

Contribution to the Theme Section '21st century paradigms for measuring and managing the effects of anthropogenic ocean noise'



REVIEW

Experimental field studies to measure behavioral responses of cetaceans to sonar

Brandon L. Southall^{1,2,3,*}, Douglas P. Nowacek^{3,4}, Patrick J. O. Miller⁵,
Peter L. Tyack⁵

¹Southall Environmental Associates, Aptos, CA 95003, USA

²University of California Santa Cruz, Institute of Marine Sciences, Long Marine Laboratory Santa Cruz, CA 95060, USA

³Duke University Marine Laboratory, Duke University, Beaufort, NC 28516, USA

⁴Pratt School of Engineering, Duke University, Beaufort, NC 27705, USA

⁵Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, Fife KY16 9LB, UK

ABSTRACT: Substantial recent progress has been made in directly measuring behavioral responses of free-ranging marine mammals to sound using controlled exposure experiments. Many studies were motivated by concerns about observed and potential negative effects of military sonar, including stranding events. Well-established experimental methods and increasingly sophisticated technologies have enabled fine-resolution measurement of many aspects of baseline behavior and responses to sonar. Studies have considered increasingly diverse taxa, but primarily odontocete and mysticete cetaceans that are endangered, particularly sensitive, or frequently exposed to sonar. This review focuses on recent field experiments studying cetacean responses to simulated or actual active military sonars in the 1 to 8 kHz band. Overall results demonstrate that some individuals of different species display clear yet varied responses, some of which have negative implications, while others appear to tolerate relatively high levels, although such exposures may have other consequences not measured. Responses were highly variable and may not be fully predictable with simple acoustic exposure metrics (e.g. received sound level). Rather, differences among species and individuals along with contextual aspects of exposure (e.g. behavioral state) appear to affect response probability. These controlled experiments provide critically needed documentation of identified behavioral responses occurring upon known sonar exposures, and they directly inform regulatory assessments of potential effects. They also inform more targeted opportunistic monitoring of potential responses of animals during sonar operations and have stimulated adaptations of field methods to consider increasingly realistic exposure scenarios and how contextual factors such as behavioral state and source proximity influence response type and probability.

KEY WORDS: Behavioral response study · Sonar · Marine mammals · Controlled exposure experiment

INTRODUCTION

Early ethologists recognized that baseline observations provide key insights for developing hypotheses about animal behavior and designing experiments to test them. Recognizing the need to separate anec-

dotal observation from controlled measurements, and building upon experimental methods in other scientific disciplines, pioneering researchers developed techniques to empirically document causal links between exposure to stimuli and behavioral responses (Lorenz 1937, Tinbergen 1963). These and other

*Corresponding author: brandon.southall@sea-inc.net

© The authors 2016. Open Access under Creative Commons by Attribution Licence. Use, distribution and reproduction are unrestricted. Authors and original publication must be credited.

early studies established experimental methods as an essential foundation for studying animal behavior. Increasingly sophisticated experimental methods have been derived and applied to understand marine mammal behavior for over half a century (e.g. experimental demonstration of echolocation in dolphins by Norris et al. 1961).

Experimental presentations of external stimuli to wild and laboratory animals have documented myriad aspects of behavioral ecology, such as how animals recognize one another, compete and succeed in mating, select feeding or breeding habitat, and respond to various stimuli (see Bradbury & Vehrencamp 1998 for a review). For species that rely heavily on sound to perform vital functions, including most vertebrates and all marine mammals, experimental methods have been fundamentally important in understanding how and why they use sound and how they may respond to it.

In recent decades, interest has grown concerning how cetacean behavior may be affected by active sonar, particularly military systems operating in the lower-frequency (~0.1–2 kHz) and mid-frequency bands (2–8 kHz). Research on possible behavioral effects of the US Navy's SURTASS low frequency active (LFA) (~0.1–0.5 kHz band) sonar (Miller et al. 2000, Fristrup et al. 2003) was initiated given the relatively high source level of this low frequency sonar, which can propagate large distances. Public concern about the possible behavioral impact of sonar on killer whales was heightened following an unintentional sonar exposure of the endangered (US Endangered Species Act [ESA], listed in 2005) southern resident population of killer whales in Haro Strait, WA, USA in 2003 (Southall & Gentry 2005). In a similar event, multi-national naval exercises in 2000 were associated with reduced killer whale presence in the Vestfjorden basin of Norway (WWF-Norway 2001). But despite these concerns and short-term observations of response, the National Research Council Ocean Studies board (NRC 2005) concluded that there were no data or methods available to quantify possible population consequences of these kinds of disturbances.

