammalogists and behavioral biologists were then sent a request to rank the behavioral responses on a severity scale of 1–100. The purpose of the scoring survey was to determine whether there was enough agreement among experts within the field of marine mammalogy over the severity of behavioral responses that could be used to better qualify the impact of a behavioral response. Accompanying the request was the description of the behavioral reaction and the overall study design. Fourteen invitees responded to the invitation and sent their severity scores to the CI. Median and mean severity scores were determined for each of the behavioral reactions and the reactions were then ordered according to the median or mean severity score (see Supplemental File 2). No other participant in the study was permitted to see the severity scores until after the study was completed and the control and experimental sessions had been evaluated for the occurrence of behaviors that could qualify as a behavioral response. It was anticipated that the median or mean scores would be applied to the behavioral responses so that a relative scale of response severity could be determined and compared to the received SPL.

2.5. Session analysis

Session files were post-processed by manually synchronizing the audio and video files digitized by the Streampix software within Adobe Premier Elements (Adobe Systems Inc., San Jose, California). For each trial within a session, a 1-second period of silence was entered into the audio stream at the point during which the dolphin

crossed the middle of the pool (i.e. the point where sound exposure occurred). This was performed for every trial of the control and experimental sessions so the stimulus playback could not be used as a cue for the type of session being evaluated during the analysis. For the highest level exposures, a notch filter which captured the frequencies of the stimulus waveform was run through the entire session audio file in order to remove any audible reverberation. Resultant session audio/video (A/V) files were then randomized and labeled with a number from 1 to 60 so that identifying information regarding the animal being tested or the type of session (control vs. exposure) was removed.

Two dolphin handlers that were unassociated with the research project were selected to review the session A/V files and determine if any behaviors that could be classified as behavioral responses occurred. The analysis was performed for each trial of each session such that a list of behaviors for each trial was created. Behaviors could occur at any point within a 30-s trial window; however, multiple occurrences of the same behavior within a trial were only counted once (i.e. three fluke slaps by the dolphin were simply marked as “fluke slap”). Thus, a session report was produced that listed all of the behavioral responses for each trial in the session with a binary indicator of whether or not the behavior was observed (0 = behavior was not observed, 1 = behavior was observed). After all of the sessions had been scored, the results of the two scorers were compared. Agreement on the observation of behavioral responses was 90%. On the occasions that a response was scored by only one reviewer, it was conservatively assumed that the behavior occurred.

Several potential behavioral responses had a quantitative basis for determining their occurrence. One of these was based on whether the dolphin's movement during the trial was directed or undirected. To make this determination, a movement analysis was performed using Movias Pro motion analysis software (NAC Image Technology, Simi Valley, California). Each dolphin had zinc oxide placed on its head before the control and exposure sessions in order to create a white spot in the video files that could be tracked. The motion analysis software was used to track the zinc oxide in each session and the total distance that the dolphin traveled on each trial was determined. A directivity index was calculated by determining twice the distance between station A and station B and dividing it by the total distance traveled by the dolphin during a trial. A perfectly directional path by a dolphin would therefore have a directivity index of 1.0. The mean and standard deviation of the directivity index were calculated across all trials of the control session. The directivity index for each of the trials, both control and session, were then compared to this value to determine whether variations in directional travel qualified as a behavioral response (see Supplemental File 1). On trials where the animal refused to participate in the trial, the directivity index analysis was not performed.

The remaining behaviors that had a quantitative basis for determination included the time spent submerged during a trial, the completion time of the ABA behavior, and the number of respirations and whistles produced during a trial. The number of whistles produced on each trial was determined by listening to the audio of the underwater recordings and visualizing the whistle contours via joint time-frequency analysis (spectrogram). Spectrograms were created in SoundForge (Sony Creative Software, Middleton, WI) and parameters for the spectrogram analyses were varied in order to best visualize the whistle contours, i.e. the settings were not constant across trials. The trial completion time, time spent submerged, and number of respirations was determined from the session A/V files. For each of these, the mean and standard deviation of the measured value were calculated from all trials within the control session for a given animal. Values for individual trials from both control and exposure sessions were then compared to the mean and standard deviation to determine whether behaviors observed on these trials qualified as a behavioral response (see Supplemental File 1).

Two acoustic metrics were removed from the analysis based on an inability to accurately quantify their occurrence. These included the

number of echolocation clicks produced by the animal as well as the number of buzzes (or burst-pulses). These metrics were not used because the use of a single monitoring hydrophone did not permit reliable and accurate detection of these sound types when the dolphin was oriented away from the hydrophone, particularly when it was located on the opposite side of the pen from the monitoring hydrophone. Therefore, these responses were not included in the analysis and are not discussed further (noted by ## in Supplemental File 1). Two behavioral responses that were not anticipated prior to conducting the study were observed during the analysis phase and these were added to the list of potential behavioral responses (noted by \$\$ in Supplemental File 1). These responses included: “Paddle miss,” where the dolphin swam the full circuit for the trial but swam by the paddle without touching it; and “Source investigation,” where the dolphin stayed at station B following a sound exposure to inspect the sound source either visually or through echolocation.

2.6. Data analysis

The probability of a behavioral response was determined for each exposure level and trial number assuming that any observed response qualified as a behavioral response. Based on the outcome of this initial analysis (see Section 3. Results), it was determined that a model was required to reduce the parameter space to incorporate only those response variables that shared a significant amount of the variance with the independent variables of received SPL, dolphin age, and trial number. Therefore, a canonical correlation analysis was utilized to determine which of the three independent and 35 response variables contributed to the interaction model.

Canonical correlation analysis (CCA) correlates linear relationships between multidimensional variables that are composites of other metric-independent variables (Hair et al., 1998). A series of functions that maximize the correlation between the composite variables, or canonical variates, can be used to determine the overall strength of the relationship between the canonical variates. By measuring the relative contribution of the independent and dependent variables to these canonical functions, a better understanding of the nature of the relationship between the independent and dependent variables can be obtained. The number of canonical variates formed is limited by the smaller of the number of variables that contribute to a variate. The first pair of derived canonical variates has the highest degree of intercorrelation between the variables while subsequent pairs of variates are maximized with respect to the residual variance not accounted for by preceding variates.

