Aquatic Mammals 2012, *38*(4), 362-401, DOI 10.1578/AM.38.4.2012.362

The Severity of Behavioral Changes Observed During Experimental Exposures of Killer (*Orcinus orca*), Long-Finned Pilot (*Globicephala melas*), and Sperm (*Physeter macrocephalus*) Whales to Naval Sonar

Patrick J. O. Miller,[1] Petter H. Kvadsheim,[2] Frans-Peter A. Lam,[3]
Paul J. Wensveen,[1] Ricardo Antunes,[1] Ana Catarina Alves,[1] Fleur Visser,[3, 4]
Lars Kleivane,[2] Peter L. Tyack,[1, 5] and Lise Doksæter Sivle[6]

[1]*Sea Mammal Research Unit, Gatty Marine Laboratory, University of St Andrews, St Andrews, Fife KY16 8LB, UK*
E-mail: pm29@st-andrews.ac.uk
[2]*Norwegian Defence Research Establishment (FFI), NO-3191, Horten, Norway*
[3]*Netherlands Organisation for Applied Scientific Research (TNO), PO Box 96864, The Hague, 2509 JG, The Netherlands*
[4]*Kelp Marine Research (KMR), 1624 CJ, Hoorn, The Netherlands*
[5]*Biology Department, Woods Hole Oceanographic Institution (WHOI), Woods Hole, MA 02543, USA*
[6]*Institute of Marine Research (IMR), Bergen, Norway*

Abstract

This study describes behavioral changes of wild cetaceans observed during controlled exposures of naval sonar. In 2006 through 2009, 14 experiments were conducted with killer (n = 4), long-finned pilot (n = 6), and sperm (n = 4) whales. A total of 14 6-7 kHz upsweep, 13 1-2 kHz upsweep, and five 1-2 kHz downsweep sonar exposures, as well as seven Silent vessel control exposure sessions and eight playbacks of killer whale sounds were conducted. Sonar signals were transmitted by a towable source that approached each tagged subject from a starting distance of 6 to 8 km with a ramp up of source levels (from 152 to 158 to a maximum of 198 to 214 dB re: 1 µPa m). This procedure resulted in a gradual escalation of the sonar received level at the whale, measured by towed hydrophones and by tags that record movement and sound (Dtags). Observers tracked the position of each tagged animal and recorded group-level surface behavior. Two expert panels independently scored the severity of diverse behavioral changes observed during each sonar and control exposure, using the 0 to 9 point severity scale of Southall et al. (2007), and then reached consensus with a third-party moderator. The most severe responses scored (i.e., most likely to affect vital rates) included a temporary separation of a calf from its group, cessation of feeding or resting, and avoidance movements that continued after the sonar stopped transmitting. Higher severity scores were more common during sonar exposure than during Silent control sessions. Scored responses started at lower sound pressure levels (SPLs) for killer whales and were more severe during sonar exposures to killer and sperm whales than to long-finned pilot whales. Exposure sessions with the higher source level of 1 to 2 kHz sonar had more changes and a trend for higher maximum severity than 6 to 7 kHz sessions, but the order of the sessions had no effect. This approach is helpful to standardize the description of behavioral changes that occurred during our experiments and to identify and describe the severity of potential responses of free-ranging cetaceans to sonar.

Key Words: effects of sonar, controlled exposure experiments, severity scoring, cetaceans, ocean noise

Introduction

There has been increasing concern about the effects of underwater sound on marine mammals (Richardson et al., 1995; Nowacek et al., 2007), especially since naval sonar exercises were linked to atypical mass strandings of several cetacean species (Frantzis, 1998; Balcomb & Claridge, 2001; D'Amico et al., 2009). In addition to causing lethal strandings, sonar activities can also disrupt behavior of animals (Maybaum, 1993; Miller et al., 2000).

Southall et al. (2007) extensively reviewed the existing literature on behavioral effects of sound and the sound pressure levels (SPLs) associated with changes in behavior. An important innovation of Southall et al. is that responses were scaled according to their perceived severity for the affected animals by a panel of experts in marine mammal behavioral biology. Southall et al. applied a 10 point scale (see Table 1) ranging from no effect (0), effects not likely to influence vital rates (1 to 3), effects that could affect vital rates (4 to 6), to effects that were thought likely to influence vital rates (7 to 9). Implicit in the severity scale is the consideration that the severity of a response depends upon its duration in relation to the exposure duration. Thus, a behavioral response that continued beyond the end of an exposure (*prolonged*) was considered more severe than a response that stopped when the exposure stopped (*moderate*). Though the distinction between these categories is somewhat subjective, it provides a framework for evaluating observations of how animals change their behavior in response to anthropogenic noise.

Observations of cetaceans during actual sonar exercises are critical to advancing our understanding of how actual sonar usage affects them (Moretti et al., 2010; Tyack et al., 2011), but experimental approaches can enable critical measurements and controls that are difficult to accomplish with observational research (Tyack et al., 2004; Tyack, 2009). When experimenters control the location and timing of the sonar source, they are able to assure that adequate baseline behavior is collected for each individual animal before the sonar stimulus commences. This experimental control improves the ability to describe whether and how behavior is altered as a consequence of the sound exposure, and to test whether any change was caused by the sonar exposure. Control of the sound source also allows experimenters to control the sound dose to which the subject is exposed. One concern for observational studies in areas where sonar is repeatedly used is that animal subjects may be biased towards those individuals that remain in the area; if more sensitive animals leave the area, they might never be sampled (Bejder et al., 2006a, 2006b). In short-term experiments, subjects are targeted within a selected study area, and their inclusion in an experiment is only influenced by the experimenter's ability to record their behavior. Ideally, the subjects in the experiments should be representative for the species, and care should be taken to identify individual subjects using natural markings to document whether or not they had been previously exposed to the same experimental procedure. An experimental approach makes it possible to test for reactions to various control stimuli, including Silent vessel approaches and biologically relevant sounds such as those of predators. Presentation of control stimuli, seldom possible in actual exercises, can help us understand what features of the sonar exposure are most relevant to the subject animals.

Attachment of animal-borne tags and visual and/or acoustic tracking of the movements and social behavior of the tagged subject and its associated group increase the power to observe the precise outcomes of each exposure. A number of studies use high-resolution tag technology like the acoustic and motion recording data-logger Dtag (Johnson & Tyack, 2003). This approach enables data collection at more detailed spatial and temporal scales than most studies considered by Southall et al. (2007). High-resolution observation of behavior above and below the sea surface allows detection of subtle or short-duration behavioral changes that are not incorporated in the existing severity scale. The ability to interpret behaviors, such as foraging (Miller et al., 2004a; Simon et al., 2005) or resting (Miller et al., 2008), in terms of their functions increases our ability to judge the potential biological consequences of behavioral disturbance.

Studies using controlled experiments (Malme et al., 1983, 1984; Richardson et al., 1985, 1990; McCauley et al., 2000; Miller et al., 2000, 2009; Nowacek et al., 2004; Tyack et al., 2011) have reported behavioral responses in free-ranging cetaceans at exposure levels that are often much lower than those that impair hearing (Finneran et al., 2005, 2010). The literature review of Southall et al. (2007) indicates that marine mammals may respond at very different sound levels depending on the context of the exposure (e.g., behavioral state, previous experience with sound type, cost/benefit of staying in the area).

In this study, multiple behavioral parameters and an expert scoring system are used to describe whether and how behavior was altered in a set of experiments in which cetaceans were exposed to a controlled presentation of sonar signals, Silent vessel approaches, and playback of killer whale sounds. Killer whale sound playbacks were conducted in order to contrast behavioral changes to biologically relevant natural signals with changes during sonar exposures. The behavior of the animals was studied using a combination of tags that record sound and movement, visual tracking of individual animals, and visual observations of each tagged whale's group. Summaries of these observational data streams, associated to experimental exposure to sonar and control stimuli, have been compiled into a comprehensive technical report describing each of the experiments (Miller et al., 2011). Visual tracking of a tagged whale and its associated group, aided by radio-tracking of the VHF beacon of the tag and data recorded with the movement-recording Dtag, allow us to reconstruct a detailed 3-dimensional (3-d) movement track of each subject (e.g., Miller et al., 2009). Functional behavioral states such as foraging, resting, or travelling are reflected in combinations of horizontal movement, vertical movement (diving), interaction, synchrony and coordination between group members, and sound production of the tagged whales and other nearby animals (Miller et al., 2004a, 2008; Simon et al., 2005). The suite of measured parameters allows us to describe specific functions for behavioral states that changed during exposure to the sonar. For example, cessation of foraging would be reflected in a change of diving behavior and altered movement patterns along with a change in the production of sounds associated with feeding such as echolocation clicks or tailslaps. In group-living animals, changes in behavioral state are likely to be reflected in differences in group spacing and swimming synchrony.

Descriptive analysis of our experimental data set indicated diverse behavioral changes that could have been reactions to the sonar (Miller et al., 2011). In order to interpret all observed behavioral changes as possible responses, we have here adopted the approach taken by Southall et al. (2007) to follow an objective structured framework to judge the severity of changes in behavior that scale with increasing cost or risk to the subject, and to relate observations of such changes with the received level of the sonar immediately before the behavioral changes started to occur. As much as possible, we sought to follow the criteria already developed in Southall et al., but some modifications, including shorter-duration behavioral changes and specific categories for alterations of feeding or resting behaviors, were necessary to make the conclusions more appropriate for the high temporal resolution and broad suite of measures inherent in our data.

Materials and Methods

Animal Welfare Considerations

All of our research activities were licenced under a permit provided by the Norwegian Animal Research Authority (NARA Permit No. 2004/20607 and S-2007/61201) and were approved by institutional animal welfare committees (the St Andrews Animal Welfare and Ethics Committee and the Woods Hole Oceanographic Institution [WHOI] Institutional Animal Care and Use Committee). During the course of our experiments, we sought to protect the welfare of our research subjects with a safety plan that included specific procedures to reduce risk to each research subject (and any other animals or humans diving in the area). Visual observers continuously scanned for subjects and other whales throughout the exposures with a detailed plan in place to stop sonar transmissions if potentially dangerous responses occurred or if cetaceans came too close to the sonar source. Other aspects of our research design also reduced the risk of disturbance, including limited duration of exposure periods, changing of subjects between experiments, and collecting a relatively small number of samples.

Target Species and Research Area

Sonar experiments were conducted during three cruises in 2006, 2008, and 2009 (Table 2). The experiments were conducted along the coast of Northern Norway between 66 and 70° northern latitude in the winter of 2006, and the late springs of 2008 and 2009. In November 2006, the study species was restricted to herring-feeding killer whales (*Orcinus orca*) within the Vesfjord-Ofotfjord-Tysfjord area. In 2008 and 2009, the operational area included offshore areas extending to, and somewhat beyond, the continental shelf break, and the study species were expanded to include sperm whales (*Physeter macrocephalus*) and long-finned pilot whales (*Globicephala melas*).

Research Platforms

Full details of the field trial methodologies were reported in Miller et al. (2011). All experiments were conducted using two ships: a source vessel and an independent tracking and observation vessel. In all three trials (2006, 2008, and 2009), the 55-m R/V *H.U. Sverdrup II* was the vessel from which the sonar source was deployed and operated. The *Sverdrup* (hereafter "source vessel") served as a command and control center throughout the trials and had visual and VHF tracking stations on the bridge and flying bridge. Boats dedicated to tagging were launched and recovered from the *Sverdrup*. Tracking and observations of the tagged animals during experiments were made using the *Sverdrup*'s workboat in 2006 or the 29-m *MS Strønstad* (hereafter "observation vessel") in 2008 and 2009.

Passive Acoustic Monitoring

Passive acoustic monitoring was used to locate cetaceans for research and for sperm whales to track the subject whale during each experiment. Passive acoustic observation and tracking were carried out using a single 54-m line array (Delphinus, TNO, The Netherlands) deployed from the source vessel containing 18 hydrophones capable of recording frequencies up to 20 kHz (sampled at 48 kHz), three hydrophones up to 160 kHz (sampled at 400 kHz), and a depth sensor. The hydrophone section contains 16 closely spaced hydrophones and three widely spaced hydrophones to perform real-time beamforming to detect, classify, and localize vocalizations < 12 kHz and > 12 kHz, respectively. Another passive array was towed from the observation vessel and consisted of 16 hydrophones at 13-cm spacing (Miller & Tyack, 1998). The acoustic sensitivity of this array for frequencies of 2.2 to 20 kHz was 172 ± 2 dB re: 1 V µPa^{-1}. Signals from 12 channels of the array were recorded at 96 kHz with 24-bit resolution.

Sonar Source

The sonar source (Socrates, TNO, The Netherlands) deployed from the source vessel during the experiments was a multipurpose towed acoustic source that is part of the prototype Low Frequency Active Sonar (LFAS) system being tested on board the multipurpose frigates of the Royal Netherlands Navy. The system contains one hydrophone and depth, pitch, roll, and temperature sensors. The system was designed to enable ramp-up procedures. Waveform, pulse length, duty cycle, and source level were specified by an operator. The operating software had a shutdown function that allowed transmissions to be terminated.

Three types of sonar signals were transmitted during exposure experiments:

1. MFAS-UP (6 to 7 kHz hyperbolic upsweep)
2. LFAS-UP (1 to 2 kHz hyperbolic upsweep)
3. LFAS-DN (1 to 2 kHz hyperbolic downsweep)

The sonar signal rms source level started at 152 dB re: 1 µPa m for 1 to 2 kHz and 158 dB for 6 to 7 kHz sessions, and increased to full power over 10 min. The signal duration was 1 s, and the interval between signals was 20 s during both ramp-up and full power transmission, except that the interval during ramp up was 10 s in 2006. Maximum rms source levels in 2006 were 209 dB and 197 dB re: 1 µPa m for the 1 to 2 kHz and 6 to 7 kHz bands, respectively. In 2008 and 2009, they were 214 dB and 199 dB re: 1 µPa m, respectively. For Silent control approaches, the source was towed and the timing of mock sonar *pings* were recorded, but no sounds were transmitted. We confirmed that no signals were transmitted in the Silent condition by monitoring a hydrophone on the source and in one case by attaching a Dtag directly onto the source during a test run.

Data-Recording Tag

A criterion for each experiment was that one or more of the subjects had to be tagged with a data-recording tag that also contained a VHF transmitter. Miniature high-resolution movement-and-sound-recording Dtags were utilized (Johnson & Tyack, 2003). The Dtag was attached to the whale with suction cups using a handheld carbon fiber pole or a pneumatic remote deployment system (ARTS; Kvadsheim et al., 2009) modified for the suction cup attachment. The Dtag recorded stereo sound at the whale with 16-bit resolution at a 96 or 192 kHz sampling rate. The tag also recorded depth, temperature, 3-d acceleration, and 3-d magnetometer information. The non-acoustic Dtag sensors were sampled at 50 Hz, which allows fine-scale reconstruction of whale behavior before, during, and after sonar transmissions. At a pre-set time, the suction was removed from the suction cups, releasing the tag from the whale which then floated to the surface. The tag contains a VHF transmitter used to track the tagged whale and to retrieve the tag after release.

Visual Tracking

Visual tracking of the whales' surfacing locations and observations of group behavior were conducted from the dedicated observation vessel, occasionally aided by visual spotters on the source vessel. On the observation vessel, the visual observer team typically consisted of four observers: (1) the focal animal tracker, (2) a behavioral observer, (3) a data recorder, and (4) a VHF tracker. Whale positions were determined using the distance from the vessel to the whale, the azimuth to the whale relative to the ship's heading, and the ship's magnetic or true heading. Distance was measured using laser-range finders or was estimated by eye when this was not possible. The relative bearing to the whale was measured using a protractor with a pointer. The observation boat's heading at the time of each sighting was measured with a Seagate fluxgate compass, a gyrocompass, or by course over ground measured with a GPS when speed over ground was above 3 kts. Visual observations were recorded using *Logger 2000* software.

Group Behavioral Observations

Group-level observations of behavior were made on an *ad-libitum* basis in 2006 and with a protocol of continuous, systematic collection of group-level observations in 2008 and 2009. Before tagging, the focal group was chosen by selecting a well-recognizable focal individual. The focal group was then defined by the individuals most closely associated with the focal individual, the smallest sample of individuals that formed a distinct unit with the focal individual, either forming a subgroup within a larger group or forming the total group. We attempted to target the focal group for tagging, with pre-tagging observations to enable analyses of the effects of tagging. Once a tag was attached, the tagged animal became the focal individual. Long-finned pilot whale groups, in particular, often formed several subgroups, repeatedly merging and splitting during an experiment. If the focal group split into several subgroups, the subgroup holding the focal individual became the focal group. In addition to data collection for the focal group, the number of subgroups and individuals in close vicinity of the focal group was systematically recorded by counting the number of subgroups and individuals within a 200-m radius of the focal individual.

Group-level behavioral parameters were defined quantitatively and consisted of low-level behavioral types (Miller et al., 2011). Parameters collected in 2008 were group size, calf presence, group spacing, number of subgroups and individuals in the focal area (all states), line swimming, and surface displays (events). In 2009, the protocol was extended with the additional parameters of surfacing synchrony, milling index, and distance to the nearest other subgroup (states).

Mitigation Procedure

Mitigation observations were conducted from both vessels. During transmissions, visual observers on the source vessel assured that no other whales were so close to the source that they might be exposed to sounds over 200 dB re: 1 μPa rms as required by the permit. A shutdown zone of 100 m was implemented, which added additional margin to this threshold as the range to 200 dB re: 1 μPa rms was always < 100 m. A mitigation observer on the observation vessel was tasked to order that sonar transmissions cease immediately if any cetacean approached 100 m from the source; showed any signs of pathological effects such as response to direct injury, disorientation, or severe behavioral reactions; or if any animals swam too close to the shore or entered confined areas that might limit escape routes.

Experimental Protocol

The protocol consisted of several phases: (1) searching, (2) tagging, (3) baseline pre-exposure data collection, (4) experimental exposure phase, (5) secondary exposure sessions, and (6) post-exposure data collection. The source and observation vessel teams searched for whales in the study area using towed array acoustics and visual observations. Once whales were located, and conditions were acceptable to attempt an experiment, the tag boat(s) were launched with tagging and photo-identification capability. During tagging attempts, the observer teams provided visual and acoustic tracking support to the tag boats, or they searched for new animals depending upon the situation.