Perhaps the strongest motivation for additional behavioral research arose from a series of lethal cetacean strandings coinciding with active sonar (~0.5–8 kHz) testing and training. This phenomenon was first noted from mass strandings in the Canary Islands and Greece (Simmonds & Lopez-Jurado 1991, Frantzis 1998) and became more widely recognized following a well documented, highly publicized mass stranding in the Bahamas in 2000 (Bal-

comb & Claridge 2001), for which the US Navy acknowledged a causal association (Evans & England 2001). These and subsequent events, including continuing strandings in the Mediterranean (e.g. Tethys 2014, www.tethys.org/tethys/strandedwhalesupdate/), have raised awareness, interest, and debate within scientific, military, regulatory, and environmental communities about the nature and extent of the association between active sonar and strandings. Broad analyses have identified other potential sonar-associated strandings globally, observing commonalities and raising key questions, including the potential role of behavioral responses (Brownell et al. 2004, Cox et al. 2006, Filadelfo et al. 2009).

Sonar-associated strandings have typically included multiple animals found within hours to several days, spread over tens of kilometers, in places with deep water near shore, and have predominately involved several species of deep-diving, pelagic beaked whales (family Ziphiidae), most commonly Cuvier's beaked whale *Ziphius cavirostris* (Cox et al. 2006, Filadelfo et al. 2009). The Bahamas 2000 stranding involved some of the most powerful tactical military mid-frequency active sonar (MFAS) systems (SQS-53C) with primary energy in the 2.5–5 kHz range (as did a number of others based on unpublished reports). However, a lower-frequency (0.45–0.7 kHz) sonar system was involved in the 1998 Greece stranding (Frantzis 1998, D'Amico et al. 2009, Filadelfo et al. 2009), as well as a 3 kHz sonar transmitted at the same time. In some instances, physical trauma associated with the formation of gas or fat emboli have been documented in sonar-associated strandings (e.g. Jepson et al. 2003, Fernández et al. 2005). The underlying cause of injury remains unknown. However, from the relatively tight temporal and broad spatial distribution pattern of most strandings and the quite high exposure levels required to induce changes in sensitive hearing systems in most cetaceans tested in laboratories (see Finneran 2015), it is unlikely that they resulted directly from physiological consequences of exposure (e.g. tissue damage caused directly by sound). Stranding patterns suggest that other less direct mechanisms were involved, unless by some unlikely coincidence all animals involved were within a few hundreds of meters of sound sources. Specifically, exposure to sonar may have resulted in behavioral responses that ultimately resulted in a cascade of events leading to lethal stranding, potentially including decompression injuries caused by alteration of dive behavior (Cox et al. 2006). Anecdotal observations of strong and potentially harmful behavioral responses in several other

(non-beaked whale) cetaceans associated with or identified as likely resulting from exposure to approximately 2–8 kHz tactical sonars have provided additional support for the notion that behavioral reactions may play a key role in the sequence of events leading to injury or mortality (Southall & Gentry 2005, Southall et al. 2006).

Given a general lack of empirical data on cetacean behavioral responses to sonar systems involved in stranding events, focused research was clearly needed. This was explicitly identified by international scientific and governmental regulatory bodies in the mid-to late 2000s. These recommended controlled experiments to measure behavioral response in cetaceans, including but not limited to beaked whales, to improve the scientific basis for understanding how they may respond to and potentially be harmed by active sonar (ICES 2005, NRC 2005, IACMST 2006, Nowacek et al. 2007, Southall et al. 2007, 2009). Efforts to formulate science-based noise exposure criteria (e.g. Southall et al. 2007) were motivated by the need to improve regulatory assessments and mitigation of potential harm from human sound sources, including sonar. This increased interest and concern has resulted in major advances in both basic and applied scientific understanding of marine mammal behavior and potential effects of noise.

A series of independent but cross-pollinating international research efforts have focused on cetacean behavioral responses to different sonar systems. Studies have applied well established scientific principles to develop robust and adaptive experimental methods. The methods, results, and broad conclusions from a decade of focused research, financially supported largely by the US, Norwegian, and Dutch defense organizations, are described here. These studies have substantial scientific and regulatory implications for endangered, threatened, and protected species.