For the purpose of this analysis, the canonical variates were composed of the independent variables (received SPL, dolphin age, trial number) and 23 response variables. Prior to performing the CCA, all response variables for which a response was not observed were removed from the data set. (Response variables removed from the data set prior to the CCA are noted with && in Supplemental File 1). Three pairs of canonical variates were formed and the canonical correlation between paired variates was determined to assess the strength of the relationship between them. An *F* statistic was determined for each canonical correlation with $\alpha = 0.05$. Canonical variates with significant correlations were subsequently investigated to determine which of the original independent and dependent variables contributed most to explaining the variance within the model. This was done in two steps. First, the canonical loading for each of the original variables was calculated. The canonical loading reflects the amount of the variance in a canonical variate that is shared with an original variable from which it was derived. Second, the canonical cross-loading for each of the variables was determined. The canonical cross-loading is the correlation between an original independent variable and its dependent canonical variate, and vice-versa. Variables where the absolute value of the canonical loading > 0.32 and the absolute value of the canonical cross-loading > 0.23 were deemed as contributing sufficiently to the observed variance of

the model to warrant inclusion in subsequent analyses. The decision regarding canonical loading and cross-loading thresholds is subjective and raising or lowering the threshold would decrease or increase the number of variables included in subsequent analyses. However, in the analysis presented here, the minimum amount of variance explained within and between variates is ~10% and 5%, respectively. Lowering the threshold for inclusion would increase the number of variables propagated forward but with diminishing returns on the contribution of individual variables to the total variance of the model. For the purpose of describing the remaining analyses, the response variables that met the threshold criteria are termed “contributing responses.”

Once the contributing responses were identified, the data set was reduced to include only the contributing responses and the independent variables. A severity score was applied to each trial, both control and exposure, based on the occurrence of the contributing response. Each response was assigned a severity score and the highest ranking severity score associated with an observed response on a given trial was used as the score for that trial. Severity score assignments were initially attempted using mean and median severity scores derived from the collection of severity scores proposed by the independent review panel. However, because the mean and median scores of the complete set of response variables allowed for little differentiation in the severity of the individual responses (see Section 3. Results), this approach was abandoned and the severity scores were derived from those previously proposed (Southall et al., 2007). The scores and their derivations are described in Table 2.

An average control severity score was calculated for each dolphin by determining the mean severity score across all control trials for a given dolphin. The average severity score was then subtracted from the severity score of each of the dolphin's exposure trials so that each exposure trial had an adjusted severity score. If the score was less than zero for a given trial, then it was set equal to zero for that trial. Thus, each dolphin had ten adjusted severity scores corresponding to the ten trials of the exposure session. A mean-adjusted severity score (\bar{E}_{ij}) was then calculated for each trial (i) and exposure condition (j) by calculating the mean of the adjusted severity scores for all individuals receiving the same exposure level. Thus, for each trial and exposure condition,

$$\bar{E}_{ij} = (1/m) \sum_{k=1}^m E_{ijk}, \quad (1)$$

where k corresponds to an individual dolphin participating in the particular trial and exposure condition, m is the number of dolphins participating in exposure condition j , and E_{ijk} is the adjusted severity score for the dolphin for trial i . For purposes of curve fitting, \bar{E}_{ij} was then normalized to the maximum severity score measured across all exposure trials.

Severity scores are subjective in nature and are likely inappropriate for direct comparisons across species and studies. However, within a study, they provide an opportunity to investigate how responses to a stimulus scale with increasing levels of the stimulus. To this end,

Table 2

Severity of response scores applied to the contributing responses of this study. Scores were derived from the severity scale provided in Southall et al. (2007). Where a description of the behavior did not exist for the “Laboratory subjects” category, the corresponding score from the “Free-ranging subjects” category was utilized.

Behavioral response	Response score	Corresponding behavior (from Table 4 of Southall et al., 2007)
Increase in respiration (2)	3	Moderate change in respiration rate
Fluke/pectoral flipper slap	6	Aggressive behavior related to noise exposure (e.g. tail/flipper slapping)
Refusal to participate	7	Avoidance of experimental situation or retreat to refuge area

an asymmetric function was fit to the distribution of severity scores by received SPL for each trial such that ten dose-severity functions were created, one for each trial. An asymmetric function was chosen as it better fit the distribution of the data than did available symmetric dose functions. The form of the asymmetric function was:

$$E_X = B + \left[1 + 10^{H(X_b - X)}\right]^{-S}, \quad (2)$$

where E_X is the predicted severity score, B is the basal value, X is the received SPL, and X_b (also in SPL), H and S are optimally fit through a least-squared error minimization. Since it was assumed that under non-exposure conditions no aversive reaction related to the exposure could occur, B was set equal to zero.

A similar procedure was performed to determine the probability of response as a function of received SPL, i.e. a dose-response function. First, the probability of a response for all k subjects in a given trial and exposure level was determined. Using the same asymmetric equation previously described, the function describing the probability of a behavioral response on a given trial and exposure level (P_{ik}) was fit via least-squared error minimization. The potential for a behavior to occur during a control session that would be categorized as a contributing response was also determined to provide context to the lower limit of reliable response detection and categorization.

3. Results

3.1. Characterization of the received SPL

The difference between the mean received SPL and that expected across the region of the test enclosure where sound exposures were triggered is shown in Fig. 2. Except for one recording location (designated Δ), the mean received SPL was within ± 3 dB of the expected exposure level for all trials. Accurate received levels were most difficult to achieve at the lowest and highest sound transmission levels but were generally within a couple of dB of the desired exposure level.

3.2. Initial analysis – behavioral responses and severity scores

The initial analysis considering all of the potential behavioral responses indicated that behavioral responses were observed on 83% of the exposure trials (see Supplemental File 3). However, behaviors that qualified as behavioral responses were also observed on 61% of the control trials suggesting that the proposed response indicators were insufficient to adequately capture responses due to the sound exposure. For this reason, a canonical correlation analysis was performed (see Section 2. Materials and methods) to reduce the parameter space to those variables which contributed to explaining the greatest amount of shared variance within the canonical model.