Once a tag was attached, one tag boat followed the tagged whale to take identification photographs while the other continued attempts to tag a second animal. Second tagging was only conducted for killer whales and long-finned pilot whales and was always attempted on another animal within the same group in order to gain more information on group-level behavior such as dive synchrony, and to assure that one tag would remain attached long enough for the experimental protocol to be completed. Attempts to attach a second tag were stopped 1 h after the first tag was attached. Acoustic recordings from the observation vessel always started once a tag was attached. If more than one animal was tagged, a primary focal animal was specified based upon tag placement position, and each primary focal whale was visually monitored in a consistent fashion until all tags were detached. The observer team was kept blind to the experimental condition to the maximum extent possible, but it was not possible for observers to be blind to the approaching source vessel. Opportunistic sightings were made of other tagged animals, and specific sighting codes were used for each tagged animal. Tags were typically programmed to release after 15 to 18 h, enabling collection of 2 to 8 h pre-exposure baseline data, ~5 h for three sequentially presented exposure sessions (MFAS-UP, LFAS-UP, and Silent). The remaining 3 to 8 h of tag-attachment time was used to conduct playback of killer whale sounds, presentation of LFAS-DS signals, and recording of post-exposure data.

After a period of baseline pre-exposure data collection, the source vessel moved into position to start the first exposure run. The position of the tagged whale was estimated in real-time and relayed to the source vessel roughly every 5 min. The primary goal of the source vessel team prior to starting exposure was to place the source in a position about 6 to 7 km from the tagged animal in front of or to the side of the whales' direction of movement. During the planned 30-min transmission cycle, the source vessel approached the whales at a speed sufficient to move to a ~100 m range from the focal whale by the end of the 30-min period (7 to 8 kts), following a 10-min ramp-up period. Once the approach was started, the course of the source vessel was altered based upon updated positions of the tagged focal whale or bearings to clicking sperm whales determined from the Delphinus system. The course of the source vessel was fixed once it was 1 km from the focal whale, and it continued to transmit sonar pulses while moving in a straight direction for about 5 min after passing the tagged whale.

In November 2006, whales could only be approached once for a single sonar exposure session due to limited daylight. In 2008 and 2009, whales were approached for multiple exposure sessions, including Silent control approaches. The order of the sonar exposure and Silent control sessions was systematically altered for each experiment within each species (Table 2). The exposure schedule (see Table III in Miller et al., 2011) also reflected our consideration that Silent control passes, while important, would only be useful once a minimum number of experimental exposures have been conducted. Exposure sessions were planned to start at least 1 h following the end of the previous session once the source vessel was in a new acceptable location.

Playback of killer whale sounds and secondary exposure sessions with LFAS-DN signals were conducted after the controlled exposure experiments (CEEs) of sonar signals. For killer whale sound playbacks, a dedicated playback boat was deployed from the observation vessel. For long-finned pilot whales and killer whales, sounds of herring-feeding killer whales recorded within the study area were played, and the sound source was positioned to the side of the tagged whale's path, at a distance of 2,400 ± 943 m. For sperm whales, sounds of mammal-feeding killer whales in Alaska were played. The source was positioned 500 to 100 m ahead of the whale, and playback started just after the whale dove. Different specific sound sequences from the relevant killer whale ecotype were used among the tested whales. Each playback lasted 15 min and was broadcast at a depth of 8 m. The source level of the stimuli was monitored 1 m from the source and was set to 140 to 155 dB_{rms} re: 1 μPa m, which corresponds to the source level of killer whale vocalizations observed in natural conditions (Miller, 2006). Post-exposure data collection continued after the final stimulus presentations. The team on the observation vessel continued to track and observe the tagged whale and its group until the tag detached from the whale.

Behavioral Data Processing and Analysis

Upon recovery of the tags, pressure data were converted to depth using calibration values compensating for temperature effects. Similarly, the accelerometer and magnetometer outputs were converted to field strength on each axis (Johnson & Tyack, 2003). Pitch, roll, and heading of the whale were calculated following published methods (Johnson & Tyack, 2003; Miller et al., 2004b).

For all long-finned pilot whale and most killer whale experiments, a detailed track of the tagged animals could be made from the visual tracking data. A dead-reckoned track was produced for all sperm whale experiments as well as for those experiments with killer whales in which sightings were made less frequently. A constant speed dead-reckoned track was first calculated for the periods between consecutive sightings (Miller et al., 2009). Dead-reckoning started at the first sighting using an average speed calculated by the ratio of the distance and the difference in time between sightings. The dead-reckoned track points were then modified by adding a vector whose magnitude and angle would make the point corresponding to the next surfacing match the sighting position at the next surfacing. Points leading up to the next surfacing were adjusted by interpolating the magnitude of the correction vector linearly against time, from zero correction at the previous surfacing to the full correction vector for the next surfacing and thereafter. This process was repeated sequentially from the first to the last sighting, resulting in a dead-reckoned track that matched the locations determined from sightings with interpolated positions in between the sightings. For sperm whales, corrections were made by only using sightings of the location of the whale when it raised its flukes prior to diving as those were the highest quality sightings for each surfacing sequence of sperm whales, and speed was set to 0 ms^{-1} when the whale was at the surface. The applied correction vectors tended to be fairly small, and in a consistent direction indicating that the deviation between the dead-reckoned and sighting tracks arises due to water currents or variations in the actual speed of the whale.

Speed and direction of movement of the tagged whale were calculated from the horizontal location obtained from sightings at the surface or from the corrected dead-reckoned track. Speed was calculated using the surfacing locations just prior to and after each surfacing point. Speed was calculated as the total great circle distance travelled over the three surfacings divided by the total time between them. Direction of motion of the whale was calculated as the true bearing from the previous surfacing.

Sound files of biological sounds recorded by the Dtags were viewed as spectrograms with a 4096 point Blackman-Harris window for a time resolution of 21.3 ms. The onset time and duration of each sound produced by the tagged species was recorded. Each sound was assigned to a type using a letter code and to an amplitude index (see Table IV in Miller et al., 2011). The amplitude of biological sounds was classified as 1 for faint, barely detectable sounds; 3 for sounds that were loud and clear in the recording and spectrogram; and 2 for sounds that were intermediate between those two extremes. Vocalization types were defined and coded. The code SS (social sounds) sound type was considered to include any sound produced by the animals that resulted in horizontal bands in the spectrogram. SS were considered as part of the same event if they were separated by less than 0.2 s. The specific code SSMIMIC was given to any SS with frequency modulation similar to sonar pulses that occurred within 5 min after the last sonar transmission. Click sounds (CS) were considered the same sequence if clicks were separated by < 2 s. This rule was applied in the same way to rapid click series or buzzes (BUZZ), and linked sequences of clicks and buzzes (BUZZCS).

Processing and Analysis of Received Sonar Signals

Sonar signals received on the tag or hydrophone arrays were processed by means of a custom *Matlab*, Version 7.5 (The Mathworks, Natick, MA, USA) program, including a strict analysis protocol to control for noise from animal movements, vocal behavior, echolocation, and ship propulsion. A flip-template matched filter (Burdic, 1991) identified the start of a ping's first arrival, and a time cue for the first arrival was stored. Waveform and spectrogram views of the signal guided every step of the analysis. After signal inspection, a 200-ms window of stationary noise preceding the ping was marked. Where needed, echolocation clicks were removed for the estimation of ping levels. Following the recommendations of Southall et al. (2007), we quantified the sonar signals in terms of sound pressure level (SPL) and sound exposure level (SEL). Level definitions are based on the terminology in Morfey (2001) and Southall et al. (2007).

To account for the time-varying pressure envelope of the sonar signal, we report the maximum SPL (SPL_{max}; dB re: 1 μPa, rms); the highest value of SPL that occurs during a specified time interval after a running average is performed on the instantaneous or mean-square pressures. We used two sliding windows with averaging times of 10 and 200 ms, which resulted in two time-weighted SPLs—SPL_{10} and SPL_{200}, respectively. The maximum of the latter is reported as SPL_{max}. For the frequencies of interest here (1 to 2 and 6 to 7 kHz), the mammalian ear integrates sound intensity over a time window of 100 to 400 ms duration for perception of loudness (Plomp & Bouman, 1959; Fay, 1988). Comparable integration time constants are reported for the bottlenose dolphin (*Tursiops truncatus*; 1 to 4 kHz, ~200 ms; Johnson, 1968), and the harbour porpoise (*Phocoena phocoena*; 1 to 8 kHz, ~200 to 600 ms; Kastelein et al., 2010), which suggest that SPL_{200} is likely to be a relevant quantity in terms of sensation of loudness for odontocete cetaceans.

Signal duration τ_{20dB} is defined as the time during which the SPL exceeds a 20 dB threshold below the highest SPL_{10}. The first crossing with increasing SPL and the last crossing with decreasing SPL were selected that occurred over a 10-s period starting from the first-arrival time cue. For a few low source level pings, the signal never exceeded the background noise by 20 dB, so the start and end of the ping were selected manually by visual inspection of the spectrogram. For analysis windows that were partially overlapped by noise transients, an alternative start-time cue was used. This cue was taken as the start point, and the 10 s window was shortened to prevent the noise from influencing the duration measure. The final values reported for SPL_{max} and SEL (see next paragraph) were computed using τ_{20dB} as integration time T (Equation 1).

SEL (dB re: 1 μPa^2-s) is defined as the level of the cumulative sum-of-square pressures. As it integrates acoustic exposures over time, the SEL metric is also very useful for assessing cumulative exposure to intermittent sonar signals:

$$\mathrm{SEL} = 10\log_{10}\left(\frac{\sum_{n=1}^{N}\int_{0}^{T} p_n^2(t)dt}{p_{ref}^2 t_{ref}}\right) \qquad (1)$$

where N is the number of transmitted pings, T is the sonar signal ping duration (in s), and $p_n^2(t)$ is the square pressure of the n^{th} transmission as function of time (in μPa^2). Reference pressure p_{ref}^2 and reference time t_{ref} are 1 μPa^2 and 1 s, respectively. The single-ping SEL (N=1) and the cumulative SEL (SEL_{cum}) (n = all pings up to that point) per exposure session were calculated. As a consequence of the click removal procedure, SELs were computed by cumulative summation of the mean square pressures $p_{rms}^2(t)$. To eliminate the influence of background noise on the exposure levels, the mean square pressure of the noise segment preceding the ping was subtracted from $p_{rms}^2(t)$ before each SEL was calculated.

If reverberation had not completely vanished after 20 s, the signal-to-noise ratio (SNR) (defined here as the difference between the SPL_{max} and the SPL of the 200-ms noise segment preceding the ping) was in fact a signal-to-reverberation ratio. The lowest observed signal-to-reverberation ratio was about 40 dB, thus the noise subtraction procedure did not significantly influence the level of such pings.

Sometimes a signal could not be measured but was still likely to have been received by the animal. A ping was assigned as "received at full level" when a tagged sperm whale rested at the surface, or when pilot or killer whale vocalisations or splashing water sounds coincided with the signal. It is possible that some animals may use their surfacing to reduce sound exposure. A ping was scored as "not received at full level" by the animal when a tag on a killer or long-finned pilot whale was completely out of the water over the full duration of the signal. For pings that could not be measured but were assigned as "received at full level," single-ping levels were estimated from the adjacent ping levels by linear interpolation on the decibel scale, and the SPL_{cum} over the experiment was recalculated including the level of that ping. To estimate the received level in the beginning of the ramp-up period when the source vessel was quite distant from the whale, the first measured ping level was extrapolated, and levels were corrected for differences in source level. This approach was taken because one group of animals (Oo09_144ab) appeared to respond to the sonar before any ping could be measured.

Scoring Panel Methods

Miller et al. (2011) describe each experiment with a complete series of data plots. Based upon the information within this 3S-technical report, behavioral changes that could putatively have been responses to the experimental exposure were described and scored by two independent groups in accordance with the Southall et al. (2007) severity scale (Figure 1). One group consisted of authors PM, AA, RA, and PW and the second of authors PK, FL, LS, LK, and FV. The two groups each conducted an initial scoring, blind to each other's scoring but not blind to the type/presence of acoustic stimulus. Thereafter, the results of the two groups were tabulated before they met with an adjudicator (PT) to reach a consensus scoring (Figure 1). The consensus meeting started with a discussion about how to specifically apply the severity scale. Most differences between the groups were due to a difference in interpretation(s) and use of the severity scale, and not disagreements about what type of behavioral changes actually occurred or whether or not an observed behavioral change was a potential response to the experimental exposure. The judgment of whether behavioral changes were scored as responses took into consideration the variation in parameters recorded during the baseline period.

Two types of data plots were considered by the panels for each exposure session: a geographic track plot (GIS plot) and a time series plot. The GIS plot (e.g., Figures 2, 4 & 6) showed sighting locations of each tagged whale, GPS-tracked locations of the source vessel or location of the source boat used for killer whale playbacks, and the GPS-tracked location of the observation boat. The time series plot (e.g., Figures 3, 5 & 7) showed (1) a subset of the behavioral observation record, plotted as the raw values of scores of the behavior at the time they were recorded for group spacing, surfacing synchrony, and breaching and tailslap events; (2) the horizontal speed of the whale calculated from movement track; (3) the direction of movement of the tracked whale; (4) the received level of the sonar transmissions (SPL_{max} and SEL_{cum}); and (5) the time-depth profile of the tagged animal(s). Colors or symbols indicating acoustic clicking, buzzing, and tailslap sounds were overlaid on the dive profile, while the timing of social sounds were indicated by a symbol above the dive profile.

Some minor changes were made to the original Southall et al. (2007) severity scale to deal with additional capabilities in our data. Southall et al. reviewed all relevant scientific literature on responses of marine mammals to sound. They then developed a scale to score the severity of behavioral reaction based on a very broad range of observations, many of which used methods with limited temporal resolution. By contrast, our experiments studied and analyzed changes in behavior using methodology that gave a very high resolution of the acoustic exposure and scored responses both temporally and spatially. We therefore added the duration category *brief*, defined as substantially shorter than the exposure duration and also a shorter duration than the *minor* category. In addition, two behavioral responses that

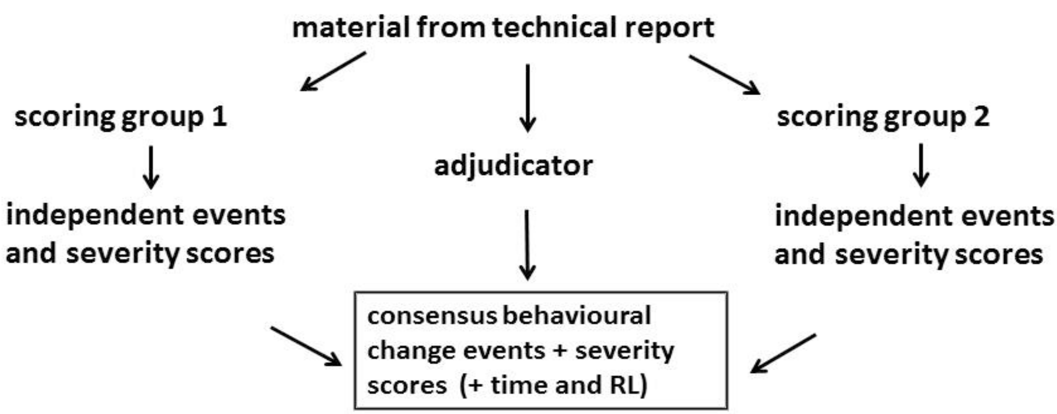

Figure 1. The structure of the severity scoring procedure

we scored were not listed on the original Southall et al. severity scale:

1. *Cessation of Feeding* – Cessation of feeding was considered to have the same severity as cessation of reproductive effort, and less severe than separation of dependent offspring. As for all scored responses, we distinguished between *moderate* (severity of 6) and *prolonged* (severity of 7) cessation of feeding.

2. *Cessation of Resting* – Resting is an important life function, and disruption of resting was considered to be as equivalently severe as cessation of feeding. We only scored cessation of resting to have occurred when stereotypical resting behavior was interrupted (Miller et al., 2008). Moderate cessation of resting behavior scored a severity of 6.

After these revisions were made to the scoring criteria (Table 1), we reviewed the scoring of each experiment. When the two groups disagreed in their independent scoring, we revisited the original data plots and came to an agreement on the type of response and the corresponding severity score. The final consensus on all severity scores was agreed to by both independent scoring panels

Table 1. The Southall et al. (2007) severity scale with modifications used in this study; brief < minor < moderate < prolonged in terms of both duration and severity of the behavioral change, where moderate has a duration roughly equal to the duration of the exposure.

Score	Behavioral responses	
0	• No observable response	
1	• Brief orientation response	
2	• Moderate or multiple orientation responses • *Brief or minor changes in respiration rates*	• Brief cessation/modification of vocal behavior
3	• *Prolonged orientation behavior* • Minor change in locomotion (speed/direction) and or dive profile but no avoidance of sound source • Minor cessation/modification of vocal behavior	• *Individual alert behavior* • *Moderate change of respiration rate* • **Brief change in group distribution**
4	• Moderate change in locomotion (speed/direction) and or dive profile but no avoidance of sound source • **Brief avoidance of sound source**	• Minor shift in group distribution • Moderate cessation/modification of vocal behavior
5	• Extended change in locomotion (speed/direction) and or dive profile but no avoidance of sound source • **Minor avoidance of sound source**	• Moderate shift in group distribution • *Change in inter-animal distance and/or group size* • Prolonged cessation or modification of vocal behavior
6	• Moderate avoidance of sound source • Extended cessation or modification of vocal behavior • Visible startle response • **Moderate cessation of feeding** • Prolonged shift in group distribution	• Brief or minor separation of female and dependent offspring • *Aggressive behavior related to noise exposure* • *Brief cessation of reproductive behavior* • **Moderate cessation of resting behavior**
7	• **Prolonged cessation of feeding** • Moderate separation of female and dependent offspring • Severe and or sustained avoidance of sound source	• *Extensive or prolonged aggressive behavior* • *Clear anti-predator response* • *Moderate cessation of reproductive behavior*
8	• *Obvious aversion and/or progressive sensitization* • *Long-term avoidance of area*	• Prolonged or significant separation of female and dependent offspring with disruption of acoustic reunion mechanisms • *Prolonged cessation of reproductive behavior*
9	• *Outright panic, flight, stampede, attack of conspecifics, or stranding event*	• *Avoidance related to predator detection*

Original responses as described and scored in the Southall et al. (2007) table. Behavioral changes in *italic* type were described in the Southall et al. table but were never scored in this study. Behavioral changes in **bold** type were amendments or modifications made to the scale.

and the adjudicator. The data were then inspected in detail to determine the precise onset time of the behavioral change. For sonar exposures, the SPL_{max} received up to the onset of the behavioral change was used as the received level associated with the scored response. We also report the SEL_{cum} at the start of each scored response (Appendix Tables AI-AIII). Proximity to the source during the sonar exposure sessions is available in Miller et al. (2011).