EVOLUTION OF EXPERIMENTAL APPROACHES AND METHODS

Experimental approaches using controlled exposure experiments (CEEs) to study marine mammal behavioral responses to noise have been derived from traditional sound playback experiments to study natural communication. These experiments involve the presentation of natural signals and often synthetic control signals in different contexts (e.g. known versus unfamiliar conspecifics) to investigate aspects of behavior such as individual recognition, territoriality,

or parental attendance of young (e.g. McGregor 1992, 2013, Falls 1992, Hopp & Morton 1997). Results from naturalistic observations informed by empirical playback experiments have formed the foundation of much of what is known about communication systems and behavior in many taxa, including birds (e.g. Falls et al. 1982), terrestrial mammals (e.g. Cheney et al. 1995), and fish (e.g. McGregor 1992) as well as marine mammals (e.g. Tyack 1983, Janik et al. 2006, Holt et al. 2010). Playback methods have been adapted and applied to free-ranging marine mammals using CEEs to test potential responses to various human noise sources, including examples involving low frequency coded signals (Frankel & Clark 2000), LFA sonar (Miller et al. 2000, Fristrup et al. 2003), seismic airgun surveys (Miller et al. 2009, Cato et al. 2013, Dunlop et al. 2015), and low-level tonal signals (Nowacek et al. 2004, Dunlop et al. 2013). Their application to studying behavioral responses to naval sonar in the 1–8 kHz range is an area of active research considered here.

The basic approach involves measuring aspects of individual behavior within ‘pre-exposure baseline’ (no stimulus), ‘exposure’ (controlled stimulus presentation using specified protocols), and ‘post-exposure’ (after stimulus) periods (e.g. Tyack et al. 2003). During exposure, received sound at the experimental subject is controlled in order to meet research objectives. This involves a presentation of exposure stimuli with specific characteristics, often including an escalation in the exposure level to identify whether and when behavioral response(s) may occur. Behavior during exposure is compared with baseline behavior to identify potential change points. Post-exposure behavior is evaluated in terms of whether and when behavior returns to baseline.

Methodological approaches to measuring marine mammal behavioral response

Early experiments were conducted as part of the US Navy’s SURTASS-LFA scientific research program. Miller et al. (2000) demonstrated changes in humpback whale song durations during exposure to a similar system by visually following individual singers and listening with towed hydrophone arrays. Croll et al. (2001) also studied responses of foraging fin and blue whales to the same sound source, showing that whale behavior appeared more strongly linked to prey than sonar exposure. Tyack (2009) demonstrated localized avoidance of an active (0.1–0.5 kHz) sonar source by migrating grey whales using shore-

based visual tracking methods. However, broad application of these observational methods and the types of behavior that can be studied are limited.

In conjunction with increasing research effort in the early 2000s, animal-attached archival tags that record fine-scale individual movement and broadband acoustic data were developed (Johnson & Tyack 2003). These tags contain pressure sensors to measure depth, 3-axis magnetometers and accelerometers to record movement, and hydrophones to measure vocal behavior of the tagged animal along with environmental sounds. The behavioral and noise exposure recording capability of these tags combined with the ability to follow tagged individuals using integrated radio beacons were critical technological developments enabling subsequent experimental studies.

These novel tags quickly enabled researchers to record fine-scale diving, vocalization, and movement behavior in cetaceans (e.g. Nowacek et al. 2004, Johnson et al. 2006). These and other studies revealed aspects of behavior impossible to measure with surface observations, passive listening, or lower-resolution archival or satellite-linked tags. Their successful application on many species, including the poorly known, deep-diving beaked whales that were the subject of concern regarding sonar-associated strandings (Johnson et al. 2006, Tyack et al. 2006), provided critically needed baseline behavioral data. Focal study individuals (Altmann 1974) could be followed using tag-specific radio beacons, enabling measurement of responses within an experimental framework comparing the focal individual before, during, and after exposure (e.g. Nowacek et al. 2004, Miller et al. 2009). Other sensors (e.g. Argos, GPS; see Costa et al. 2010) or visual focal-follow methods describing individual or social behaviors measured (e.g. Visser et al. 2014) have been used with archival tags to provide complementary data.

A detailed understanding of baseline (undisturbed) behavior patterns from these tags has been crucial in CEE design and in meaningfully interpreting behavior and potential response. Baseline studies have focused on key behavioral parameters likely to be important in detecting potential behavioral changes during CEEs. For instance, Tyack et al. (2011) used baseline data from tagged Blainville's beaked whales *Mesoplodon densirostris* to design research protocols to present sonar exposure during a particular part of the dive cycle in which baseline behavior seldom changed. They then compared numerous dive parameters in intentionally unexposed Blainville's beaked whales to those in sonar CEEs using identical high-resolution behavioral metrics. Similarly, studies of

diving kinematics and feeding strategies of blue whales *Balaenoptera musculus* in relation to prey distribution (Friedlaender et al. 2015, Goldbogen et al. 2015, Hazen et al. 2015) have provided critical insight into the interpretation of responses of feeding whales during CEEs (Friedlaender et al. 2016). Furthermore, baseline acoustic data from the tags have contributed to understanding the typical animal distribution in study areas before, during and after exposure to sound (McCarthy et al. 2011, Tyack et al. 2011, Yack et al. 2013). Finally, novel methods to identify vocalizing individuals within groups are enabling baseline vocalization measurements and examinations of potential effects of noise on calling and echolocation (e.g. Goldbogen et al. 2014, Stimpert et al. 2015, Arranz et al. 2016).