The scores of 14 independent individuals that assigned a severity score to the list of anticipated behavioral reactions generated prior to conducting the study can be found in Supplemental File 2. The proposed severity scores varied considerably for each behavior; the range of proposed scores for individual behaviors spanned 50 points or more for 91% of the behaviors. As a result, there was high variance but a narrower range of mean and median values. Approximately 41% of the mean scores were between 40 and 60 and 76% were between 30 and 70. Because there was wide variation in individual scores but relatively little variation within mean/median scores, the mean/median severity scores obtained from the survey were not used in scoring the behavioral responses to the sound exposure. Instead, contributing responses (as determined from the CCA) were assigned severity scores derived from Southall et al. (2007) (see Section 2. Materials and methods).

3.3. Canonical correlation and dose–response/dose–severity functions

Table 3 lists the statistical output of the canonical correlation. All three of the canonical variates were significant at the $p = 0.01$ level, although the strongest canonical correlation by far was contained in the first canonical dimension (i.e., between the first pair of independent and dependent variates). Table 4 shows the loadings and cross-loadings resulting from the canonical model. However, for the dependent variables (i.e. response variables), it only reports the loadings and cross-loadings for those behaviors that qualified as contributing responses. Exposure level exhibited the most influence on the variance of the first canonical dimension, i.e. within its own variate and the corresponding dependent canonical variate. Each of the other independent variables explained a significant portion of the variance within one of the other independent variates (Age \gg Variate 3, Trial number \gg Variate 2), but substantially less in the corresponding dependent variates. Nevertheless, based on the criteria for inclusion of response variables for further analysis (absolute value of canonical loading > 0.32 and absolute value of canonical cross-loading > 0.23), the trial number appeared to influence the response of animals to the sound exposure. The negative correlation suggested a reduction in responsiveness with increasing trial number.

The only responses that met the criteria to be contributing responses were a refusal to participate in exposure trials, fluke slaps, and changes in the respiration rate that were more than two standard deviations from the mean respiration rate observed during the control session. The contributing responses were all most strongly correlated with the first independent canonical variate, which was most heavily influenced by the exposure level. Therefore, subsequent dose–response and dose–severity analyses were conducted with exposure level as the independent variable and the three response variables as dependent variables. Because the model suggested that trial number influenced the response of the dolphins to the sound exposure, each trial was assessed independently prior to analyzing all of the data within one dose–response function. To a lesser extent, age also appeared to influence the relationship. However, the issue of age is treated in the Discussion section and was not teased out of the subsequent dose–response analysis.

Table 5 shows the parameter values and R^2 for the best fit of the asymmetric function to the data demonstrating the probability of a response on each trial. The amount of variance that could be explained by the model was lowest on the first exposure trial, but improved with subsequent trials. Fig. 3 shows the family of curves describing the dose–response relationship by trial. As the trial sequence progresses the dose–response curves approximate a step function. The evolution of the dose–response curve with trial sequence is also evident in the lessening and stabilization of $\log X_b$ over the last four trials and is consistent with the reduction in responsiveness with trial progression as suggested by the output of the CCA. The steepness of functions 3–4 and 7–10 is due to the fact that minor changes in the slope cause the greatest error with increasing probability of response, which occurs at ~ 175 dB SPL. This is notable in Fig. 4, which plots the non-linear curve fits of trials 1 and 10. These trials have very different dose–response functions even though the distribution of the response probabilities shares a similar overall shape (i.e. comparing the data points indicated by the triangles). However, it should be noted that trials 3–4 and 7–9 either had

Table 3

The correlation between the paired canonical variates (the canonical dimension) and the results of the statistical relationship.

Canonical dimension	Canonical correlation	F value	Num DF	Den DF	Pr > F
1	0.61	6.14	66	1715	<0.0001
2	0.32	2.45	42	1150	<0.0001
3	0.25	1.99	20	576	0.0063

Table 4

The canonical loadings and cross-loadings for each of the dependent and independent variables that contributed to the greatest amount of shared variance within the canonical model.

	Variate 1	Variate 2	Variate 3
Canonical loadings			
<i>Correlations between the independent variables and their canonical variates</i>			
Exposure level	0.971	0.04	0.235
Age	−0.211	0.505	0.837
Trial number	−0.087	−0.863	0.498
<i>Correlations between the dependent variables and their canonical variates</i>			
Refusal to participate in the trial	0.828	−0.11	0.152
Fluke/pectoral flipper slapping	0.434	−0.053	−0.261
Change in respiration (2)	0.394	0.173	0.01
Canonical cross-loadings			
<i>Correlations between the independent variables and the dependent canonical variates</i>			
Exposure level	0.591	0.013	0.06
Age	−0.129	0.159	0.213
Trial number	−0.053	−0.272	0.127
<i>Correlations between the dependent variables and the independent canonical variates</i>			
Refusal to participate in the trial	0.504	−0.035	0.039
Fluke/pectoral flipper slapping	0.264	−0.017	−0.067
Change in respiration (2)	0.24	0.055	0.003

lower probabilities of response than trial 10 or no responses for received SPL ≤ 160 dB SPL.

Fig. 5 shows the dose–response function for all trials combined, which ignores the trial order effect but increases the sample size. This provides a compromise by increasing data set robustness at the cost of information on changes in animal behavior with repetitive exposures, but might serve as an “averaged” dose–response curve for combinations of naïve and experienced animals. The probability of observing a contributing response in the absence of a sound exposure, as determined from the control data set, was found to be 7%. This baseline value is provided in Fig. 5 for context on the limitation to the ability to relate observed behaviors as responses to a sound exposure.

Fig. 6 shows the family of dose–severity response curves corresponding to the individual trials. The increase in the severity of the response, as scored in this study, was directly related to the increase in the received SPL during exposure trials. The majority of the maximum contributing responses at a received SPL ≤ 160 dB SPL consisted of changes in the respiration rate (Table 6). Fluke slaps (no pectoral slaps observed) only occurred at moderate received SPL whereas refusals to participate increased steadily with increasing received SPL. At the highest level of exposure, every dolphin refused to participate in any of the trials following the sound exposure.