Statistical Analysis

For statistical comparison of the maximum severity of scored responses identified during Silent vessel vs sonar exposure sessions, the proportion of sessions with a maximum severity of 4 or greater was calculated. This divides sessions into those in which no risk to vital rates was indicated (maximum severity score of 0 to 3) vs sessions in which possible (severity of 4 to 6) or likely effects (severity of 7 to 9) on vital rates are expected (Southall et al., 2007). A G-test was used to test the hypothesis that the proportion of sessions with a maximum severity of 4 or greater differed between Silent vessel approaches and sonar approaches (Zar, 1984). LFAS-UP and LFAS-DN were pooled into a common 1 to 2 kHz frequency band.

To investigate potential influences of order of exposure, frequency band of sonar, and/or species on maximum severity score and number of responses scored in sonar sessions, a Generalised Estimating Equations (GEE) Model for Poisson data was built using order of exposure, frequency band, and species as explanatory variables, the last two being categorical variables. Individual identification was used as a clustering factor to take into account the fact that some observations in the data came from the same individuals. Given the low number of clusters in the data (N = 14), the Jackknife variance estimator was used. All statistical analyses were implemented using the function geeglm in the geepack package (Yan, 2002; Yan & Fine, 2004; Højsgaard et al., 2006) in *R* (Version 2.14.1; R Development Core Team, 2008) for Mac OS X.

Results

In 2006 through 2009, 14 unique experiments were conducted with killer whales (n = 4), long-finned pilot whales (n = 6), and sperm whales (n = 4). Multiple exposure sessions were conducted with most focal groups for a total of 14 6-7 kHz upsweep (MFAS-UP) exposure sessions, 13 1-2 kHz upsweep (LFAS-UP) exposure sessions, and five 1-2 kHz downsweep (LFAS-DN) sonar exposure sessions, as well as seven Silent vessel control exposure sessions and eight killer whale-playback exposure sessions (Table 2).

Table 2. Listing of all experiments conducted with the three species; the code is the Dtag code used for the deployment. The numbers indicate the order each sound type was presented during the experiment (MFAS: 6 to 7 kHz upsweep; LFAS: 1 to 2 kHz upsweep; Silent: source vessel approach but no sonar transmissions; LF-ds: 1 to 2 kHz downsweep; KW: natural killer whale sounds).

Species	Year	Code(s)	MFAS	LFAS	Silent	LF-ds	KW	Comments
Orcinus orca	2006	Oo06_317s		1				
	2006	Oo06_327s	1					
		Oo06_327t						
	2008	Oo08_149a	1 & 5	2	3		4	Narrow fjord
	2009	Oo09_144a	2	1		4	3	
		Oo09_144b						
Globicephala	2008	Gm08_150c	1	2				
melas	2008	Gm08_154c	2	1				Difficult tracking
		Gm08_154d						
	2008	Gm08_158b	3	2	1			Dtag data lost
	2008	Gm08_159a	3	2	1		4	
	2009	Gm09_138a	2	1	3	4	5	
		Gm09_138b						
	2009	Gm09_156b	3	2	1	4	5	Narrow fjord
Physeter	2008	Sw08_152a	1	2				Difficult tracking
macrocephalus	2009	Sw09_141a	2	1	3		4	
	2009	Sw09_142a	3	2	1	5	4	
	2009	Sw09_160a	1	2		4	3	

Examples of Severity Scoring of Experimental Outcomes

In this first part of the results, we detail one typical example for each of our three study species using the method by which we identified behavioral changes that were scored as responses to the sonar.

Example 1: *Killer whales Oo06_327s and Oo06_327t (Table 2)* – The description from Miller et al. (2011) in the next paragraph is linked to Figures 2 & 3.

> **Description:** *Before tagging, we followed a small group of 4 killer whales that moved quickly in a westerly direction in the Vestfjord basin to join a large group of carousel-feeding killer whales, estimated to be 70-80 animals of all age and sex classes. Two individuals were tagged within the carousel-feeding group, and we are not sure whether they were in the initial group we followed or were from the original large carousel-feeding group. Subject Oo06_327s was an adult female traveling with a small animal, and Oo06_327t was an adult male. There was a clear change in behavior at the end of a synchronous deep dive (13:56:14-13:56:25) during MFAS exposure. There were numerous indications of feeding in the early part of the record, including numerous tailslaps during the dive made just before the change in behavior. Before the change point, the whales were moving at low speed with a highly circular travel path with stunned herring observed at the surface, indicating that they were feeding on herring. At the change point, all indications of feeding by the whales (tail slaps and echolocation) stopped, and the animals' movement path became highly directional, with an increase in speed. This movement continued and the animals moved in the direction leading out of Vestfjord.*

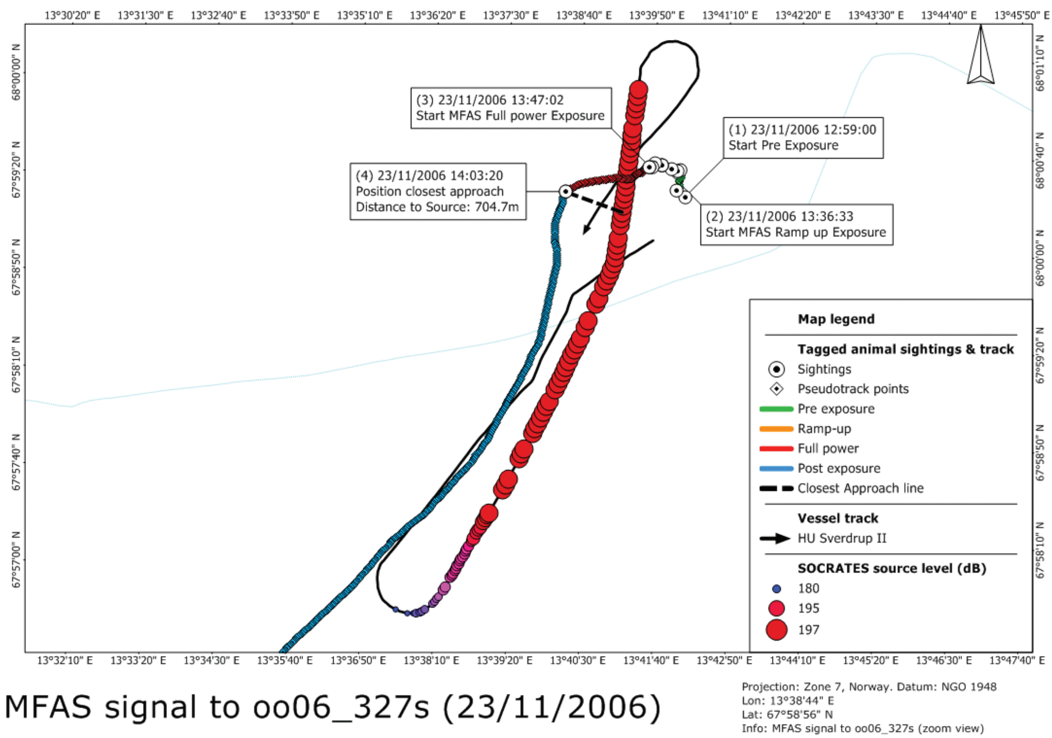

Figure 2. Horizontal track of the source vessel (thick black line) and tagged killer whale Oo06_327s before, during, and after exposure to MFAS (6 to 7 kHz) sonar signals; individual sightings of the whale are plotted as white circles with a black dot, and the lines between the sightings are derived from the dead-reckoned track (labelled "pseudotrack") of the tagged whale. Color of the track points indicates the exposure condition while the whale was at that location. The colored circles overlaid upon the source vessel track indicate the source level of sonar transmissions made in those locations. Notes are added to each plot to show locations at the point of closest approach determined from the visual sighting record and locations of the whale when the experimental condition changed.

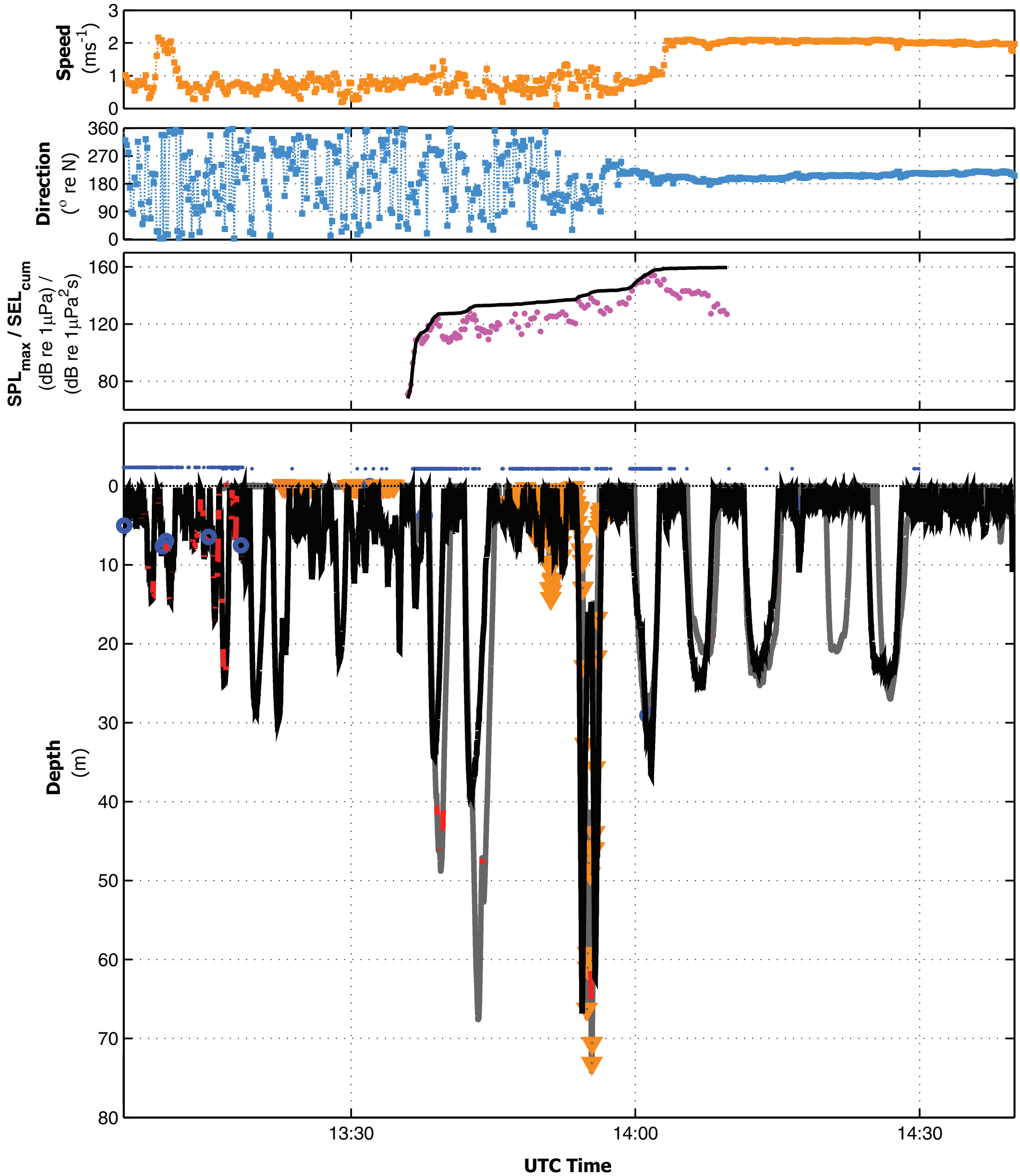

Figure 3. Time-series data plot for the MFAS exposure to killer whales Oo06_327s and Oo06_327t. Two animals were tagged, and the primary focal individual was Oo06_327s, whose dive profile is plotted in black (bottom panel) and which was the animal tracked to derive horizontal speed (top panel) and direction of movement (2nd panel). The received level of the eight MFAS sonar signals is shown in the 3rd panel. Social sounds recorded on that tag are plotted as nine blue dots above the dive profile in the bottom panel, while feeding-related sounds are plotted directly onto the dive profile. Echolocation click series are plotted in red, buzz sounds in blue, and tailslap sounds in orange. The dive profile for the 2nd tagged animal, Oo06_327t, is plotted in grey. Note that no systematic social behavior observations were collected during this experiment.

Sightings of the focal animals were made difficult by darkness at the end of the day, but we [were] able to follow the group based upon the VHF signals transmitted by the tag. The tags were recovered 28 and 30 km away, 4 to 5 hours later. Few calls were recorded during the dive before the change in behavior. Calling was recorded on the tag as the animals were moving away from the feeding location, and then stopped for several hours. Received levels of the sonar increased during dives compared to surface intervals. Sound velocity changed linearly with depth because water temperature was constant throughout the water column. No data were collected on group-level behavior.

Severity Scoring – The consensus responses scored for this session were as follows:

Change #	Severity	Description
(1)	3	Minor change in dive profile. *Time:* Started at inflection point within the final foraging dive at 1355:14 h. *Dose:* SPL_{max}: 139 dB, SEL_{cum}: 141 dB
(2)	7	Prolonged avoidance; tagged subject with its group moved sideways to oncoming source, then away from source for several hours. *Time:* Started at end of dive with inflection point at 1356:14 h and 1356:25 h for Oo06_327s and Oo06_327t, respectively. *Dose:* SPL_{max}: 139 dB, SEL_{cum}: 143 dB
(3)	7	Prolonged cessation of feeding. Tailslaps were heard in the final dive and not heard again in the tag record. The animals may have returned to feeding sometime after the end of exposure—calling was heard but not tailslaps. *Time:* Started at final tailslap at 1356:33 h. *Dose:* SPL_{max}: 139 dB, SEL_{cum}: 144 dB
(4)	5	Prolonged cessation of vocal behavior. *Time:* Started after set of four calls at 1404:08 h. *Dose:* SPL_{max}: 154 dB, SEL_{cum}: 159 dB

This example shows a typical strong change in behavior during an exposure period, which was considered likely to have been a response to the sonar. It is also likely that the four different behavioral changes identified in our records depend upon, or are related to, one another. For example, movement away from an isolated prey patch such as a school of herring, as could have been the case here, must necessitate cessation of foraging. Similarly, a change in acoustic behavior involving cessation of sounds related to foraging is likely related to cessation of foraging (Simon et al., 2005). The unusual dive profile (Change 1) occurred during the final dive of attempted foraging. This type of outcome was common in our study: a suite of changes were identified that were likely different components of a change in behavioral state, in this case from feeding to travelling, which appears to represent a complex behavioral reaction to the sonar exposure.

The avoidance behavior, cessation of foraging, and cessation of calling in this case were prolonged substantially beyond the duration of the sound exposure period. While some calls were observed near the end of the tag record starting at 1900 h UTC the same day (~5 h after exposure had ended; Miller et al., 2011), no tailslap sounds, circling movements, or other clear indications of feeding were recorded by the tags. Thus, in this case, the duration or full extent of either the avoidance or the cessation of feeding cannot be specified; however, they clearly were prolonged beyond the duration of the exposure, which indicates a more severe response than a behavioral change that reverted to normal around the time that the sound exposure ended.

Example 2: *Long-finned pilot whale Gm09_156b. 1 to 2 kHz downsweep signals (LFAS-DN; Table 2)* – The description below includes the baseline behavior period and during the exposure session, but it does not include the descriptions for the Silent, 1 to 2 kHz upsweep or 6 to 7 kHz exposure sessions that preceded it nor for the killer whale sound playback exposure session that followed it. Data series for the LFAS-DN exposure session are given in Figures 4 & 5.

Description: *A large group of 80 to 100 pilot whales with calves was sighted from the* HU Sverdrup II. *Both tag boats were deployed following a pre-tagging observation period. The whales were moving NE into Vestfjord during most of the tagging phase. Tag Gm09_156b was deployed at 1712 h (06/06/2009) on a large male using the pole system. After tagging, the whales moved into Ofotfjord where all the exposures took place. Another tag was deployed previously using* *the ARTS system, which only remained attached for 12 min.*

While the tag boats were in the water, the focal group alternated between milling and lining up at the surface moving at low speed. As the tagboats left the whales, they increased their speed and headed NE swimming fast, showing very tight group spacing and line-up swimming, with no records of surface display events except for one tailslap and one breach. Following tag deployment, the tagged whale did not make any deep-dives until 2051 h when it made a single dive to 171 m. The focal whale then kept shallow diving and moving at high speed for another 40 min until it started deep-diving (2141:50 h). During these deep dives, the whale made clicks and buzzes indicating feeding as well as social sounds. This change was also visible at the surface by increased group spacing and no further events of lined-up swimming.

The focal whales were subsequently exposed to LFAS downsweeps, 78 min after the MFAS exposure (see Miller et al., 2011, for timing of the full experiment). Following the first deep-dive at the start of the MFAS exposure, the tagged whale made several other deep dives until the start of the LFAS downsweep exposure. These were not consecutive, but they were interspersed with periods of shallow diving (8 to 25 min long). The whale kept this diving pattern during this exposure while making clicks and buzzes. During this period between the MFAS and LFAS downsweep exposures, the tagged whale was observed [alone] at the surface several times until rejoining the closely spaced focal group prior to the LFAS downsweep exposure. The source vessel Sverdrup *approached from the E (heading W) as the whales were heading NE. At 0509 h (150 dB RL, 1,345 m from* Sverdrup*), the tagged whale made a sharp turn to SW and initiated a long dive, resurfacing W of its last sighting. This was the closest surfacing to the source vessel during this exposure (170 dB RL, 154 m from* Sverdrup*). The focal whale then turned NNE, passing behind the source vessel. Group spacing remained tight to very tight throughout and following the exposure, and surfacing synchrony increased soon after the end of the exposure during which several spyhops were observed. Also at this time, several groups of whales merged. Milling and logging were observed throughout the exposure and post-exposure periods. The tagged whale*

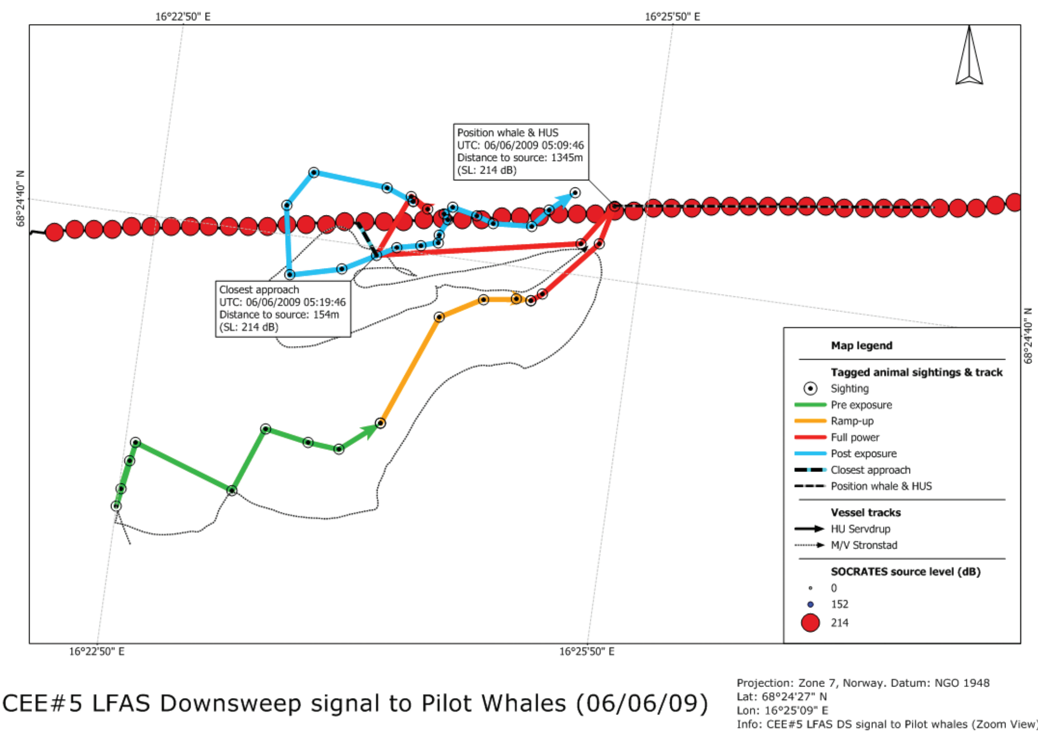

Figure 4. Horizontal track of the source vessel (thick black line) and tagged long-finned pilot whale Gm09_156b before, during, and after the LFAS-DS (1 to 2 kHz downsweep) exposure; individual sightings of the whale are plotted as white circles with a black dot, and the lines connecting the sightings are colored according to the exposure condition. Source vessel moved from right to left.