The combination of focal individual sampling with high-resolution tags to measure fine-scale movement, sound production, and sound exposure within an experimental context poses some analytical challenges. For instance, studies are often limited by small sample sizes, although this depends on the relative efficacy of locating and tagging subjects. Species such as baleen whales that are relatively amenable may have larger sample sizes (e.g. Goldbogen et al. 2013, Sivle et al. 2015), but for more challenging subjects, including the high-priority beaked whales, sample sizes have remained small (e.g. DeRuiter et al. 2013a, Stimpert et al. 2014, Miller et al. 2015). Additionally, the multivariate nature of rich time-series behavioral data, differences in temporal resolution, and differential variability in behavioral parameters, pose complex statistical challenges. Sustained development of associated analytical methods to deal with these challenges, through interactions between tag designers, acousticians, field biologists, and statisticians, has significantly improved CEE analyses. These have included the application of traditional and novel statistical methods, including expert elicitation, generalized linear mixed models, general estimating equations, time-series analysis statistical techniques, and state-space analyses (see Harris et al. 2015, 2016).

Methodological considerations of sonar CEEs

The interaction of scientific objectives with the logistical realities of obtaining behavioral response data from free-ranging cetaceans in increasingly realistic sonar exposure conditions has shaped CEE experimental design and methodology. These present a number of practical challenges, including con-

ducting multiple CEE replicates on the same wild individuals to test for potentially different effects of stimulus type, repeated exposure, variable received level or other exposure parameters. In laboratory settings, these include issues of habituation and/or sensitization in addition to the very different exposure context and consequences relative to free-ranging animals. But it is often difficult to re-acquire and test wild individuals multiple times and there are challenges in interpreting order effects where sequential noise exposures are presented within individuals over short durations (e.g. Tyack et al. 2011). Exposure dose escalation is often used where aspects of exposure (e.g. received sound levels) are increased within presentations to each individual subject by increasing the source level or by moving sound sources transmitting at a constant level to identify the lowest exposure condition at which the onset of a particular response may occur. Multiple individuals are then tested in separate experiments or using sequences of repeated exposures and controls to determine exposure conditions within which behavioral responses may occur and to describe intra-individual variability.

Numerous experimental and methodological considerations drive operational aspects of cetacean sonar CEEs. These broad considerations, which depend on species life history, general movement patterns, social structure, and geographic location, include at least the following:

(1) Sonar source type, including frequency spectra, source level, signal modulation pattern (including rise time and duty cycle), and overall duration.

(2) Source–subject range and relative movement, determined by source capabilities, desired received levels, and whether level escalation is used and is achieved by increasing stationary source level or source movement or both.

(3) Mitigation requirements of permitted research, which may include source level ramp-up (which has been required as a permit condition in some experiments and is consistent with exposure level-escalation objectives), source shut-down zones, acoustic detection of sensitive species such as beaked whales during exposure, observations for aberrant behavior that indicate a risk of harm to the subjects, and other considerations.

(4) Type of control stimuli used, which may include no-sonar control sequences, signals with acoustic features similar to those of interest (e.g. pseudo-random noise at overlapping frequencies), or presumably salient biological signals (e.g. sounds of conspecifics or predators).

(5) Whether or not the same or multiple stimuli are repeated during a CEE sequence.

These methodological considerations have been approached in slightly different ways across the diverse experimental sonar exposure studies that have been conducted over the past decade, and they have collectively yielded a large, increasing body of published results despite acknowledged inherent challenges.

Recent and ongoing marine mammal sonar CEEs

Several recent research programs from around the world have resulted in major advances in studying cetacean behavioral responses to sonar. The application of high-resolution, multi-sensor tags within a CEE context in these studies has allowed direct measurements of potential behavioral responses that are directly applicable to and are being incorporated into ongoing regulatory assessments regarding military sonar. These studies have some common goals, objectives, and methods, with considerable overlap in researchers, including the authors of this review. Each study is discussed here in some detail, as collectively they represent much of the recent progress in this field.