Table 5

Parameter values and R^2 for each of the ten dose–response functions as well as the dose–response function fit to the data when all trials were considered (the “averaged” dose–response function).

Trial	$\log X_b$	H	S	R^2
1	190.4	0.335	0.032	0.565
2	185.0	1.200	0.017	0.935
3	176.3	5.963	0.028	0.897
4	176.0	0.488	0.354	0.954
5	185.4	1.877	0.011	0.954
6	176.5	0.423	0.310	0.625
7	175.3	14.480	0.045	1.000
8	175.2	1.101	0.212	0.961
9	176.3	17.630	0.018	0.886
10	175.5	1.599	0.106	0.773
All	183.2	2.586	0.011	0.882

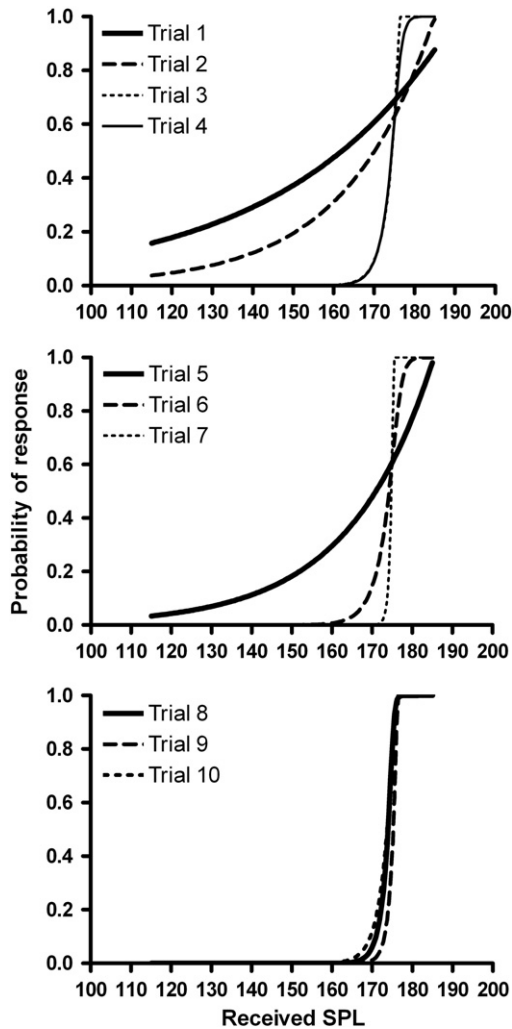


Fig. 3. The family of dose–response curves for exposure trials 1–10. The curves give the probability of a response for a given received SPL. Note that the distribution of data points was best fit by an asymmetric function, not a symmetric sigmoid function as is often used in dose–response studies.

4. Discussion

The purpose of this study was to relate the received SPL of a simulated mid-frequency tactical sonar signal to variations in the behavior of bottlenose dolphins. The dose–response functions that were determined in this study should help inform regulators charged with marine mammal management and protection and improve the application of

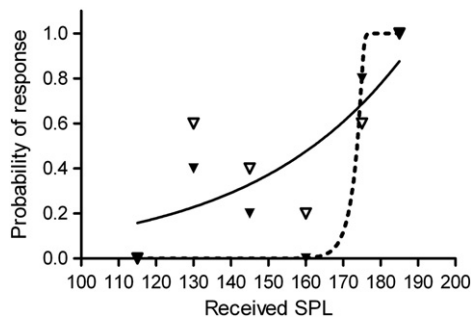


Fig. 4. A comparison of the dose–response functions for trials 1 and 10. Data points corresponding to the average probability of a response for each exposure level are also provided (open triangles = trial 1; filled triangles = trial 10). Note that for received levels ≤ 160 dB SPL, trials 3–4 and 7–9 either had lower probabilities of response than trial 10 or no response at all.

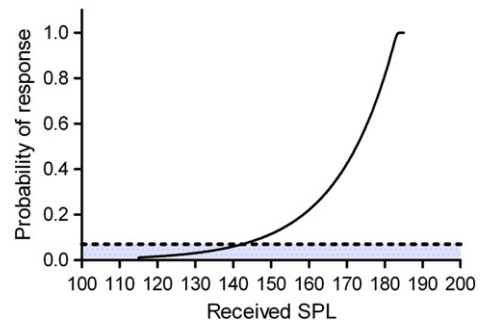


Fig. 5. The “averaged” dose–response function for all trials combined. The shaded region below the dashed line corresponds to the 7% probability of observing a behavior that would qualify as a behavioral response in the absence of a sound exposure. The intersection of the two functions occurs at ~ 143 dB SPL.

dose–function methods in predicting the harassment of marine mammals. The consequences of observed behavioral changes were not quantified, but the relationship to received SPL under a controlled context allows a better understanding of the factors affecting responses to sound exposure. The dolphins participating in the study were under human care and their responses are likely not directly transferrable to conspecifics in the wild. The dolphins have years of experience under stimulus control, which is a necessary condition for the performance of trained behaviors, and they live within an environment with significant boating activity. These factors likely impact the threshold of responsiveness to sound exposure, potentially in the direction of habituation or increased tolerance to noise. Conversely, because of limitations on Navy sonar use within San Diego Bay, the animals are not accustomed to high level noise exposures at tactical sonar frequencies (~ 3 kHz) and may be more naïve to such exposures than wild animals inhabiting regions where tactical sonar activity is more commonplace (e.g. Navy ocean training ranges).