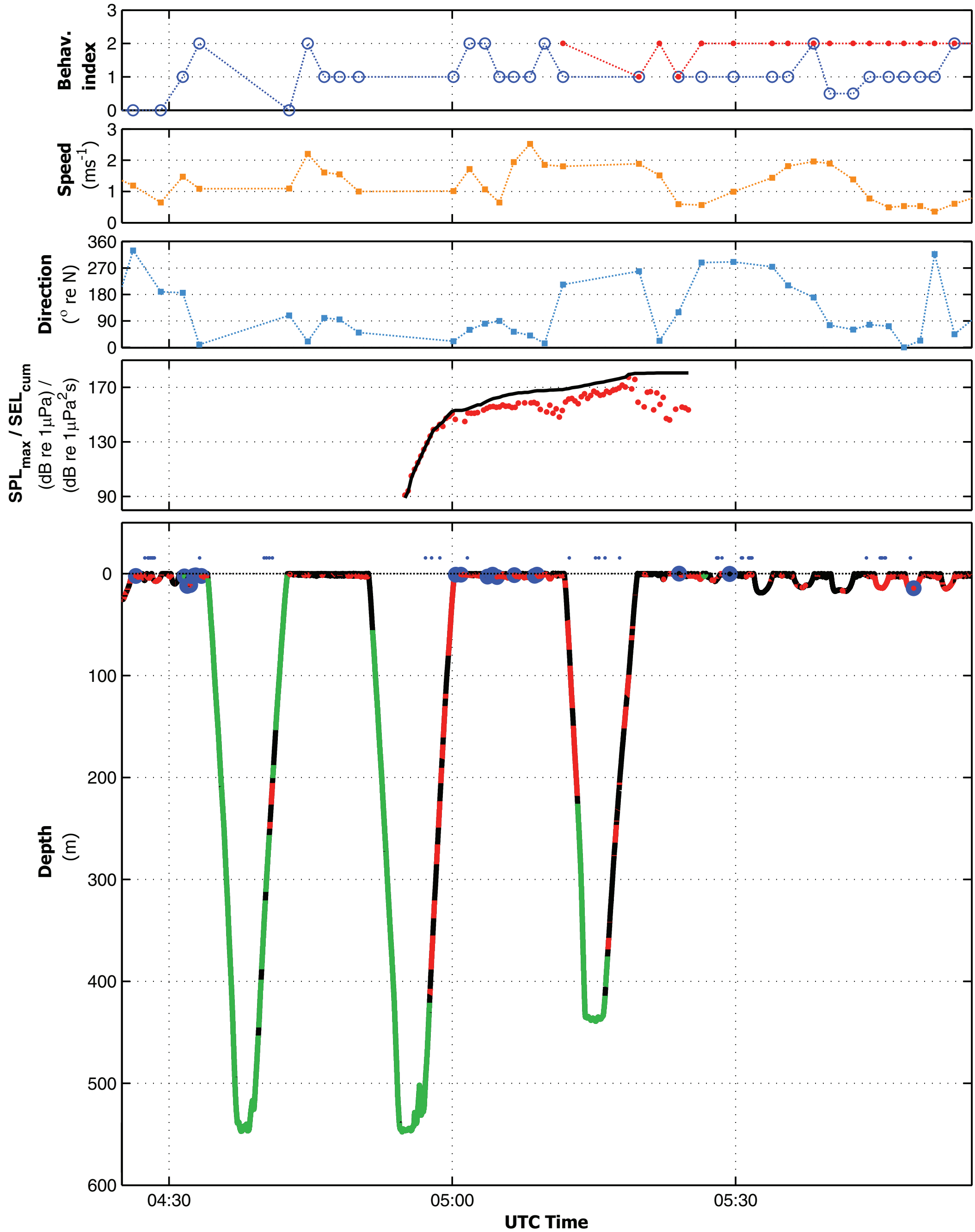

Figure 5. Data plot for the LFAS-DS exposure to long-finned pilot whale Gm09_156b; the top panel encodes group spacing (blue circles; low index means tight spacing) and group synchrony (red dots; low index means high synchrony). Low behavioral indices indicate tight group spacing. See Figure 3 for description of other panels. Green coloring of the dive profile trace indicates a combination of clicks and buzzes during that interval.

logged for 1.5 min during the very last pings of the exposure, but after the source vessel had passed. After the end of the exposure, the animals looped back along the source boat's track by heading W, then S, and then E.

Severity Scoring – The consensus response scored for this exposure session was as follows:

Change #	Severity	Description
(1)	6	Moderate avoidance. Ship approached from the front. Animals moved sideways from the ship and then 180° from ship, similar to pattern in LFAS exposure session. RL continued to increase because the source was moving faster than them. The 2nd avoidance movement was judged to be a continuation of the 1st one. *Time:* Taken just before 1st overlap surfacing at 0511:43 h. *Dose:* SPL_{max}: 159 dB, SEL_{cum}: 168 dB

In this example, the change in travel direction away from the path of the oncoming source is clear, but the precise form of the avoidance behavior may have been shaped or limited due to the relatively narrow geometry of Ofotfjord. While the logging of the tagged whale was a short-term change in behavior, it was not scored to have been a response to the sonar because there had been several similar sequences of logging during the baseline data period. Other variations in behavior patterns, such as diving, were judged to be within the variation observed during baseline behavior.

Example 3: *Sperm whale Sw09_142a Silent vessel approach (Table 2)* – Here, the description includes the baseline behavior period and the Silent vessel approach, which was the first exposure session conducted with this whale subject. Data series are shown in Figures 6 & 7.

Description: *This sperm whale was a large solitary male that was found close to the area in which we encountered Sw09_141a. He was not seen associating with other animals but did produce a coda upon tag attachment. During pre-exposure, the tagged animal made five foraging dives to 100 m with clicks and buzzes. The animal moved steadily E during the baseline period.*

During the Silent pass, the animal made another dive to 300 m with clicks and one buzz during ascent in the 100-m depth range. The closest approach was about 150 m. The tagged whale turned around before the closest point of approach, which brought the whale closer to the path of the source. The turn was after the period with no buzzes, but before the buzz itself. The whale then returned to its previous course after surfacing.

Severity Scoring – The consensus response scored for this exposure session was as follows:

Change #	Severity	Description
(1)	1	A brief turn towards source vessel at CPA, which we interpreted as an orientation response. *Time:* Taken at point where direction changed towards ship 2016:51 h.

In this example, there are several changes in movement behavior throughout the exposure period, but all are considered to be within the level of variation in movement observed during the baseline period (Miller et al., 2011). The rapid, short-duration turn from 90 to 270° (Figures 6 & 7) was judged to have been a potential orientation response to the approaching vessel. The duration was much less than the exposure period, so it was considered to be brief with an associated severity score of 1. While this change in movement direction also could have represented simply normal, undisturbed foraging behavior (Miller et al., 2004a), its directionality towards the vessel and simultaneous cessation of clicking justifies its scoring as a response to the vessel approach. Variation in the dive depth of the whale during the Silent vessel approach was judged to be within the normal range of variation in normal undisturbed dives of this whale (see baseline and post-exposure data plots in Miller et al., 2011, and other sperm whales in the study area [Teloni et al., 2008]).

Summary of Scored Responses to Exposure

A total of 84 responses were scored in the 47 different exposure sessions conducted across the 14 subject whale groups (Figure 8; Appendix Tables AI-AIII). The total number of responses scored per session ranged from zero changes in eight sessions to five changes in four sessions. One change per session was the most common (17 sessions), while two changes were observed in 11 sessions, three changes in three sessions, and four changes in four different sessions. When more than one response was scored during a session, they tended to be different manifestations of a broader change of functional behavioral state (see Oo06_327s in Example 1 above).

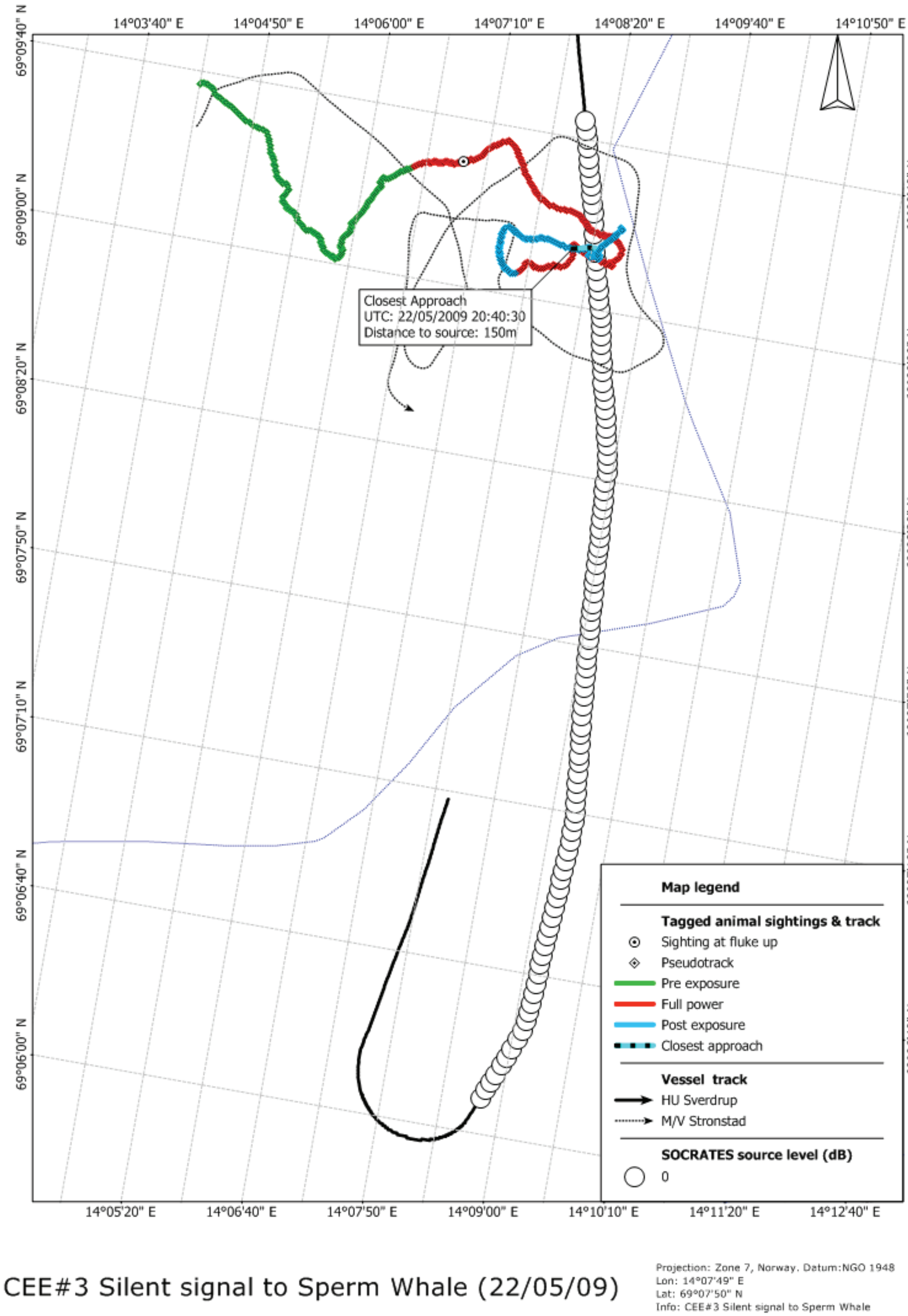

Figure 6. Horizontal track of the source vessel (thick black line) and tagged sperm whale Sw09_142a before, during, and after a Silent exposure; colored dead-reckoning track indicates the exposure condition. Empty circles on the source vessel track indicate mock Silent transmission times.

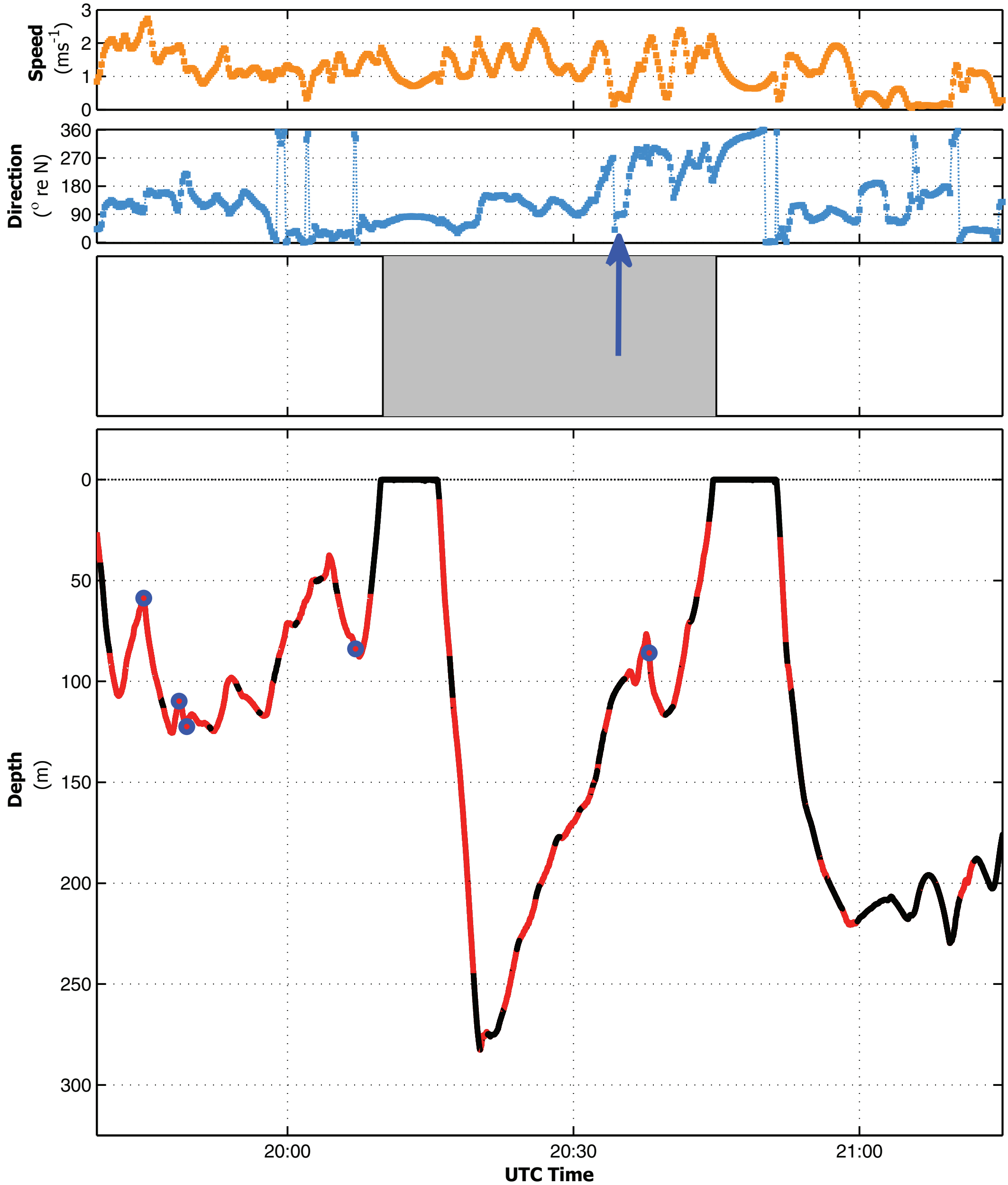

Figure 7. Data plot for the Silent exposure to sperm whale Sw09_142a; no group behavior was collected for this solitary animal. The period of the Silent exposure is marked in grey in the 3rd panel. See Figure 3 for description of other panels. The blue arrow indicates a short-duration (brief) turn toward the vessel at the time of the closest point of approach of the vessel, which was scored as a brief orientation response.