Determining which behaviors are important and reliable indicators of a response to anthropogenic sound exposure and that are therefore suitable for guiding the creation of a dose–response function is daunting. The behavioral repertoire of marine mammals is rich and varied and the variation in the occurrence of a behavior may be quite high in the absence of anthropogenic sound. Of the behaviors initially identified as potential responses to sound exposure in this study, at least one of the behaviors was observed on 83% of the sound exposure trials. However, at least one of the behaviors was observed on 61% of the control trials, as well. This suggests that the normal variation in many of the behaviors is sufficiently high as to make them unreliable indicators of a sound-induced response. The CCA utilized here distilled the set of potential behavioral responses to those that reflected the greatest predictive relationship to sound exposure. This approach was useful in the development of subsequent dose–response relationships, but even these behaviors had a sufficiently high rate of occurrence that they could be misrepresented as a behavioral response to sound exposure 7% of the time. Assuming that the “averaged” dose–response function (Fig. 5) represents a combination of response probabilities from naïve and experienced animals, the probability of reliably categorizing a behavioral response to sound would be limited at received SPLs below ~ 143 dB SPL. Despite the limitations imposed by the natural variability of potential behavioral indicators of a response to sound exposure, approaches similar to the CCA should be considered for determining relevant and reliable behavioral responses for any species under wild or laboratory conditions. This is understandably a difficult task in natural settings due to limitations in animal access, the potential for altering animal behavior through observer presence, and the often unknown context under which observations are made. Nevertheless, characterizing reliable indicators of a behavioral response is necessary in order to determine dose–response relationships with adequate predictive power.

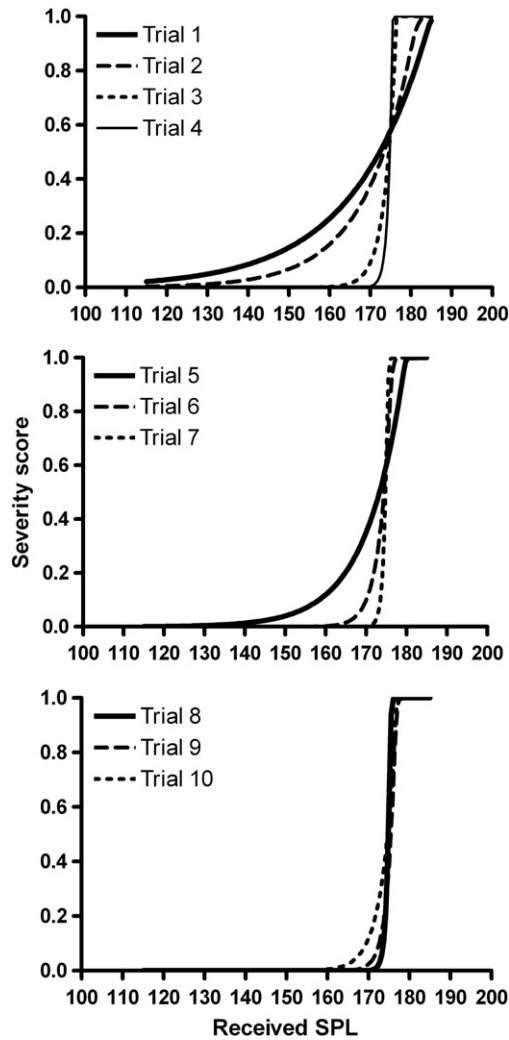


Fig. 6. The family of dose–severity curves for exposure trials 1–10. The curves describe the increase in severity of the response, as defined in this study, with increasing received SPL. As with the dose–response functions, the distribution of data points was best fit by an asymmetric function.

Bottlenose dolphins are generally considered to be a robust cetacean species. The bottlenose dolphin does well under human care and has shown variable degrees of habituation to presumably aversive acoustic stimuli in the wild (Cox et al., 2003; Diaz López and Mariño, 2011; Leeney et al., 2007). Whether or not habituation occurred in these studies was likely due to a combination of factors, including the signal type utilized, the duration over which animals were exposed, and whether other contextual factors were involved (e.g. the presence of food). In the study presented here, the 50% probability of a behavioral response occurred at a received SPL of ~162 dB SPL on the first trial but was ~174 dB SPL by the last trial, a value similar to the mid–point of curves 3–4 and 6–9. The progression and stabilization of the curves suggest

Table 6
Percentage that each of the contributing responses contributed to all observed responses for each of the sound exposure conditions.

Received SPL	Change in respiration (2)	Fluke/pectoral flipper slap	Refusal to participate in the trial
115	100.0%	0.0%	0.0%
130	92.3%	0.0%	7.7%
145	50.0%	33.3%	16.7%
160	75.0%	0.0%	25.0%
175	16.1%	19.4%	64.5%
185	0.0%	0.0%	100.0%

that the dolphins habituated to the exposure quickly over the course of the session, but only at received SPLs of 160 dB and below. At received SPLs of 175 dB and greater, no habituation was observed and every dolphin refused to complete the behavioral sequence when exposures reached 185 dB SPL. In some instances, a single exposure of 175 dB SPL or greater resulted in the dolphin not crossing the pen at the point at which the initial exposure occurred.

Similar avoidance behaviors to those observed here have been observed in captive gray seals (*Halichoerus grypus*) that associated a particular context with an aversive acoustic stimulus (Götz and Janik, 2011). In contrast to the dolphins, although seals that demonstrated startle responses and sensitization had 50% response thresholds ranging from 155 to 160 dB SPL, some seals demonstrated habituation and did not exhibit a startle response or avoidance at received levels of up to 180 dB SPL indicating short-term habituation and an overall increase in tolerance to the stimulus. The rise time of the signal (5 ms) and the sensation level (~93 dB SL) of the signal, which is the received level of sound referenced to an animal’s threshold of detection for the same sound, were critical components in affecting sensitization in those seals that demonstrated the startle response. In this study, the rise time of the exposure signal was 50 ms, which is considerably longer than that usually required to induce startle in mammals (see Götz and Janik, 2011). Under quiet testing conditions, dolphin hearing thresholds at or near 3 kHz are ~80–85 dB SPL (Finneran and Schlundt, 2007; Finneran et al., 2005; Johnson, 1967). Assuming that prior threshold estimates are representative of the group of dolphins in this study, exposure levels ≥ 175 dB SPL would have a sensation level > 90 dB. However, the interpretation of this comparison is complicated by the fact that thresholds at 3 kHz are masked by as much as 20 dB within San Diego Bay (Finneran et al., 2005); i.e. it is unclear if the rise time and sensation level are critically related, or whether the critical relationship is between the rise time and the signal to noise ratio. Given the relatively long rise time used in this study, the progression from habituation to sensitization and area avoidance with increasing received level suggests a process that is amplitude-dependent. Nevertheless, based on prior evidences, it is reasonable to expect this relationship might change toward increased sensitization at lower received levels if the rise time of the signal was sufficiently short to induce a startle response.

The degree to which a given stimulus is aversive likely varies as a function of the level at which the stimulus is perceived, the rise time of the received stimulus, the temperament of the individual exposed to the stimulus, and the degree and rate of stimulus repetition relative to the absence or occurrence of associated negative consequences. Prior work with laboratory and wild animals has demonstrated that the type and magnitude of a stimulus greatly influences whether habituation or sensitization occurs and the rate at which it is manifest (Bowles and Anderson, 2012; Ellenberg et al., 2012; Heyser and Chemero, 2012). Temperament type, which reflects genetic (species or gender) tolerances and intra-individual variability and which can be influenced by early life experience, is also an important feature impacting responsiveness to novel or aversive stimuli (Biro and Dingemans, 2009; Bowles and Anderson, 2012; Carter et al., 2012; Ruis et al., 2001; Stamps et al., 2012; Tang et al., 2012; Vandenheede and Bouissou, 1993). In this study, the statistical model suggested age was a factor affecting the responsiveness of individuals to the sound exposure. However, this finding was biased by a single animal, D20, a male dolphin that was 10 years old at the time of the study. Dolphin D20 exhibited a behavioral response on nine out of ten trials, each one categorically defined as a change in the respiration rate. The distribution of behavioral responses of other similarly aged animals (6–10 years old) was not substantially different than older animals under similar exposure conditions. Thus, factors influencing the responsiveness of D20 likely reflect either inherent sensitivities or contextual factors that were unknown to the experimenters. No other indication of age influencing response thresholds or the nature of the response was observed, although testing was limited to dolphins greater than six years of age. It cannot be ruled out, and should probably

be expected, that younger animals with less experience might respond differently at a lower received levels of the same stimulus.

One of the challenges to estimating the impact to marine mammals that result from sound-induced behavioral reactions is determining the cost of the altered behavior to the animal. There are numerous ways to quantify the costs associated with altered natural behavior, each of which has its own limitations and most of which are interdependent. There are obvious inherent difficulties in obtaining quantitative measures of cost, but even qualitative measures and the scaling of response severity seem to be challenged by differences in opinion among experts (Supplemental File 2). As reflected by the comments of some participants in our exercise in scoring behavioral response severity, context is important. For some behaviors, a change in behavior may have negligible costs in one context but may be extremely costly or beneficial in another (e.g. ceasing sound production in the absence or presence of a predator). Thus, although severity scores may have value for investigating responses within a study (i.e. a defined context), their utility in comparing across studies may be difficult and have questionable interpretive value as a result of contextual dissimilarity. In addition, as determined from the survey of experts, community agreement on whether many behavioral responses qualify as having significant impact to any species may be difficult or impossible to achieve.

The reliable predictors of a behavioral response to the simulated sonar signal used in this study were a change in respiration, fluke slaps on the water surface, and abandonment of the trial behavior. The severity scaling applied in this study, as derived from Southall et al. (2007), permitted patterns in the severity of the response to be related to both sound exposure and trial sequence. The dose–severity functions suggested a decline in the severity of the response with the progression of the trials at received SPL \leq 160 dB SPL. The particular response severity scores also provide some qualitative linkage to impacts to the individual, albeit with the caveats previously presented. An abandonment of behavior involving a food reward is the easiest to conceptually extrapolate to wild animals since abandonment of behaviors such as foraging or reproductive displays can have immediate energetic and fitness consequences directly related to the duration of the abandonment. Fluke slaps are commonly believed to be associated with aggravation or aggression in the bottlenose dolphin. Fluke slaps could possibly be linked with an increased state of stress, the accumulation of which could also have energetic and fitness consequences. However, such a relationship has not yet been ascertained in this species. A change in the respiration rate is the most difficult to qualitatively link to costs, particularly if it is of short duration and occurs in the absence of other behavioral responses. A change in respiration was the most common behavioral response to received SPL \leq 160 dB SPL. However, the standard deviation of the respiration rate across the control trials was rarely $>$ 1.0 and the behavioral response to a sound exposure most commonly observed was the cessation of breathing or an increase by one or two breaths over the 30 s trial interval. These responses were often immediately preceded or followed by periods where patterns of respiration were “normal.” Thus, due to the acute nature of the response, its high natural variability, and the absence of other behavior alterations, it becomes increasingly difficult to evaluate the cost of this behavioral change, if any.

5. Conclusions

Bottlenose dolphins performing a trained paddle-touch behavior and exposed to a simulated mid-frequency tactical sonar signal demonstrated rapid habituation to exposures \leq 160 dB SPL. Abandonment of the trained behavior increased dramatically at exposures \geq 175 dB SPL and all dolphins abandoned the behavior at received levels of 185 dB SPL. The most common behavioral response to received SPL \leq 160 dB was a change in the respiration rate, but this response was variable across trials and individuals. Dose–response functions, like those derived here, permit the probability of a behavioral reaction to a sound exposure to be predicted. Their application to estimating

the consequences of sound exposure has the potential to improve conservation efforts. However, the methodology is limited by the reliability of the behavior as a response indicator as well as the variability and occurrence of the behavior in the absence of an acoustic exposure. Furthermore, dose–response functions likely vary depending on the characteristics of the sound exposure (e.g. continuous vs. impulsive sound types) and the species studied. Extrapolating to the more critical question of biological impacts will require ingenuity in linking the severity of the response, which may scale in magnitude with the dose–response function, to individual and population costs.

Supplementary data to this article can be found online at <http://dx.doi.org/10.1016/j.jembe.2013.02.043>.