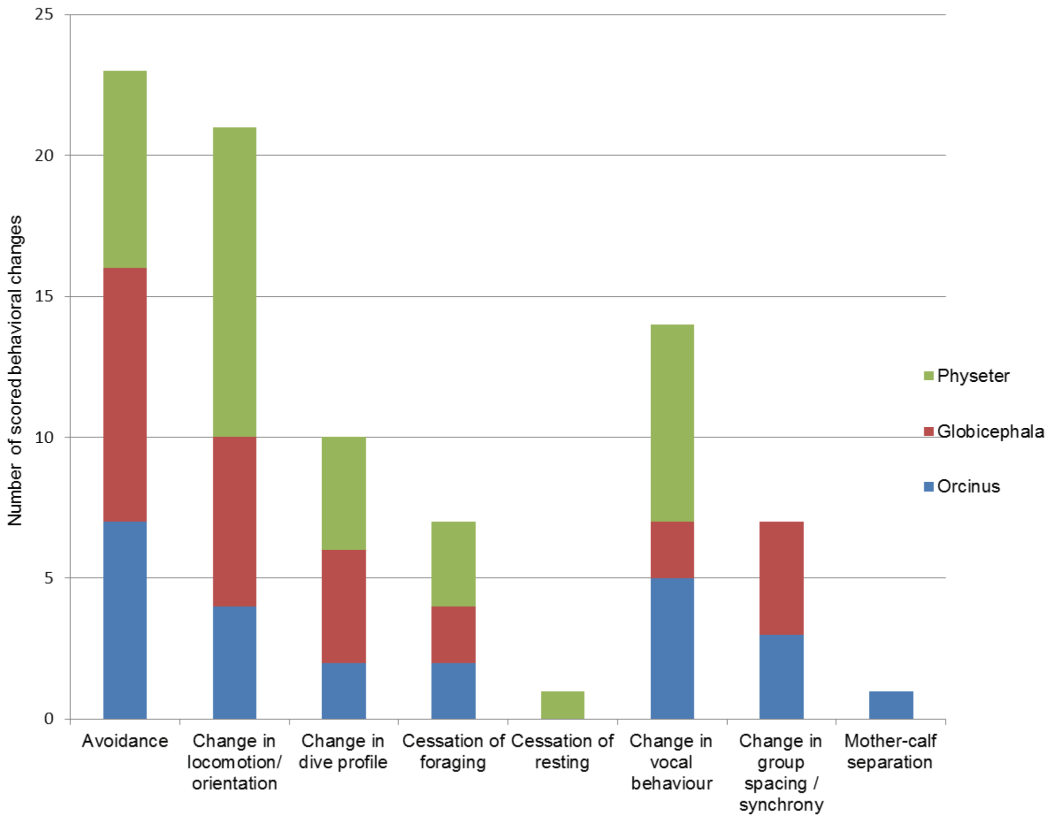

Figure 8. The types of behavioral responses scored in 47 exposure sessions in 14 unique experiments with the three target species.

The most common response scored (23 of the 84 identified changes) was *avoidance* of the source (Figure 8), including horizontal movement away from the source vessel or the path of the moving source vessel, and two cases of vertical avoidance by long-finned pilot whales surfacing in synchrony with the arrival of sonar pings (Gm08_159a_2 and Gm09_156b_6). Avoidance behaviors occurred quite evenly across the three species studied. Many avoidance events coincided with other scored responses, including changes in diving (e.g., Sw09_142a_21) and vocal behavior (e.g., Gm08_150c_1 and Oo09_144a_1).

Short-duration changes in horizontal movement that were not away from the vessel or anticipated vessel track were not scored as avoidance. However, these changes and brief dives or inflections in dive profiles were judged to be orientation movements (20 changes). In two playback sessions of killer whale sounds, long-finned pilot whales were apparently attracted to potentially familiar sounds of herring-feeding killer whales. Changes in the dive profile not thought to be orientation movements accounted for 11 scored responses. Across all three species, seven such changes were associated with cessation of foraging. In two cases (Gm09_138a_7 and Gm09_156b_7), long-finned pilot whales started deep-diving during exposure. Changes in diving were associated with cessation of resting in a killer whale playback session with a sperm whale (Sw09_142a_5), and the same whale switched from foraging to resting dives during a later sonar exposure (Sw09_142a_12).

Vocal behavior was scored to have changed in 13 cases (Figure 8) and included both increases (Oo08_149a_1 & 3 and Sw09_141a_2 & 7) and decreases in social calling rates (Gm08_150c_2). Other modifications of social sounds included production of calls closely linked to the arrival of individual sonar pings (Oo09_144a_1), modification of some calls to be more similar to the sonar signal (Gm08_150c_1 and Gm09_138a_6), and shifts towards high-frequency whistles (Oo08_149a_6; Samarra et al., 2010). In sperm whales, echolocation behavior changed in several cases (e.g., cessation of clicking during deep dives: Sw08_152a_3) including one brief change (started clicking during ascent) associated with orientation movements in the dive profile (Sw09_141a_4). In three cases (Sw08_152a_3, Sw09_142a_9, and Sw09_160a_2), reduction or cessation of clicking and buzz sounds in sperm whales was also judged to indicate cessation of foraging.

Changes in group spacing or synchrony that were scored as responses were observed seven times (Figure 8). Spacing became tighter in three putative behavioral responses of one group of killer whales to sonar (Oo09_144a_5, 7 & 10), and in two groups of long-finned pilot whales (Gm09_138a_2 and Gm09_156b_2). Group spacing was also observed to change in one killer whale playback (Gm08_159a_4) and one Silent vessel approach (Gm09_156b_1) to long-finned pilot whales.

The most severe change in our entire study (severity of 8) was an 86 min separation of a small killer whale calf from its group that extended beyond the duration of the sonar exposure and was associated with avoidance and changes in vocal behavior. As detailed in Miller et al. (2011), this exposure session was unusual in our dataset for three reasons: (1) this session was the only repeated MFAS-UP exposure presented to the same group of animals; (2) the experiment was conducted in an unusually narrow fjord roughly 1 km wide; and (3) transmissions were started unusually close to the subjects. The tagged animal group crossed to the other side of the narrow fjord during the ramp-up phase, which was scored as an avoidance response. After this, the calf was first seen travelling alone when the sonar was received by the focal whale at a SPL of roughly 152 dB re: 1 μPa, but the calf separation must have started earlier and at a lower SPL than when we first detected it. The observation boat followed the calf and noted that the calf was always oriented towards its group, which was moving down the fjord ahead of the calf. The calf later rejoined its group and was sighted together with its group for many hours after the sonar exposure. This separation seemed to have been related to the avoidance response of the group and was judged to include interference with acoustic reunion mechanisms. Ultrasonic high-frequency whistles (Samarra et al., 2010) were produced during the time period when the calf was seen away from its group, but not before the separation, indicating a shift by the whales to communication in a higher frequency band (Figure 9; see Miller et al., 2011). The high-frequency whistles were clear within the Dtag recording, so they were likely produced by animals near the tagged animal. We cannot confirm whether the calf also produced whistles as it was separated from the group at this time.

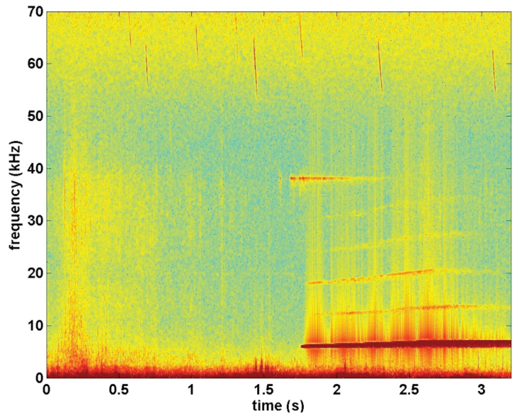

Figure 9. Spectrogram of sounds recorded on the Dtag during calf separation of experiment Oo08_149a; visible in the spectrogram is an MFAS-UP transmission (with harmonics) from 6 to 40 kHz and high-frequency whistles of killer whales above 50 kHz. A 38 kHz echosounder signal is also visible.

Severity of Scored Responses in Relation to Exposure Type

There was a clear pattern of more severe behavioral changes occurring during sonar and killer whale exposures than during Silent vessel passes (Table 3, top). Numerous scored responses involved changes in behavioral state of the whale subject. Looking across the different exposure types, the number of scored responses in an exposure session tended to be higher when the maximum severity change during a session was higher (Table 3). Sessions with one to two changes had a lower mean severity score (4.2 ± 1.6; N = 29) than sessions with three or more changes (mean score of 6.5 ± 0.7; N = 11).

Of the seven Silent-vessel approach sessions, the maximum severity behavioral change was a severity of 4 or greater in only two (29%) cases, and a severity of 5 was the highest value observed. In contrast, during sonar exposure sessions, scored responses with a severity of 4 or greater were more commonly observed (Table 3, bottom). As there were few LFAS-DN exposure sessions, they were pooled with the LFAS-UP exposure sessions, resulting in 83% of LFAS exposure sessions having a maximum severity of 4 or greater. A log-likelihood test indicated a low probability that the differences in the proportion of high (> 4) severity changes across Silent, MFAS, and LFAS exposure sessions were due to chance ($G_2 = 7.2$, $p = 0.027$). Post-hoc subdivision of the table indicated that the strongest statistical difference was between Silent and LFAS exposure sessions ($G_1 = 6.7$, $p = 0.009$), while the other two contrasts had p values greater than 0.10.

Table 3. Severity of scored responses by stimulus type: Top – the number of scored responses sorted by severity score and stimulus type for all three species combined; Bottom – the number of exposure sessions sorted by the maximum severity score during the exposure session and the stimulus type used.

Severity	Silent	MFAS-UP	LFAS-UP	LFAS-DN	KW
9					
8		1			
7		2	3		1
6		4	8	4	1
5	1	6	4	1	2
4	1	2	9	5	5
3	2	3	3	2	1
2		1	3	0	2
1	1	1	3	1	1
Total # of responses scored	5	20	33	13	13
Mean per session	0.7	1.4	2.5	2.6	1.6
Severity	Silent	MFAS-UP	LFAS-UP	LFAS-DN	KW
9					
8		1			
7		1	2		1
6		4	5	4	1
5	1	1	1		2
4	1	1	2	1	1
3	2	1	1		1
2			1		1
1	1	1			
0	2	4	1		1
Total # of sessions	7	14	13	5	8
% severity 4 to 9	29	57	77	100	63

For killer whale playback sessions, 63% had a maximum severity behavioral change of 4 or greater. This value is close to the overall proportion for sonar exposures, indicating that sounds of killer whales and sonar had effects with similar severity on the target species. Behavioral changes were evident in playback sessions of herring-feeding killer whale sounds to long-finned pilot and mammal-feeding killer whale sounds to sperm whales. In contrast, killer whales showed little behavioral change during playback of herring-feeding conspecific sounds (Oo08_149a and Oo09_144ab).

Relating Severity to Sonar Received Sound Pressure Level by Species

All of the behavioral changes recorded during sonar exposure sessions were separated by species and tabulated in relation to the SPL_{max} prior to the onset of the behavioral change (Table 4). The SPL_{max} associated with the highest severity behavioral change was also tabulated by species, including sessions in which no putative behavioral response was detected (Table 4). Looking for trends within each of the species, it is important to recognize that the difference in maximum source level of the sonar system (199 and 214 dB re: 1 μPa m for MFAS and LFAS, respectively) resulted in a large difference in the SPL_{max} to which animals were exposed in each session (Table 4). Few MFAS exposure sessions had a SPL_{max} beyond the 150 to 159 dB re: 1 μPa bin, whereas most LFAS sessions had a SPL_{max} in the 160 to 169 dB re: 1 μPa bin (Table 4). Thus, the number of bins with received levels of 160 dB re: 1 μPa or greater should be ignored when relating severity to received level differences for LFAS vs MFAS exposure sessions.

Number and Severity of Scored Responses Versus Species, Sonar Frequency, and Order

Across the three species, killer whales had relatively more scored responses to MFAS (3.0/session) than LFAS (2.0/session). In contrast, both long-finned pilot and sperm whales had more scored responses to LFAS (1.9/session and 3.5/session, respectively) than MFAS (0.8/session and 0.8/session, respectively). Killer whales had a response scored in all sonar sessions, with a similar distribution of received levels for the two frequency bands (Table 4). In contrast, sperm whales and long-finned pilot whales were more

Table 4. Severity of all scored behavioral responses vs the maximum received SPL at the onset of the behavioral change by species; M and L indicate transmissions in the MFAS (6 to 7 kHz) and LFAS (1 to 2 kHz) frequency bands, respectively. Underlining indicates behavioral changes that occurred in the first sonar exposure session for each subject whale, and bold typeface indicates the maximum severity change within a session. The number of MFAS and LFAS exposure sessions during which the received SPL_{max} was at least the level of the bin is indicated in the two bottom rows for each species.

Experiments with Killer Whales

Severity score	Received SPL_{max} sorted into 10 dB bins: 80-89	90-99	100-109	110-119	120-129	130-139	140-149	150-159	160-169	170-179	180-189	190-199
9												
8								M				
7		**L**L				**M**M						
6		L**M**						L	L			
5		L		M		M	M	ML	L			
4												
3						M	**M**M					
2				M				L				
1												
0												
#MFAS	4	4	4	4	4	4	4	3	0	0	0	0
#LFAS	4	4	4	4	4	4	4	4	3	2	0	0

Experiments with Long-Finned Pilot Whales

Severity score	Received SPL_{max} sorted into 10 dB bins: 80-89	90-99	100-109	110-119	120-129	130-139	140-149	150-159	160-169	170-179	180-189	190-199
9												
8												
7								L				
6	L				M		**L**	L**L**				
5				M				**LM**L				
4	**M**						L	**L**L		**L**	L	
3								**L**	LL			
2												
1								M				
0								M	M	L		
#MFAS	6	6	6	6	6	6	6	6	1	0	0	0
#LFAS	8	8	8	8	8	8	8	8	8	6	1	0

Experiments with Sperm Whales

Severity score	Received SPL_{max} sorted into 10 dB bins: 80-89	90-99	100-109	110-119	120-129	130-139	140-149	150-159	160-169	170-179	180-189	190-199
9												
8												
7												
6					ML	M	L	LL	L			
5												
4					LL		LLL	LM	**L**LL			
3							L	L				
2								LL				
1				L		L	L	L				
0							M	M				
#MFAS	4	4	4	4	4	4	4	1	0	0	0	0
#LFAS	6	6	6	6	6	6	6	6	4	0	0	0

likely to have no response scored during MFAS than LFAS sonar transmissions, but responses that were scored tended to start at lower received levels for MFAS than LFAS exposure sessions. Killer whales had higher severity scores than the other species, with four at a severity of 7 and one at a severity of 8 of the responses scored.

The GEE Model indicated possible differences between species and frequency, but no influence of order on either maximum severity score or number of responses scored during a session (Table 5). Long-finned pilot whales had lower maximum severity scores within a session than killer whales (Estimate: -0.44 ± 0.19; Wald statistic: 5.33; $p = 0.021$), but no differences were apparent between sperm whales and long-finned pilot whales or sperm whales and killer whales. There was a trend for severity scores to be lower for MFAS than for LFAS (Estimate: -0.30 ± 0.17; Wald statistic: 3.12; $p = 0.077$). The number of responses scored during a session was lower for long-finned pilot whales than killer whales (Estimate: -0.67 ± 0.18; Wald statistic: 13.63; $p = 0.0002$) and greater for sperm whales than long-finned pilot whales (Estimate: 0.54 ± 0.19; Wald statistic: 7.94; $p = 0.005$). The number of changes during a session was significantly less for MFAS than for LFAS (Estimate: -0.56 ± 0.25; Wald statistic: 5.01; $p = 0.025$) (Table 5).

Discussion

This study was designed to record behavioral changes of cetaceans that could be considered responses to sonar signals, and to define the received level at which those putative responses were scored. The primary motivation for the study was the need for a dose-response function to evaluate the risk of behavioral effects of sonar on free-ranging cetaceans. Our exposure protocol increased the received level of the sonar throughout each exposure session by ramping up the source level to 197 to 214 dB re: 1 µPa m, close to that of a number of operational sonars (D'Amico & Pittenger, 2009; Ainslie, 2010). We started our exposure sessions independent of the context or behavioral state of the subject whale, which matches what would be expected to occur in operational sonar use. However, the approach of the vessel from a starting distance of 6 to 8 km probably led to a more intense exposure than would be typical for actual exercises, where the motion of sonar vessels is independent of whale location. All of these factors make the experiments a realistic though possibly worse than normal scenario for sonar exposures from real navy activities. The same movement pattern was repeated with no sonar transmissions for the Silent vessel exposure. Recordings of the calls of killer whales were played back using a sound source from a small boat at a source level typical for killer whales (Miller, 2006), much lower than the source level of the sonar.

We used a complex suite of tools (visual tracking, group-level observations, and Dtags [Johnson & Tyack, 2003] attached to whales) to monitor multiple behaviors and to be able to consistently identify a wide range of potential changes in behavior (Table 1). The combination of a large suite of response parameters, variable contexts for each exposure, and a relatively small sample size of subjects tested makes it difficult to use standard experimental data analysis and statistical techniques to test null hypotheses of no effect.

Table 5. Results of the GEE Model fit for maximum severity score within a session (top section) and number of changes scored within a session (bottom)

Variable	Estimate ± SE	Wald	p value
(Intercept)	1.77 ± 0.23	58.26	< 0.0001
Order	0.03 ± 0.06	0.30	0.582
MFAS vs LFAS	-0.30 ± 0.17	3.12	0.077
Pilot whale vs killer whale	-0.44 ± 0.19	5.33	0.021
Sperm whale vs killer whale	-0.23 ± 0.20	1.35	0.246
Sperm whale vs pilot whale	0.21 ± 0.21	1.01	0.315

Variable	Estimate ± SE	Wald	p value
(Intercept)	1.24 ± 0.34	13.58	0.00023
Order	0.02 ± 0.13	0.03	0.855
MFAS vs LFAS	-0.56 ± 0.25	5.01	0.025
Pilot whale vs killer whale	-0.67 ± 0.18	13.63	0.00022
Sperm whale vs killer whale	-0.13 ± 0.15	0.71	0.399
Sperm whale vs pilot whale	0.54 ± 0.19	7.94	0.00485

Instead, the approach taken here was to use expert inspection and interpretation of each dataset in its entirety, with independent scoring of the data by different groups, to identify behavioral changes that were not commonly identified in the baseline record for each subject and could therefore be scored as putative responses to the different exposure conditions (sonar, Silent, killer whale playback).

As our goal was to score responses to the sonar, we did not score changes in behavior as responses if similar changes were identified within the baseline records for each animal or if the behavioral change could have occurred before the start of the exposure (e.g., Sw09_141a Silent). However, this method does not allow us to conclude whether or not each identified change in behavior was truly a response to the sonar or not, nor to connect each scored response to a statistical likelihood that a change in behavior happened during exposure by chance. A turn away from the source vessel path is scored as an avoidance response, but the turn could have been caused by natural factors such as following a bathymetry contour (see Sw08_152a MFAS). Detailed examination of specific behaviors and their variability in baseline vs exposure conditions may be possible in some cases, but statistical analysis of any but the largest and most consistent responses will likely require larger data sets than we have obtained in our study.

The list of responses scored in this study is therefore a precautionary interpretation of the outcomes of our experiments. Very few behavioral changes, particularly those with high severity, would be missed in our analyses, but some changes scored as responses may not have actually been in response to the sonar. While these caveats are important, many policy applications require this kind of precautionary approach, and our consistent application of the descriptive method in this study allowed us to evaluate the number and severity of potential responses of cetaceans to the experimental exposures used in this study. We examined the influence of the exposure type (sonar, Silent, killer whale playback) on the number and severity of responses scored, the potential influence of frequency and order of sonar transmissions, and the pattern of severity vs received level. The Southall et al. (2007) severity scale considers both the type of response and its duration relative to the duration of the exposure, which allows us to extrapolate severity from these short experimental exposures to real exposure scenarios during full-scale naval exercises at sea.