Acknowledgments

The authors wish to thank all of the trainers within the MMP that participated in the study and trained the behavior required for testing. Special gratitude is extended to R. Dear, L. Green, and M. Tormey for serving as trainers during the sessions; R. Jauck and E. Bauer for providing engineering support in the construction of the test enclosure; and P. Banks and K. Bakhtiari for analyzing the control and exposure trials for behavioral responses. Additional thanks are given to the group of marine mammal experts that provided severity scores for the list of anticipated behavioral responses. This study was funded by the Office of Naval Research (award #N0001409WX20853) and the Chief of Naval Operations Energy and Environmental Readiness Division (award #N0002409WX11974). All procedures involved in this study were approved by the Institutional and Animal Care Utilization Committee of the MMP. [SS]

References

- Biro, P.A., Dingemans, N.J., 2009. Sampling bias resulting from animal personality. *Trends Ecol. Evol.* 24 (2), 66–67.
- Bowles, A.E., Anderson, R.C., 2012. Behavioral responses and habituation of pinnipeds and small cetaceans to novel objects and simulated fishing gear with and without a pinger. *Aquat. Mamm.* 38 (2), 161–188.
- Carter, A.J., Heinssohn, R., Goldizen, A.W., Biro, P.A., 2012. Boldness, trappability and sampling bias in wild lizards. *Anim. Behav.* 83 (4), 1051–1058.
- Cox, T.M., Read, A.J., Swanner, D., Urian, K., Waples, D., 2003. Behavioral responses of bottlenose dolphins, *Tursiops truncatus*, to gillnets and acoustic alarms. *Biol. Conserv.* 115 (2), 203–212.
- D'Amico, A., Gisiner, R.C., Ketten, D.R., Hammock, J.A., Johnson, C., Tyack, P.L., Mead, J., 2009. Beaked whale strandings and naval exercises. *Aquat. Mamm.* 35 (4), 452–472.
- Delaney, D.K., Grubb, T.G., Beier, P., Pater, L.L., Reiser, M.H., 1999. Effects of helicopter noise on Mexican spotted owls. *J. Wildl. Manag.* 63 (1), 60–76.
- Delaney, D.K., Pater, L.L., Carlile, L.D., Spadgenske, E.W., Beaty, T.A., Melton, R.H., 2011. Response of red-cockaded woodpeckers to military training operations. *Wildl. Monogr.* 177, 1–38.
- Department of the Navy (DoN), 2008a. Hawaii Range Complex Final Environmental Impact Statement/Overseas Environmental Impact Statement (EIS/OEIS). U.S. Department of Defense, Washington, DC.
- Department of the Navy (DoN), 2008b. Southern California Range Complex: Final Environmental Impact Statement/Overseas Environmental Impact Statement. Department of the Navy, Washington, DC.
- Department of the Navy (DoN), 2008c. Atlantic Fleet Active Sonar Training: Final Environmental Impact Statement/Overseas Environmental Impact Statement (FEIS/OEIS). Department of the Navy, Washington, DC.
- Diaz López, B., Mariño, F., 2011. A trial of acoustic harassment device efficacy on free-ranging bottlenose dolphins in Sardinia, Italy. *Mar. Freshw. Behav. Physiol.* 44 (4), 197–208.
- Ellenberg, U., Mattern, T., Houston, D., Davis, L., Seddon, P., 2012. Previous experiences with humans affect responses of Snares Penguins to experimental disturbance. *J. Ornithol.* 153 (3), 621–631.
- Fernández, A., Edwards, J., Martín, V., Rodríguez, F., Espinosa de los Monteros, A., Herráez, P., Castro, P., Jaber, J.R., Arbelo, M., 2005. “Gas and fat embolic syndrome” involving a mass stranding of beaked whales (family *Ziphiidae*) exposed to anthropogenic sonar signals. *J. Vet. Pathol.* 42, 446–457.
- Finneran, J.J., Schlundt, C.E., 2007. Underwater sound pressure variation and bottlenose dolphin (*Tursiops truncatus*) hearing thresholds in a small pool. *J. Acoust. Soc. Am.* 122 (1), 606–614.
- Finneran, J.J., Dear, R., Carder, D.A., Ridgway, S.H., 2003. Auditory and behavioral responses of California sea lions (*Zalophus californianus*) to single underwater impulses from an arc-gap transducer. *J. Acoust. Soc. Am.* 114 (3), 1667–1677.