Patterns of Severity by Exposure Type

There was a strong difference in the number and maximum severity of responses scored within a session depending upon the exposure type. The Silent control passes had both the fewest responses scored per session and the lowest proportion of sessions with maximum severity scores of 4 or greater (Table 3). MFAS (6 to 7 kHz) exposure and killer whale playback sessions had the next highest rate of responses and maximum severity. The killer whale sounds played to sperm whales were unfamiliar sounds from mammal-feeding killer whales, and the scored responses to killer whale playback were very similar in type and severity as responses scored to the sonar. Because the herring-feeding sounds played to the long-finned pilot whales were recorded nearby, those sounds may have been familiar sounds, so the reaction of the long-finned pilot whales may have been shaped by experience (Deecke et al., 2002). Two of the killer whale playbacks to long-finned pilot whales resulted in clear approach towards the speaker of the same duration as the exposure, with a severity of 5. An avoidance behavior of the same duration would have received a higher severity score of 6.

We conducted five exposure sessions in which we transmitted the LFAS (1 to 2 kHz) waveform as a hyperbolic downsweep rather than the upsweep used in the other exposures. Though the sample was too small to warrant statistical comparison, the number of changes and maximum severity within a session were quite similar for upsweep vs downsweep sessions within the LFAS (1 to 2 kHz) band. This finding is consistent with recent work with captive animals that found no difference in the startle thresholds for upsweep and downsweep forms of similar signals (Kastelein et al., 2012). We therefore pooled the results for the upsweep and downsweep LFAS exposure sessions, which then had the highest rate of changes per session (2.5) and the greatest proportion of changes at a severity of 4 or greater (83%).

The highest severity response scored during a Silent approach was 5, and the proportion of sessions with maximum severity scores of 4 or greater was significantly lower for Silent exposure sessions than sonar exposure sessions ($G_2 = 7.2$, $p = 0.027$). The greatest difference was between Silent exposure sessions and pooled upsweep and downsweep 1 to 2 kHz exposure sessions. Higher severity changes in behavior were more common during sonar exposure sessions than Silent exposure sessions, suggesting that it was the sonar (rather than the approaching vessel) that led to the highest severity responses scored in our dataset. This result is important as it clearly suggests that the additional changes observed during sonar exposure represented responses to the sonar.

Responses Scored During Sonar Exposure Sessions

Our analysis of the behavior of the three cetacean species indicated a rich array of behavioral changes that were scored as responses to the sonar signals (Figure 8). The relevant table for comparison is Table 17 in Southall et al. (2007) for mid-frequency cetaceans reported to have had behavioral responses to non-pulsed signals. Our data add a substantial number of observations of behavioral changes across the range of received SPLs from 80 to 180 dB re: 1 µPa. Much like Table 17 in Southall et al., there is little indication in our results of a dose-response pattern in which higher severity changes are less common at lower received levels and more common at higher received levels. Instead, we scored behavioral responses to have occurred across a wide range of received levels.

Seven scored responses to sonar started at received SPLs of < 110 dB re: 1 µPa. A scored response is considered to be less likely to have happened by chance when a suite of changes occur together, and we therefore have more confidence that behavioral changes may have been responses to the sonar when several changes occurred together. This was the case for the five changes that occurred in two < 110 dB re: 1 µPa exposure sessions (LFAS-UP and MFAS) with killer whale group Oo09_144ab (Appendix Table A1). In contrast, the two scored responses for long-finned pilot whales were each from different < 110 dB re: 1 µPa sessions, and we therefore have less confidence that these truly represent responses to the sonar at such low received levels. Ten responses were scored at the opposite end of the received level table (> 160 dB re: 1 µPa). Eight of these involved behavioral changes also identified at lower received levels (avoidance, group spacing, vocal changes). The other two responses scored were novel observations of long-finned pilot whales surfacing in near-perfect synchrony with arrivals of three or four successive sonar pings. The synchronous surfacing with the duty cycle of the sonar occurred with two different whales, both when the vessel was near its closest point of approach and received SPLs were 175 to 180 dB re: 1 µPa. We interpret these synchronous surfacings with the signal interval of the sonar as a type of vertical avoidance of the received level of the sonar signal as the surface pressure release should strongly reduce SPLs (Jensen, 1981).

The highest severity response scored to the sonar in our dataset was a separation of a calf from its group that extended beyond the end of the exposure period. Confirmation of the calf travelling alone triggered a mitigation stop of the sonar transmissions, which coincidentally were scheduled to be stopped at that time. We took care in subsequent experiments to carefully count the number of any calves within our subject groups throughout observational follows, and we did not identify any other cases in which a calf separated from its group. It is possible the unique context of this exposure session in experiment Oo08_149a, involving restricted bathymetry of the narrow fjord and the relatively close start distance from the whales at the start of the exposure (Miller et al., 2011), contributed to the calf separation.

Five scores with a severity of 7 and 16 scores with a severity of 6 were identified during sonar exposure sessions. Prolonged avoidance of the sound source (that continued for at least 5 h after the end of the sonar transmissions) accounted for two severity of 7 changes, both with killer whales. Prolonged cessation of feeding accounted for the other three severity of 7 scores, two with the same killer whales that also showed prolonged avoidance and one with long-finned pilot whales associated with moderate avoidance that ended soon after exposure ended. Moderate cessation of foraging (severity of 6) was observed in three sperm whale sessions and in one long-finned pilot whale session. The tendency for cessation of foraging to extend beyond the duration of the exposure for killer whales, but not for long-finned and sperm whales, may reflect the patchy nature of the schooling herring, which were the prey of killer whales, vs more ubiquitous benthic and deep-water prey of long-finned pilot and sperm whales.

Influence of Sonar Frequency, Session Order, and Species on Putative Responses

Killer whales showed a combination of the highest severity scores and the widest distribution in received SPL at the onset of scored responses, including many at received levels below 120 dB re: 1 µPa (Table 4). One experiment (Oo09_144ab) yielded strong behavioral changes, with prolonged avoidance and cessation of foraging (severity of 7) and moderate avoidance (severity of 6) commencing very early in the LFAS and MFAS exposure sessions. High severity behavioral responses (of 6 or greater) were judged to have occurred in six of eight sonar exposure sessions, including a separation of a calf judged a severity of 8 (Oo08_149a), and two cases of prolonged cessation of foraging judged a severity of 7 (Oo06_327s and Oo09_144a). There were no sessions with killer whales during which no response to the sonar was scored.

Long-finned pilot whales, in contrast, had a strong clustering of scored responses at received levels between 150 to 159 dB re: 1 µPa, with a few cases below and above this level (Table 4). One case of avoidance during an MFAS trial was observed at a received SPL of 120 to 129 dB re: 1 µPa (Gm08_150c), which also had cases of calls matching the sonar signal just prior to the start of the avoidance. One prolonged increase in group cohesion began at the very start of an LFAS-UP sonar session in which prolonged cessation of foraging was also noted. Indeed, it is possible that the change in group spacing occurred just prior to the start of the sonar exposure itself, in-between two sightings; however, the change also may have been part of the change to diving and foraging behavior. Another switch to deep-diving occurred very early in another session (Gm09_156b_7), with a received SPL before the change of 80 to 89 dB re: 1 µPa. No response to the sonar was scored in three MFAS and one LFAS exposure sessions to long-finned pilot whales (Table 4).

Sperm whales showed a different pattern than long-finned pilot whales, with scored responses to sonar commencing across a wide range of received levels from 120 to 169 dB re: 1 µPa (Table 4). Seven of 10 exposure sessions resulted in maximum severity scores of 6, with moderate avoidance occurring in four cases, and moderate cessation of foraging in three other cases. Avoidance and cessation of foraging were not linked in most cases. Scored responses with a severity of 4 or greater were detected in all LFAS sonar sessions with sperm whales, while no responses were scored in two of the four MFAS exposure sessions.

Separate statistical analysis of the number of responses scored and the maximum severity within a session supported the descriptive analysis given above. For both the number of scored responses and the maximum severity within a session, the order of the sonar session (1st session to a subject, 2nd, or 3rd) had no effect (Table 5). In one case (Oo09_144a), the continuing avoidance and cessation of feeding responses to exposures in the first session may have influenced how the animals responded in the second and third sessions. However, there was no overall tendency for order to influence the number or maximum severity of scored responses within sonar exposure sessions. Thus, the subjects in our experiments did not seem to consistently habituate or become sensitized to the sonar.

There was an overall tendency for both the number ($p = 0.025$) and maximum severity ($p = 0.077$) of scored responses within a session to be greater for LFAS (1 to 2 kHz) exposure sessions than for MFAS (6 to 7 kHz) sessions. This result is somewhat surprising because the behavioral and electrophysiological audiograms of three killer whales indicated a 10 to 40 dB lower sensitivity at 1 to 2 kHz than 6 to 7 kHz (Hall & Johnson, 1972; Szymanski et al., 1999). Less is known about the hearing of long-finned pilot and sperm whales, although electrophysiological data for these two species (Ridgway & Carder, 2001; Pacini et al., 2010), in combination with the typical U-shaped audiograms of odontocetes (Johnson, 1966; White et al., 1978; Thomas et al, 1988; Kastelein et al., 2010), also indicate greater sensitivity at 6 to 7 kHz than 1 to 2 kHz. Given the limited sample size, it was not possible to test all potential interaction terms in the GEE analysis; however, the tendency for scored responses to the sonar to be more numerous and more severe in 1 to 2 kHz than 6 to 7 kHz sessions seems to be stronger for sperm and long-finned pilot whales than for killer whales (Table 4).

Finneran & Schlundt (2011) suggested that weighting received levels by equal-loudness contours following the shape of the audiogram may be a useful way to reflect the perceived loudness of tonal signals by bottlenose dolphins. In line with this view are the results of Kastelein et al. (2012) who found that startle response thresholds of a captive harbor porpoise to 6 to 7 kHz sweeps without harmonics were lower than to 1 to 2 kHz sweeps without harmonics but similar to 1 to 2 kHz sweeps with many high-frequency harmonics. Our results suggest that for the free-ranging cetaceans studied here, differences in the sonar frequency alone did not drive the number or severity of putative responses. Instead, we feel it is likely that the higher source level in the LFAS (1 to 2 kHz) band used in our experiments could explain why the number and maximum severity of scored responses was higher during LFAS exposure sessions than during MFAS (6 to 7 kHz) sessions. The higher source level used in LFAS sonar sessions led to higher received levels at comparable distances from the source. Though there was no obvious pattern of severity scaling with received level, it remains plausible that higher received levels during LFAS sessions could have led to more responses and more severe responses. Alternatively, we cannot rule out that the higher source level itself, or different patterns of reverberation and/or harmonics, were salient features of the source to which subject whales were more likely to respond with higher severity levels. For source level itself to influence responsiveness, the whale would need to be able to integrate the received level and the distance to the source in order to recognize the source level. The ability to judge the distance to a sound source has been experimentally demonstrated in birds and humans (Naguib & Wiley, 2001). Though research to assess the ability of cetaceans to judge the distance to a sound source has not been feasible to date, it is plausible that they can judge the distance to sound sources using the same cues (e.g., reverberation) that are used by the better-studied terrestrial species.

The statistical analysis also indicated significant differences between species (Table 4). Long-finned pilot whales had lower maximum severity scores within a session than killer whales, but no differences were apparent between sperm vs long-finned pilot whale or sperm vs killer whales. The number of behavioral changes during a session was lower for long-finned pilot whales than killer whales and greater for sperm whales than long-finned pilot whales. There was no difference between sperm whales and killer whales in the number of responses scored per exposure session. Taken together, these two statistical results indicate that long-finned pilot whales were somewhat less sensitive to the sonar exposures we conducted than sperm whales and killer whales.

The statistical analysis did not take into account the received level at which scored responses started to occur; however, one killer whale (Oo09_144ab) showed clear indications of severe responses to the sonar that started at very low received levels (< 100 dB re: 1 μPa). This may indicate that killer whales have a risk to be particularly sensitive, with some groups responding strongly to sonar at received SPLs just loud enough to be audible. Given the small sample size tested, however, we cannot conclude that such rare extreme responses do not potentially occur in the other tested species. Tyack et al. (2011) reported behavioral responses of beaked whales at very low received SPLs.

Over all the responses scored, the relationship between sonar dose and severity appeared to be highly variable, likely due to a wide degree of variation in response thresholds. While the received level of an anthropogenic sound like naval sonar is likely to be one important factor, unexplained variation in response thresholds are also likely to result from specific contextual variables, such as previous exposure history of the group, detailed behavioral context, or condition of individuals within the group, which are difficult to predict, measure, or control with free-ranging animals. Ellison et al. (2012) proposed that behavioral responses at low received levels were likely to be particularly influenced by context, whereas responses at higher received levels were likely to be driven by a stronger dose-response function. Some of the responses scored in our study seem to conform to this idea. The synchronous surfacings of long-finned pilot whales to arrivals of sonar pings (Gm08_159a_2 and Gm09_156b_6) were the highest received level responses scored in our dataset, and we interpret such synchronous surfacings to be a direct attempt to reduce the received SPL of the sonar pings. On the lower end of the received level scale, it seems unlikely that the strong response of killer whale group Oo09_144ab was driven by the perceived loudness of the sonar which had a very low SPL when the response began. Because responses to sonar at very low received levels can potentially occur over a wide area of habitat affected by anthropogenic noise, a better understanding of what factors drive such responses is important for evaluating the risk of harm to cetaceans from sonar usage in the oceans.

Applications and Conclusions

The results of our experiments indicate that naval sonar potentially affects the behavior of cetaceans in a number of different ways and over a wide range of received levels, and that the severity of these behavioral changes range from unlikely to likely to affect vital rates. Our descriptive results enable evaluation of the potential impact of experimental sound exposures as a research activity, which is important for permitting and animal welfare evaluation of this type of research (e.g., Hooker et al., 2001). Our experiments were short in duration but otherwise represented realistic exposure scenarios with a moving source transmitting at high source levels. The duration of our experimental exposures makes it unlikely that the experiments themselves had any biologically significant effect on the research subjects; however, had the exposure been extended from minutes to days, as could be the case during full-scale sonar exercises, there is a risk that some of the behavioral changes observed could escalate to severe biological changes over longer durations with a risk of significant impact on the exposed animal(s). The severity scale used on our experimental data considers the duration of the change in behavior in relationship to the duration of the exposure. To evaluate the impact of actual naval sonar exercises, a careful extrapolation from responses observed during these experiments is needed to make an estimate of impact accounting for duration, exposure levels, and the number of affected animals.

Several aspects of our results have relevance to managers of the behavioral effects of anthropogenic sound on marine mammals in general and to naval operational planners in particular. Our results suggest that different species may have different sensitivity to sonar exposure, and mitigation measures should take this into account. Killer whales had higher severity responses at lower received SPLs during sonar exposures than long-finned pilot whales and sperm whales. Taxonomic proximity did not predict similarity in sensitivity to sonar disturbance as we have found a higher discrepancy between long-finned pilot whales and killer whales, the taxonomically closest species in our study, compared to sperm whales. This calls for great care during the extrapolation of results from experimental studies on a particular species to other closely related species.

Though there was an overall tendency for increased risk of a severe behavioral response above 120 to 130 dB re: 1 µPa received SPL_{max}, our results do imply that any signal audible to the animal can represent some risk of a behavioral response at any severity level between 0 and 7. As discussed above, the behavioral context of animals is likely to be an important factor influencing the severity of a response (Ellison et al., 2012). In fact, our results show that when animals are engaged in important functional behaviors, such as feeding or resting or taking care of young offspring, there was a higher risk that disruption of this behavior would be scored as more severe. Within a specific behavioral context, the dose vs severity relationship might be stronger, but our low sample size and our random context experimental design limit our ability to resolve this. For naval planning purposes, however, it is important to not only consider the abundance of marine mammals but also the functional activities (e.g., feeding, migrating, and breeding) of animals in the area.

Ultimately, the most important effects of a stressor such as noise are those that impact survival, growth, and reproduction. Some of these effects can be stratified by season and location. For example, reproduction is more likely to be disrupted in the breeding grounds and season; feeding is more likely to be disrupted in feeding areas and times. However, it may be even more important to investigate effects when animals are already stressed. It is well-recognized that the impact of disruption of feeding may be higher on animals with higher energetic demands, such as lactating females, or in poorer condition, such as males at the end of a reproductive season (McEwen & Wingfield, 2003). Paradoxically, the most stressed animals may have less capacity to alter their behavior—for example, a starving animal may choose to keep feeding when exposed to a disturbing stimulus rather than stop feeding and move away (Beale & Monaghan, 2004). Thus, the animals most vulnerable to a disturbing stimulus may not be the most responsive. While our study indicates context-specific dose-response relationships for behavioral responses, the task of protecting wildlife demands combining these dose-response relationships with knowledge of the contexts and life history stages when animals are most vulnerable.

Acknowledgments

Thanks to all 3S-team members and ships' crews for their efforts in the field, to René Dekeling and two anonymous reviewers for manuscript review, and to Filipa Samarra for statistical support. Visual data were collected using *Logger 2000*, developed by the International Fund for Animal Welfare (IFAW) to promote benign and non-invasive research. This study was funded by the U.S. Office of Naval Research, the Royal Norwegian Navy and the Norwegian Ministry of Defence, the Defence Research and Development Department of the Netherlands Ministry of Defence, and by WWF Norway.

Literature Cited

Ainslie, M. A. (2010). *Principles of sonar performance modeling*. Chichester, UK: Springer-Praxis. http://dx.doi.org/10.1007/978-3-540-87662-5

Balcomb, K. C., & Claridge, D. E. (2001). A mass stranding of cetaceans caused by naval sonar in the Bahamas. *Bahamas Journal of Science*, *5*, 2-12.

Beale, C. M., & Monaghan, P. (2004). Behavioural responses to human disturbance: A matter of choice? *Animal Behaviour*, *68*, 1065-1096. http://dx.doi.org/10.1016/j.anbehav.2004.07.002

Bejder, L., Samuels, A., Whitehead, H., & Gales, N. (2006a). Interpreting short-term behavioural responses to disturbance within a longitudinal perspective. *Animal Behaviour*, *72*, 1149-1158. http://dx.doi.org/10.1016/j.anbehav.2006.04.003

Bejder, L., Samuels, A., Whitehead, H., Gales, N., Mann, J., Connor, R., . . . Krutzen, M. (2006b). Decline in relative abundance of bottlenose dolphins exposed to long-term disturbance. *Conservation Biology*, *20*, 1791-1798. http://dx.doi.org/10.1111/j.1523-1739.2006.00540.x

Burdic, W. S. (1991). *Underwater acoustic system analysis* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

D'Amico, A., & Pittenger, R. (2009). A brief history of active sonar. *Aquatic Mammals*, *35*(4), 426-434. http://dx.doi.org/10.1578/AM.35.4.2009.426

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. *Aquatic Mammals, 35*(4), 452-472. http://dx.doi.org/10.1578/AM. 35.4.2009.452

Deecke, V. B., Slater, P. J. B., & Ford, J. K. B. (2002). Selective habituation shapes acoustic predator recognition in harbour seals. *Nature*, *420*, 171-173. http://dx.doi.org/10.1038/nature01030

Ellison, W. T., Southall, B. L., Clark, C. W., & Frankel, A. S. (2012). A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Conservation Biology*, *26*, 21-28. http://dx.doi.org/10.1111/j.1523-1739.2011.01803.x

Fay, R. R. (1988). *Hearing in vertebrates: A psychophysics databook*. Winnetka, IL: Hill-Fay Associates. http://dx.doi.org/10.1038/32068; http://dx.doi.org/10.1038/32071

Finneran, J. J., & Schlundt, C. E. (2011). Subjective loudness level measurements and equal loudness contours in a bottlenose dolphin (*Tursiops truncatus*). *The Journal of the Acoustical Society of America*, *130*, 3124. http://dx.doi.org/10.1121/1.3641449

Finneran, J. J., Carder, D. A., Schlundt, C. E., & Dear, R. L. (2010). Growth and recovery of temporary threshold shift at 3 kHz in bottlenose dolphins: Experimental data and mathematical models. *The Journal of the Acoustical Society of America*, *127*, 3256-3266. http://dx.doi.org/10.1121/1.3372710

Finneran, J. J., Carder, D. A., Schlundt, C. E., & Ridgway, S. H. (2005). Temporary threshold shift in bottlenose dolphins (*Tursiops truncatus*) exposed to mid-frequency tones. *The Journal of the Acoustical Society of America*, *118*, 2696-2705. http://dx.doi.org/10.1121/1.2032087

Frantzis, A. (1998). Does acoustic testing strand whales? *Nature*, *392*, 29.

Hall, J. D., & Johnson, C. S. (1972). Auditory thresholds of a killer whale, *Orcinus orca* Linnaeus. *The Journal of the Acoustical Society of America*, *51*, 515-517. http://dx.doi.org/10.1121/1.1912871

Højsgaard, S., Halekoh, U., & Yan, J. (2006). The R package geepack for generalized estimating equations. *Journal of Statistical Software*, *15*, 1-11.

Hooker, S. K., Baird, R. W., Al-Omari, S., Gowans, S., & Whitehead, H. (2001). Behavioral reactions of northern bottlenose whales (*Hyperoodon ampullatus*) to biopsy darting and tag attachment procedures. *Fishery Bulletin*, *99*, 303-308.

Jensen, E. B. (1981). Sound propagation in shallow water: A detailed description of the acoustic field close to surface and bottom. *The Journal of the Acoustical Society of America*, *70*, 1397-1406. http://dx.doi.org/10.1121/1.387130

Johnson, C. S. (1966). *Auditory thresholds of the bottlenose porpoise (*Tursiops truncatus*, Montagu)*. China Lake, CA: U.S. Naval Ordnance Test Station.

Johnson, C. S. (1968). Relation between absolute threshold- and duration-of-tone pulses in the bottlenosed porpoise. *The Journal of the Acoustical Society of America*, *43*(4), 757-763. http://dx.doi.org/10.1121/1.1910893

Johnson, M. P., & Tyack, P. L. (2003). A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering*, *28*, 3-12. http://dx.doi.org/10.1109/Joe.2002.808212

Kastelein, R. A., Hoek, L., de Jong, C. A., & Wensveen, P. J. (2010). The effect of signal duration on the underwater detection thresholds of a harbor porpoise (*Phocoena phocoena*) for single frequency-modulated tonal signals between 0.25 and 160 kHz. *The Journal of the Acoustical Society of America*, *128*, 3211-3222. http://dx.doi.org/10.1121/1.3493435

Kastelein, R. A., Steen, N., Gransier, R., Wensveen, P. J., & de Jong, C. A. (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*). *The Journal of the Acoustical Society of America*, *131*, 2325-2333. http://dx.doi.org/10.1121/1.3682032

Kvadsheim, P. H., Lam, F. P. A., Miller, P. J. O., Alves, A. C., Antunes, R., Bocconcelli, A., . . . Visser, F. (2009). *Cetaceans and naval sonar – The 3S-2009 cruise report* (FFI-rapport 2009/01140 ed.). Kjeller, Norway: Forsvarets Forskningsintitutt.

Malme, C. I., Miles, P. R., Clark, C. W., Tyack, P., & Bird, J. E. (1983). *Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behavior. Final report for the period of 7 June 1982 - 31 July 1983*. Anchorage, AK: U.S. Minerals Management Service.

Malme, C. I., Miles, P. R., Clark, C. W., Tyack, P., & Bird, J. E. (1984). *Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behavior. Phase II: January 1984 migration*. Anchorage, AK: U.S. Minerals Management Service.

Maybaum, H. L. (1993). Responses of humpback whales to sonar sounds. *The Journal of the Acoustical Society of America*, *94*, 1848-1849. http://dx.doi.org/10.1121/1.407710

McCauley, R. D., Fewtrell, J., Duncan, A. J., Jenner, C., Jenner, M-N., Penrose, J. D., . . . McCabe, K. (2000). Marine seismic surveys: A study of environmental implications. *The Australian Petroleum Production and Exploration Association Journal*, *40*, 692-708.

McEwen, B. S., & Wingfield, J. C. (2003). The concept of allostasis in biology and biomedicine. *Hormones and Behavior*, *43*, 2-15. http://dx.doi.org/10.1016/S0018-506X(02)00024-7

Miller, P. J. (2006). Diversity in sound pressure levels and estimated active space of resident killer whale vocalizations. *Journal of Comparative Physiology A*, *192*, 449-459. http://dx.doi.org/10.1007/s00359-005-0085-2

Miller, P. J., Johnson, M. P., & Tyack, P. L. (2004a). Sperm whale behaviour indicates the use of echolocation click buzzes "creaks" in prey capture. *Proceedings of the Royal Society B: Biological Sciences*, *271*, 2239-2247. http://dx.doi.org/10.1098/rspb.2004.2863

Miller, P. J., Aoki, K., Rendell, L. E., & Amano, M. (2008). Stereotypical resting behavior of the sperm whale. *Current Biology*, *18*, R21-23. http://dx.doi.org/10.1016/j.cub.2007.11.003

Miller, P. J., Biassoni, N., Samuels, A., & Tyack, P. L. (2000). Whale songs lengthen in response to sonar. *Nature*, *405*, 903. http://dx.doi.org/10.1038/35016148

Miller, P. J., Johnson, M. P., Tyack, P. L., & Terray, E. A. (2004b). Swimming gaits, passive drag and buoyancy of diving sperm whales *Physeter macrocephalus*. *Journal of Experimental Biology*, *207*, 1953-1967. http://dx.doi.org/10.1242/jeb.00993

Miller, P. J., 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 Research I*, *56*, 1168-1181. http://dx.doi.org/10.1016/j.dsr.2009.02.008

Miller, P. J., Antunes, R., Alves, A. C., Wensveen, P., Kvadsheim, P., Kleivane, L., . . . Tyack, P. L. (2011). *The 3S experiments: Studying the behavioural effects* *of naval sonar on killer whales (*Orcinus orca*), sperm whales (*Physeter macrocephalus*), and long-finned pilot whales (*Globicephala melas*) in Norwegian waters* (Scottish Oceans Institute Technical Report SOI-2011-001).

Moretti, D., Marques, T. A., Thomas, L., DiMarzio, N., Dilley, A., Morrissey, R., . . . Jarvis, S. (2010). A dive counting density estimation method for Blainville's beaked whale (*Mesoplodon densirostris*) using a bottom-mounted hydrophone field as applied to a Mid-Frequency Active (MFA) sonar operation. *Applied Acoustics*, *71*, 1036-1042. http://dx.doi.org/10.1016/j.apacoust.2010.04.011

Morfey, C. L. (2001). *Dictionary of acoustics*. San Diego: Academic Press.

Naguib, M., & Wiley, H. (2001). Estimating the distance to a source of sound: Mechanisms and adaptations for long-range communication. *Animal Behaviour*, *62*, 825-837. http://dx.doi.org/10.1006/anbe.2001.1860

Nowacek, D. P., Johnson, M. P., & Tyack, P. L. (2004). North Atlantic right whales (*Eubalaena glacialis*) ignore ships but respond to alerting stimuli. *Proceedings of the Royal Society B: Biological Sciences*, *271*, 227-231. http://dx.doi.org/10.1098/rspb.2003.2570

Nowacek, D. P., Thorne, L. H., Johnston, D. W., & Tyack, P. L. (2007). Responses of cetaceans to anthropogenic noise. *Mammal Review*, *37*, 81-115. http://dx.doi.org/10.1111/j.1365-2907.2007.00104.x

Pacini, A. F., Nachtigall, P. E., Kloepper, L. N., Linnenschmidt, M., Sogorb, A., & Matias, S. (2010). Audiogram of a formerly stranded long-finned pilot whale (*Globicephala melas*) measured using auditory evoked potentials. *Journal of Experimental Biology*, *213*, 3138-3143. http://dx.doi.org/10.1242/Jeb.044636

Plomp, R., & Bouman, M. A. (1959). Relation between hearing threshold and duration for tone pulses. *The Journal of the Acoustical Society of America*, *31*, 749-758. http://dx.doi.org/10.1121/1.1907781

R Development Core Team. (2008). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Richardson, W. J., Würsig, B., & Greene, C. R., Jr. (1990). Reactions of bowhead whales *Balaena mysticetus*, to drilling and dredging noise in the Canadian Beaufort Sea. *Marine Environmental Research*, *29*, 135-160. http://dx.doi.org/10.1016/0141-1136(90)90032-J

Richardson, W. J., Fraker, M. A., Würsig, B., & Wells, R. S. (1985). Behaviour of bowhead whales *Balaena mysticetus* summering in the Beaufort Sea: Reactions to industrial activities. *Biological Conservation*, *32*, 195-230. http://dx.doi.org/10.1016/0006-3207(85)90111-9

Richardson, W. J., Greene, C. R., Jr., Malme, C. I., & Thomson, D. H. (1995). *Marine mammals and noise*. San Diego: Academic Press. http://dx.doi.org/10.1111/j.1748-7692.1995.tb00272.x

Ridgway, S. H., & Carder, D. A. (2001). Assessing hearing and sound production in cetaceans not available for behavioral audiograms: Experiences with sperm, pygmy sperm, and gray whales. *Aquatic Mammals*, *27*(3), 267-276.

Samarra, F. I. P., Deecke, V. B., Vinding, K., Rasmussen, M. H., Swift, R. J., & Miller, P. J. O. (2010). Killer whales (*Orcinus orca*) produce ultrasonic whistles. *The Journal of the Acoustical Society of America*, *128*, El205-El210. http://dx.doi.org/10.1121/1.3462235

Simon, M., Wahlberg, M., Ugarte, F., & Miller, L. A. (2005). Acoustic characteristics of underwater tail slaps used by Norwegian and Icelandic killer whales (*Orcinus orca*) to debilitate herring (*Clupea harengus*). *Journal of Experimental Biology*, *208*, 2459-2466. http://dx.doi.org/10.1242/jeb.01619

Southall, B. L., Bowles, A. E., Ellison, W. T., Finneran, J. J., Gentry, R. L., Greene, C. R., Jr., . . . Tyack, P. L. (2007). Marine mammal noise exposure criteria: Initial scientific recommendations. *Aquatic Mammals*, *33*(4), 411-521. http://dx.doi.org/10.1578/AM.33.4.2007.411

Szymanski, M. D., Bain, D. E., Kiehl, K., Pennington, S., Wong, S., & Henry, K. R. (1999). Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioral audiograms. *The Journal of the Acoustical Society of America*, *106*, 1134-1141. http://dx.doi.org/10.1121/1.427121

Teloni, V., Johnson, M. P., Miller, P. J. O., & Madsen, P. T. (2008). Shallow food for deep divers: Dynamic foraging behavior of male sperm whales in a high latitude habitat. *Journal of Experimental Marine Biology and Ecology*, *354*, 119-131. http://dx.doi.org/10.1016/j.jembe.2007.10.010

Thomas, J. A., Chun, N., Au, W. W. L., & Pugh, K. (1988). Underwater audiogram of a false killer whale (*Pseudorca crassidens*). *The Journal of the Acoustical Society of America*, *84*, 936-940. http://dx.doi.org/10.1121/1.396662

Tyack, P. L. (2009). Acoustic playback experiments to study behavioral responses of free-ranging marine animals to anthropogenic sound. *Marine Ecology Progress Series*, *395*, 187-200. http://dx.doi.org/10.3354/meps08363

Tyack, P. L., Gordon, J., & Thompson, D. (2004). Controlled-exposure experiments to determine the effects of noise on marine mammals. *Marine Technology Society Journal*, *37*, 39-51.

Tyack, P. L., Zimmer, W. M. X., Moretti, D., Southall, B. L., Claridge, D. E., Durban, J. W., . . . Boyd, I. L. (2011). Beaked whales respond to simulated and actual navy sonar. *PLoS One*, *6*, e17009. http://dx.doi.org/10.1371/journal.pone.0017009

White, M. J., Norris, J., Ljungblad, D., Baron, K., & Notarbartolo di Sciara, G. (1978). *Auditory thresholds of two beluga whales (*Delphinapterus leucas*)*. San Diego: Hubbs/Sea World Research Institute.

Yan, J. (2002). Geepack: Yet another package for generalized estimating equations. *R-News*, *2/3*, 12-14.

Yan, J., & Fine, J. P. (2004). Estimating equations for association structures. *Statistics in Medicine*, *23*, 859-880. http://dx.doi.org/10.1002/sim.1650

Zar, J. H. (1984). *Biostatistical analysis*. Englewood Cliffs, NJ: Prentice Hall.

Appendix A

In the following three tables, each exposure session is listed by the tag ID and sound exposure type. Behavioral changes that were judged by the expert panel to have been responses are numbered chronologically in brackets (), and the agreed severity score for each response is given. A short description of the change is listed, along with the time at which the change was judged to have begun. The maximum sound pressure level (SPL_{max}) and cumulative sound exposure level (SEL_{cum}) associated with the scored response are also reported in decibels. If no responses were judged to have occurred, the exposure session is listed with "No scored response" given as explanatory text. A full description of each experimental exposure and data plots are given in Miller et al. (2011).

Table AI. *Justification of Severity Scoring of Behavioral Responses of* ***Killer Whales***

Exp ID (#)	Severity	Justification, timing of change, and received levels
Oo06_317s LFAS		
(1)	6	Minor change in speed to avoid source; increase in speed and turn away from the path of the approaching source vessel, may also be a directional response. *Time:* Change occurred between two sightings, after the sighting indicating an increase in speed. Flow noise on the tag increased by more than 6 dB at 1431:30 h. *Dose:* SPL_{max}: 150, SEL_{cum}: 158
Oo06_327s and Oo06_327t MFAS		
(1)	3	Minor change in dive profile. *Time:* Started at inflection point within the final foraging dive at 1355:14 h. *Dose:* SPL_{max}: 139, SEL_{cum}: 141
(2)	7	Prolonged avoidance; tagged animal with its group moved sideways to oncoming source, then away from source for several hours. *Time:* Started at end of dive with inflection point at 1356:14 to 1356:25 h for Oo06_327s and Oo06_327t, respectively. *Dose:* SPL_{max}: 139, SEL_{cum}: 143
(3)	7	Prolonged cessation of feeding. Tailslaps were heard in the final dive, and not heard again in the tag record. The animals may have returned to feeding some time after end of exposure—calling was heard but not tailslaps. *Time:* Started at final tailslap at 1356:33 h. *Dose:* SPL_{max}: 139, SEL_{cum}: 144
(4)	5	Prolonged cessation of vocal behavior. *Time:* Started after set of four calls at 1404:08 h. *Dose:* SPL_{max}: 154, SEL_{cum}: 159
Oo08_149a MFAS 1 (Note there were two different MFAS exposures for this experiment.)		
(1)	2	An increased calling rate during approaches of duration less than the exposure period. Minor modification of vocal behavior. *Time:* Started at group of calls at 1252:58 h. *Dose:* SPL_{max}: 101, SEL_{cum}: 103
(2)	3	Brief and small change in speed, but not necessarily to avoid. Speed goes back to normal as soon as exposure stops. *Time:* Sightings were regularly taken; increase in speed on sighting at 1337:15 h. *Dose:* SPL_{max}: 142, SEL_{cum}: 148
Oo08_149a LFAS		
(3)	2	An increased calling rate during approaches, but much shorter than the duration of the exposure. Brief modification in vocal response. *Time:* Started at group of calls at 1524:16 h. *Dose:* SPL_{max}: 157, SEL_{cum}: 167

Oo08_149a Silent

(4) 4 Brief group avoidance of sound source as tagged animal moved somewhat out of the path of the oncoming source; duration of observation is limited by intermittent sightings. *Time:* Turn observed at sighting, 1705:03 h.

Oo08_149a killer whale sound playback

No scored response during this exposure session.

Oo08_149a MFAS 2 (Note there were two different MFAS exposures for this experiment.)