- Finneran, J.J., Carder, D.A., Schlundt, C.E., Ridgway, S.H., 2005. Temporary threshold shift (TTS) in bottlenose dolphins (*Tursiops truncatus*) exposed to mid-frequency tones. *J. Acoust. Soc. Am.* 118 (4), 2696–2705.
- Götz, T., Janik, V.M., 2011. Repeated elicitation of the acoustic startle reflex leads to sensation in subsequent avoidance behaviour and induces fear conditioning. *BMC Neurosci.* 12 (30), 13.
- Goudie, R.L., Jones, I.L., 2004. Dose-response relationships of harlequin duck behaviour to noise from low-level military jet over-flights in central Labrador. *Environ. Conserv.* 31 (4), 289–298.
- Hair, J.F., Anderson, R.E., Tatham, R.L., Black, W.C., 1998. Canonical correlation analysis. *Multivariate Data Analysis.* 768.
- Heyser, C.J., Chemero, A., 2012. Novel object exploration in mice: not all objects are created equal. *Behav. Process.* 89 (3), 232–238.
- Holst, M., Greene Jr., C.R., Richardson, W.J., McDonald, T.L., Bay, K., Schwartz, S.J., Smith, G., 2011. Responses of pinnipeds to Navy missile launches at San Nicolas Island, California. *Aquat. Mamm.* 37 (2), 139–150.
- Holt, M.M., Noren, D.P., Emmons, C.K., 2011. Effects of noise levels and call types on the source levels of killer whale calls. *J. Acoust. Soc. Am.* 130 (5), 3100–3106.
- Jepson, P.D., Arbelo, M., Deaville, R., Patterson, I.A.R., Castro, P., Baker, J.R., Degollada, E., Ross, H.M., Herráez, P., Pocknell, A.M., Rodriguez, E., Howie, F.E., Espinosa, A., Reid, R.J., Jaber, J.R., Martin, V., Cunningham, A.A., Fernandez, A., 2003. Gas-bubble lesions in stranded cetaceans: was sonar responsible for a spate of whale deaths after an Atlantic military exercise? *Nature* 425, 575–576.
- Johnson, C.S., 1967. Sound detection thresholds in marine mammals. In: Tavolga, W.N. (Ed.), *Marine Bioacoustics*. Pergamon Press, Oxford, pp. 247–260.
- Johnston, D.W., 2002. The effect of acoustic harassment devices on harbour porpoises (*Phocoena phocoena*) in the Bay of Fundy, Canada. *Biol. Conserv.* 108 (1), 113–118.
- Kastelein, R.A., Steen, N., Gransier, R., Wensveen, P.J., de Jong, C.A.F., 2012. Threshold received sound pressure levels of single 1–2 kHz and 6–7 kHz up-sweeps and down-sweeps causing startle responses in a harbor porpoise (*Phocoena phocoena*). *J. Acoust. Soc. Am.* 131 (3), 2325–2333.
- Leeney, R.H., Berrow, S., McGrath, D., O'Brien, J., Cosgrove, R., Godley, B.J., 2007. Effects of pingers on the behaviour of bottlenose dolphins. *J. Mar. Biol. Assoc. UK* 87 (01), 129–133.
- McCarthy, E., Moretti, D., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, S., Ward, J., Izzii, A., Dilley, A., 2011. Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar. *Mar. Mammal Sci.* 27 (3), E206–E226.
- Miksis-Olds, J.L., Wagner, T., 2011. Behavioral response of manatees to variations in environmental sound levels. *Mar. Mammal Sci.* 27 (1), 130–148.
- Miksis-Olds, J.L., Donaghay, P.L., Miller, J.H., Tyack, P.L., Nystuen, J.A., 2007. Noise level correlates with manatee use of foraging habitats. *J. Acoust. Soc. Am.* 121 (5), 3011–3020.
- Miller, P.J.O., Johnson, M.P., Madsen, P.T., Biassoni, N., Quero, M., Tyack, P.L., 2009. Using at-sea experiments to study the effects of airguns on the foraging behavior of sperm whales in the Gulf of Mexico. *Deep-Sea Res.* 1 (56), 1168–1181.
- Morton, A.B., Symonds, H.K., 2002. Displacement of *Orcinus orca* (L.) by high amplitude sound in British Columbia, Canada. *ICES J. Mar. Sci.* 59, 71–80.
- National Research Council (NRC), 2005. *Marine Mammal Populations and Ocean Noise*. National Academies Press, Washington, DC.
- Niu, F.-q., Liu, Z.-w., Wen, H.-t., Xu, D.-w., Yang, Y.-m., 2012. Behavioral responses of two captive bottlenose dolphins (*Tursiops truncatus*) to a continuous 50 kHz tone. *J. Acoust. Soc. Am.* 131 (2), 1643–1649.
- Nowacek, D.P., Johnson, M.P., Tyack, P.L., 2004. North Atlantic right whales (*Eubalaena glacialis*) ignore ships but respond to alerting stimuli. *Proc. Roy. Soc. B* 271, 227–231.
- Olesiuk, P.F., Nichol, L.M., Sowden, M.J., Ford, J.K.B., 2002. Effect of the sound generated by an acoustic harassment device on the relative abundance and distribution of harbor porpoises (*Phocoena phocoena*) in retreat passage, British Columbia. *Mar. Mammal Sci.* 18 (4), 843–862.
- Parks, S.E., Clark, C.W., Tyack, P.L., 2007. Short- and long-term changes in right whale calling behavior: the potential effects of noise on acoustic communication. *J. Acoust. Soc. Am.* 122 (6), 3725–3731.
- Pater, L.L., Grubb, T.G., Delaney, D.K., 2009. Recommendations for improved assessment of noise impacts on wildlife. *J. Wildl. Manag.* 73 (5), 788–795.
- Ruis, M.A.W., de Groot, J., te Brake, J.H.A., Dinand Ekkel, E., van de Burgwal, J.A., Erkens, J.H.F., Engel, B., Buist, W.G., Blokhuis, H.J., Koolhaas, J.M., 2001. Behavioural and physiological consequences of acute social defeat in growing gilts: effects of the social environment. *Appl. Anim. Behav. Sci.* 70 (3), 201–225.
- Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J.J., Gentry, R.L., Greene Jr., C.R., Kastak, D., Ketten, D.R., Miller, J.H., Nachtigall, P.E., Richardson, W.J., Thomas, J.A., Tyack, P.L., 2007. Marine mammal noise exposure criteria: initial scientific recommendations. *Aquat. Mamm.* 33 (4), 411–521.
- Stamps, J.A., Briffa, M., Biro, P.A., 2012. Unpredictable animals: individual differences in intraindividual variability (IIV). *Anim. Behav.* 83 (6), 1325–1334.
- Tang, A.C., Reeb-Sutherland, B.C., Romeo, R.D., McEwen, B.S., 2012. Reducing behavioral inhibition to novelty via systematic neonatal novelty exposure: the influence of maternal hypothalamic-pituitary-adrenal regulation. *Biol. Psychiatry* 72 (2), 150–156.
- Tyack, P.L., 2008. Implications for marine mammals of large-scale changes in the marine acoustic environment. *J. Mammal.* 89 (3), 549–558.
- Tyack, P., 2009. Acoustic playback experiments to study behavioral responses of free-ranging marine animals to anthropogenic sound. *Mar. Ecol. Prog. Ser.* 395, 187–200.
- Tyack, P.L., Zimmer, W.M.X., Morretti, D., Southall, B.L., Claridge, D.E., Durban, J.W., Clark, C.W., D'Amico, A., DiMarzio, N., Jarvis, S., McCarthy, E., Morrissey, R., Ward, J., Boyd, I.L., 2011. Beaked whales respond to simulated and actual Navy sonar. *PLoS One* 6 (3).
- Vandenheede, M., Bouissou, M.F., 1993. Sex differences in fear reactions in sheep. *Appl. Anim. Behav. Sci.* 37 (1), 39–55.
- Wöhr, M., Borta, A., Schwarting, R.K.W., 2005. Overt behavior and ultrasonic vocalization in a fear conditioning paradigm: a dose–response study in the rat. *Neurobiol. Learn. Mem.* 84, 228–240.