(5) 5 Minor avoidance of sound source as tagged animal with its group (except the calf) crossed to opposite side of fjord. The avoidance is likely to have been affected both by the calf separation and the narrow shape of the fjord. *Time:* Was between sightings; turn recorded on magnetometer during dive at 2241:49 h. *Dose:* SPL_{max}: 133, SEL_{cum}: 133

(6) 5 Prolonged overall change in vocal behavior, including production of high-frequency whistles (apparently related to the calf separation). *Time:* Started at group of calls at 2253:34 h. *Dose:* SPL_{max}: 144, SEL_{cum}: 153

(7) 3 Minor change in speed and direction (slow down and turn) but not to avoid. This change may also have been related to the calf which was separated from, and behind, the main group of animals. *Time:* Scored at minimum speed point with turnaround at 2255:53 h. *Dose:* SPL_{max}: 144, SEL_{cum}: 154

(8) 8 Prolonged separation of dependent calf. Calf was sighted travelling alone for 86 min but had never been seen travelling alone for over 30 h of prior observation. Disruption of acoustic reunion mechanisms was judged to have occurred on the basis of the presence of high-frequency (> 20 kHz) whistles suggesting attempted compensation for interference from the sonar. *Time:* Scored at time calf was first sighted alone, 2302:00 h. *Dose:* SPL_{max}: 152, SEL_{cum}: 159

Oo09_144ab 1 LFAS

(1) 6 Prolonged modification of vocal behavior, with groups of calls synchronized with the arrival of sonar pings. *Time:* Scored at start of intensive calling, 1413:15 h. *Dose:* SPL_{max}: 91, SEL_{cum}: 89

(2) 7 Dramatic increase in swim speed and change in direction sideways to the oncoming source. Prolonged group avoidance. *Time:* Started between sightings, taken at point of increased fluke stroking between 2nd and 3rd sonar pings, 1413:30 h. *Dose:* SPL_{max}: 94, SEL_{cum}: 94

(3) 5 Prolonged change in dive behavior; deep foraging dives ceased. *Time:* Taken at the start of the shallow dive just after the start of exposure, 1413:34 h. *Dose:* SPL_{max}: 94, SEL_{cum}: 94

(4) 7 Prolonged cessation of feeding. Tailslaps were heard in deep dive just prior to the exposure but never again throughout the tag deployment. *Time:* Taken at the time of the change in diving behavior, 1413:34 h. *Dose:* SPL_{max}: 94, SEL_{cum}: 94

(5)	5	Surface displays stop associated with increased surface synchrony; prolonged change of group spacing (tighter). *Time:* Scored at sighting with tighter spacing, 1420:01 h. *Dose:* SPL_{max}: 152, SEL_{cum}: 156

Oo09_144ab MFAS

(6)	6	Turn away from the oncoming source with occasional turns sideways to approaching source during exposure period. Moderate increase in swim speed and change in direction to avoid source. *Time:* Taken at first turning point, 1617:40 h. *Dose:* SPL_{max}: 92, SEL_{cum}: 93
(7)	5	Overall tighter group spacing and high surface synchrony which extended beyond the end of the exposure period. Prolonged change in group spacing. *Time:* Taken at time when both are at "1," 1623:41 h. *Dose:* SPL_{max}: 116, SEL_{cum}: 119

Oo09_144ab killer whale sound playback

(8)	2	Multiple deeper dives were judged to be orientation responses, possibly to find better sound paths to listen. *Time:* Taken at start of 1st deeper dive, 1901:39 h.
(9)	3	Moderate change in movement direction and speed but not avoidance. *Time:* Taken at sighting with clear direction change, 1928:51 h.

Oo09_144ab LFAS-DS

(10)	5	Striking change in group spacing (tighter) and surface synchrony (increased). *Time:* Taken at sighting after deep dive when value were both "1," 2133:07 h. *Dose:* SPL_{max}: 163, SEL_{cum}: 167
(11)	6	Increased swim speed and change of direction away and sideways to the oncoming source of same duration as the exposure. Moderate avoidance. *Time:* Taken at sighting following deep dive, 2133:10 h. *Dose:* SPL_{max}: 163, SEL_{cum}: 167

Table AII. *Justification of Severity Scoring of Behavioral Responses of* ***Long-Finned Pilot Whales***

Exp ID (#)	Severity	Justification, timing of change, and received levels
Gm08_150c MFAS		
(1)	5	Production of calls similar to the sonar (vocal matching) and reduction of calling rate. Modification and prolonged cessation of vocal behavior. *Time:* Taken at end of calls during exposure, 1621:33 h. *Dose:* SPL_{max}: 115, SEL_{cum}: 118
(2)	6	Clear turn away then sideways from the oncoming source vessel returned to pre-exposure direction of travel. Moderate avoidance of sound source by tagged whale and its group. *Time:* Taken at sighting following deep dive, 1624:55 h. *Dose:* SPL_{max}: 124, SEL_{cum}: 127
Gm08_150c LFAS		*Note:* Small tourist boat close to the animals
(3)	3	Minor change in speed and direction near the point of closest approach by the source vessel but not characteristic of avoidance. *Time:* Taken at sighting with increase in speed and start of erratic direction, 1820:39 h. *Dose:* SPL_{max}: 159, SEL_{cum}: 162
Gm08_154d LFAS		
(1)	5	Intermittent change in direction and speed at the closest point of approach coincided with a rapid drop in the received level of the sonar pings on the tags. Judged to be minor avoidance. *Time:* Taken at point where speed reaches zero 1st time, 0223:38 h. *Dose:* SPL_{max}: 159, SEL_{cum}: 166
Gm08_154d MFAS		No scored response. *Max exposure in session:* SPL_{max}: 152, SEL_{cum}: 153
Gm08_158b		*Note:* No dtag data; tag failure.
GM08_158b Silent		
(1)	3	Minor change in locomotion and speed (the loop in the track plot) not judged to be avoidance of the source. *Time:* Taken at point where direction changes, 1510:15 h.
Gm08_158b LFAS		No scored response. Max exposure in session calculated using 19log(range), corresponding to TL fit during MFAS exposure sessions. Range calculated as distance from sighting location of whale to the source vessel. *Max exposure in session:* SPL_{max}: 172, SEL_{cum}: 179
Gm08_158b MFAS		
(2)	5	Minor avoidance; turned away from the path of the oncoming source. *Time:* Taken at point where direction and speed starts to change, 1816:00 h. *Dose:* SPL_{max}: 153, SEL_{cum}: 155

Gm08_159a Silent No scored response.

Gm08_159a LFAS

(1) 3 A minor change in direction and speed (reduced), not judged to avoidance. *Time:* Taken at point where direction and speed starts to change, 0100:23 h. *Dose:* SPL_{max}: 160, SEL_{cum}: 168

(2) 4 Brief vertical avoidance. The tagged whale made four synchronous surfacings with the arrival of sonar pings which corresponded to a dramatic (> 20 dB) drop in RL on the tags. Not visible on the plot without zoom. *Time:* Taken at point after the last ping before 1st synchronous surfacing, 0104:50 h. *Dose:* SPL_{max}: 175, SEL_{cum}: 176

Gm08_159a MFAS

(3) 1 Brief period of ascent during descent at CPA—judged to represent an orientation response similar to that observed in Sw09_160a LFAS-DS. Not visible on the plot without zoom. *Time:* Taken at point of the inflection, 0241:44 h. *Dose:* SPL_{max}: 159, SEL_{cum}: 159

Gm08_159a killer whale sound playback

(4) 5 Moderate shift (decrease) in group spacing of similar duration as the playback. *Time:* Taken at the 1st sighting where spacing remained at a value of one, 0458:27 h.

(5) 4 Moderate change in speed and direction but not avoidance (movement was towards the source and, therefore, attraction); roughly of same duration as the playback. *Time:* Taken at the 1st sighting approaching source, 0516:02 h.

Gm09_138ab LFAS

(1) 4 A brief avoidance; change in direction and speed (reduced) away from the path of the approaching source ship. *Time:* Taken at the 1st sighting with speed < 1 m/s, 1459:15 h. *Dose:* SPL_{max}: 156, SEL_{cum}: 164

(2) 3 Brief change in group spacing or social cohesion in synchrony with the avoidance. *Time:* Taken at the 1st sighting with group spacing = 1, 1505:36 h. *Dose:* SPL_{max}: 161, SEL_{cum}: 168

Gm09_138ab MFAS No scored response. *Max exposure in session:* SPL_{max}: 163, SEL_{cum}: 164

Gm09_138ab Silent

(3) 5 Minor avoidance; change of direction to avoid the path of the source of duration less than that of the exposure session. *Time:* Taken at the sighting with direction change = 1, 1904:29 h.

Gm09_138ab LFAS-DS

(4) 6 Moderate cessation of feeding as benthic foraging dive behaviors ceased. *Time:* Taken at the bottom of the short dive which would have fallen where a deep dive would have been expected, 2035:46 h. *Dose:* SPL_{max}: 145, SEL_{cum}: 144

(5)	4	Moderate change in dive behavior. *Time:* Taken at the bottom of the short dive which would have fallen where a subsequent deep dive would have been expected, 2035:46 h. *Dose:* SPL_{max}: 145, SEL_{cum}: 144
(6)	4	Moderate change in vocal behavior, reduced number of buzzes but clicking continues at the surface. Apparent vocal mimicry during exposure. *Time:* Taken at time of 1st high-quality matching sound, 2044:53 h. *Dose:* SPL_{max}: 154, SEL_{cum}: 158
Gm09_138ab killer whale sound playback		
(7)	4	Moderate change in dive profile as animal started deep diving. *Time:* Taken at the start of the 1st deep dive, 0224:31 h.
Gm09_156b Silent		
(1)	3	Brief decrease in group spacing towards the start of approach. *Time:* Taken at 1st sighting with group spacing = 2, 2330:06 h.
Gm09_156b LFAS		
(2)	6	Prolonged change in group distribution (reduced group spacing). (Given the sampling times, this change could have occurred before the exposure started but was judged to plausibly be related to the simultaneous change in diving activity.) *Time:* Taken at 1st sighting with group spacing = 2, 0136:54 h. *Dose:* SPL_{max}: 82, SEL_{cum}: 80
(3)	5	Prolonged change in dive behavior from active deep diving to surface mode at onset of exposure. *Time:* Taken 6 min after the end of the previous deep dive, which is a typical surfacing interval between deep dives, 0142:49 h. *Dose:* SPL_{max}: 152, SEL_{cum}: 153
(4)	7	Prolonged cessation of feeding; cessation of deep foraging dives exceeds exposure duration. *Time:* 6 min after the end of the previous deep dive, 0142:49 h. *Dose:* SPL_{max}: 152, SEL_{cum}: 153
(5)	6	Ship approached from the front. Tagged animal with its group moved sideways from the ship and then directly away from ship. Moderate avoidance. RL continue to increase because the source is moving faster than the animals. *Time:* Taken at 1st sighting with change in direction, 0155:41 h. *Dose:* SPL_{max}: 158, SEL_{cum}: 167
(6)	4	Brief vertical avoidance (synchronous surfacing with sonar not visible without zoom). *Time:* Taken just before 1st overlap surfacing, 0203:08 h. *Dose:* SPL_{max}: 180, SEL_{cum}: 185
Gm09_156b MFAS		
(7)	4	Moderate change in dive behavior from shallow to deep diving of same duration as the exposure. *Time:* Start of 1st deep dive, 0310:28 h. *Dose:* SPL_{max}: 85, SEL_{cum}: 87
Gm09_156b LFAS-DS		
(8)	6	Moderate avoidance; ship approached from the front. Animals moved sideways from the ship and then 180° from ship, similar to pattern in LFAS exposure session. RL continue to increase because the source is moving

		faster than them. The 2nd avoidance movement is judged to be a continuation of the 1st one. *Time:* Taken just before 1st overlap surfacing, 0511:43 h. *Dose:* SPL_{max}: 159, SEL_{cum}: 168

Gm09_156b killer whale sound playback

(9)	5	Prolonged change in movement direction but not avoidance (animals were apparently attracted to source location). *Time:* Taken at 1st sighting with change in direction, 0702:09 h.

Table AIII. *Justification of Severity Scoring of Behavioral Responses of* ***Sperm Whales***

Exp ID	Severity	Justification, timing of change, and received levels
Sw08_152a MFAS		
(1)	6	Moderate avoidance as animal turned sideways from oncoming source. *Time:* Taken at direction change at start of deep dive, 0150:14 h. *Dose:* SPL_{max}: 128, SEL_{cum}: 129
Sw08_152a LFAS		
(2)	1	Brief orientation response (wiggle on ascent phase of dive profile in the very beginning of exposure). *Time:* Taken at point when whale starts to descend, 0411:34 h. *Dose:* SPL_{max}: 132, SEL_{cum}: 129
(3)	4	Moderate change in vocal behavior (no clicking during deep dives). *Time:* Taken 20 s after start of dive—previous two surfacing clicks started 22.5 and 16.5 s after dive start, 0418:31 h. *Dose:* SPL_{max}: 156, SEL_{cum}: 156
(4)	6	Moderate cessation of feeding indicated by lack of echolocation clicks or buzzes during dives. *Time:* Taken 20 s after start of dive—previous two surfacing clicks started 22.5 and 16.5 s after dive start, 0418:31 h. *Dose:* SPL_{max}: 156, SEL_{cum}: 156
Sw09_141a LFAS		
(1)	1	Brief orientation responses; dive ascent with a simultaneous change of direction. *Time:* Taken at wiggle in dive profile, 1221:44 h. *Dose:* SPL_{max}: 140, SEL_{cum}: 144
(2)	4	Social sounds (codas and slow clicks) during exposure also happened before but the rate increases. Moderate modification of vocal behavior. *Time:* Taken at start of slow click bout, 1224:00 h. *Dose:* SPL_{max}: 141, SEL_{cum}: 147
(3)	2	Wiggle on the dive profile (1230 h); brief change of dive profile. *Time:* Taken at start of re-descent, 1229:54 h. *Dose:* SPL_{max}: 158, SEL_{cum}: 161
(4)	2	Brief change in vocalization (he started clicking for a brief period during the wiggle in the dive profile). *Time:* Taken at start of re-descent, 1229:54 h. *Dose:* SPL_{max}: 158, SEL_{cum}: 161
(5)	6	Moderate avoidance of sound source as animal turned sideways to the source. The direction change coincided with the wiggle in ascent dive profile. *Time:* 1229:54 h. *Dose:* SPL_{max}: 158, SEL_{cum}: 161
Sw09_141a MFAS		
(6)	6	Moderate avoidance of sound source as animal turned sideways to the path of the oncoming source. *Time:* Direction change in heading data to 270° is indicated at the start of deep dive, 1419:23 h. *Dose:* SPL_{max}: 135, SEL_{cum}: 138
(7)	4	The social sounds (codas and slow clicks) during exposure also happened before but the rate increases. Moderate modification of vocal behavior. *Time:* Taken at start of slow-click bout, 1445:56 h. *Dose:* SPL_{max}: 150, SEL_{cum}: 158

Sw09_141a Silent		No scored response during exposure session. The animals separated out at the end of the pre-exposure period, but this was judged to have occurred prior to the start of the deep dive in pre-exposure as sperm whales typically separate during dives. No other changes in behavior that were putative responses were detected.
Sw09_141a killer whale sound playback		
(8)	4	Regular clicking becomes intermittent (buzzes indicate foraging continued), and slow clicking rate increases. Moderate change in vocal behavior. *Time:* Taken at point where regular clicking stops during descent, 1902:04 h.
(9)	7	Change in direction; moderate avoidance. *Time:* Taken at point when direction stabilizes to 90°, 1915:38 h.
Sw09_142a Silent		
(1)	1	A brief turn towards ship at CPA which we interpreted as an orientation response. *Time:* Taken at point where direction changed towards ship, 2016:51 h.
Sw09_142a LFAS		
(2)	4	Moderate change in dive behavior with abnormal dive profile, with strong changes in depth during prolonged ascent phase with no clicking. *Time:* Taken at point where unusual dive starts, 2148:09 h. *Dose:* SPL_{max}: 120, SEL_{cum}: 123
(3)	6	Moderate avoidance of sound source as animal turned sideways to path of oncoming source. *Time:* Also begins at start of the unusual dive, 2148:09 h. *Dose:* SPL_{max}: 120, SEL_{cum}: 129
(4)	1	Orientation response midway through the exposure (during avoidance); sharp turn in horizontal movement seemed to be associated with the wiggles in the dive profile. *Time:* Taken at time of the 1st turn, 2158:20 h. *Dose:* SPL_{max}: 154, SEL_{cum}: 159
Sw09_142a MFAS		No scored response. *Max exposure in session:* SPL_{max}: 146, SEL_{cum}: 156
Sw09_142a killer whale sound playback		
(5)	4	Moderate change in dive behavior. *Time:* Taken at point where normal resting dive is changed, 0235:10 h.
(6)	6	Interruption of resting dives (whale rests later in the tag record) leads to initiation of feeding after exposure (thus severity difficult to judge). Considered to be a moderate cessation of resting as foraging is an important life function. *Time:* Taken at point where normal resting dive is changed, 0235:10 h.
(7)	4	Moderate change in locomotion (increased speed) but not avoidance of the source. *Time:* Time of 1st speed increase in TR plot, 0235:40 h.
(8)	1	Spy hop (brief orientation response). *Time:* Time of spyhop in dive record, 0237:04 h.

Sw09_142a LFAS-DS		
(9)	4	Moderate change in vocal behavior (silence). *Time:* Taken when clicking stops, 0409:47 h. *Dose:* SPL_{max}: 141, SEL_{cum}: 147
(10)	6	Moderate cessation of foraging; switch to travel. *Time:* When clicking stops, 0409:47 h. *Dose:* SPL_{max}: 141, SEL_{cum}: 147
(11)	3	Minor change in locomotion direction; not avoidance. *Time:* Corresponds with time when clicking stops, 0409:47 h. *Dose:* SPL_{max}: 141, SEL_{cum}: 147
(12)	4	Moderate change in dive profile (breaks off descent of foraging dive and switches to resting dives). *Time:* Time when dive inflects, 0410:32 h. *Dose:* SPL_{max}: 141, SEL_{cum}: 147
(13)	3	Another minor change in locomotion direction, not avoidance, at closest point of approach. *Time:* Time of turn late in the exposure, 0447:37 h. *Dose:* SPL_{max}: 154, SEL_{cum}: 166

Sw09_160a MFAS		No scored response. *Max exposure in session:* SPL_{max}: 151, SEL_{cum}: 157
Sw09_160a LFAS		
(1)	4	Moderate change in dive profile; unusual shallow dives during surfacing period. *Time:* Time of 1st shallow surfacing dive, 1447:10 h. *Dose:* SPL_{max}: 126, SEL_{cum}: 126
(2)	4	Brief avoidance of sound source; whale turned around to move away from path of oncoming source. *Time:* Time of turn during exposure, 1514:15 h. *Dose:* SPL_{max}: 161, SEL_{cum}: 167
(3)	4	Moderate change of vocal behavior (less clicking and no buzzes during exposure). *Time:* Time of 1st gap in clicking during ascent 1515:34 h. *Dose:* SPL_{max}: 161, SEL_{cum}: 168
(4)	6	Aborted feeding dive with little clicking during ascent. Moderate cessation of feeding; switch to travel. *Time:* Time of 1st gap in clicking during ascent, 1515:34 h. *Dose:* SPL_{max}: 161, SEL_{cum}: 168
Sw09_160a killer whale sound playback		
(5)	2	Multiple wiggles in the descent phase of the dive; judged to be orientation responses, not avoidance. *Time:* Time of first wiggle, 1825:42 h.
Sw09_160a LFAS-DS		
(6)	1	Brief orientation response. Wiggle on the dive profile; the rest of the dive looks like a normal foraging dive. *Time:* Time of 1st wiggle, 2014:02 h. *Dose:* SPL_{max}: 110, SEL_{cum}: 110
(7)	4	Brief avoidance at CPA (change of direction away from source). *Time:* Time of direction change at start of descent, 2106:24 h. *Dose:* SPL_{max}: 166, SEL_{cum}: 176