AUTEC-BRS

Tyack et al. (2011) studied Blainville's beaked whale behavior on and around a navy training and testing range in the Bahamas called the Atlantic Undersea Testing and Evaluation Center (AUTEC). This was the first study to measure beaked whale behavioral responses to simulated and real naval sonar, including the potential mechanisms underlying the adverse effects of specific active military sonar systems (AN/SQS-53C and AN/SQS-56C tactical sonars) that use 2–8 kHz sonar signals and were involved in previous mass stranding events. It was supported by the US Navy along with the primary US federal regulatory agency for marine mammals (US National Oceanic and Atmospheric Administration [NOAA]) and included a multidisciplinary collaboration of academic, private sector, and government scientists. The study was directly motivated by the 2000 beaked whale stranding in a relatively nearby area of the Bahamas (see Filadelfo et al. 2009) and was designed to investigate behavioral responses of beaked whales using 3 different complementary methods suited to measure potential responses on different time, space, and data resolution scales.

The experimental component of the Tyack et al. (2011) study was referred to as the AUTECH Behavioral Response Study (AUTECH-BRS). Empirical measurements of acoustic exposure, behavior, and potential changes in behavior were made using CEEs to compare baseline behavior with that occurring during and following exposure. Three different sound stimuli were projected from an experimental sound source deployed from a research vessel: (1) sonar waveforms used in operational navy systems; (2) pseudorandom noise in the same frequency band; and (3) sounds from mammal-eating killer whales, a potential predator (see Tyack et al. 2011 for more details). The maximum source level used for the sonar and pseudorandom noise stimuli was 211 to 212 dB re 1 μ Pa at 1 m, and that for the killer whale playback was 190 to 203 dB re 1 μ Pa at 1 m. As the first study was specifically designed to expose beaked whales to sonar signals simulating those for which lethal strandings had previously occurred, extensive protective measures for conducting CEEs were applied (Southall et al. in press). The focal species was Blainville's beaked whales *M. densirostris*, but several delphinid cetaceans, including short-finned pilot whales *Globicephala macrorhynchus*, melon-headed whales *Peponocephala electra*, and false killer whales *Pseudorca crassidens*, were included as comparison species and to evaluate whether differential social responses to potential threats might affect the probability of flight reactions and potential associated risk of stranding. Insufficient samples were obtained across these species to fully evaluate differences between species, but some observations of behavioral changes are illustrative and are discussed below. The beaked whale CEE measurements provide extremely detailed, high-resolution measurements of individual behavior for relatively short periods (hours) before, during, and after CEEs using experimental protocols designed to identify potential behavioral response onset.

A second method employed by Tyack et al. (2011) was to deploy satellite tags on individual Blainville's beaked whales to opportunistically monitor movements during a period of days before, during, and after a sonar training exercise conducted on the AUTECH range. These measurements provided relatively coarse measures of surface locations, but extended the total duration monitored to many days. These data included a baseline (before exposure) period of several days, followed by a period of intense sonar operation in the general vicinity of the tagged individual, followed by a period with no sonar transmissions.

The third approach also involved observational monitoring of marine mammal behavior before, during, and after sonar transmissions, but expanded the spatial scale to include the entire AUTECH range. Passive acoustic monitoring was conducted using the array of bottom-mounted hydrophones modified to track cetaceans vocalizing within the AUTECH range. Each of nearly 80 hydrophones in an array that covers hundreds of square kilometers was simultaneously monitored to detect echolocation clicks of multiple beaked whales. The duration of clicking for groups of beaked whales was called a group clicking period (GCP); their location and duration were quantified for comparable periods (20 to 23 h total) before, during, and after active sonar from realistic training operations. This approach lacks an experimental framework and resolution of individual responses, and is limited to measurements of vocalizing animals. However, it provides a broader perspective on potential responses to real operations and movement patterns for the local population of animals on the range.

The combined results of these experimental and observational methods provide a unique perspective into the nature of beaked whale responses to active military sonar. As shown in Fig. 1 (from Tyack et al. 2011), beaked whales in all 3 scenarios clearly responded to sonar exposure, but each method provides different kinds of insights. For the beaked whale exposed to simulated sonar (upper left panel), individual dive behavior was significantly affected. The animal ceased vocalizing early in the dive and demonstrated a long shallow ascent moving horizontally away from the sound source. This effect was more pronounced following subsequent exposure to killer whale signals, after which the animal completely avoided the area. Similarly, the individual tagged with a satellite-linked transmitter (upper right panel) moved tens of kilometers away from the range area during sonar operations, but returned several days following sonar cessation. Finally, passive acoustic monitoring of the entire range to document GCPs before, during, and after active sonar (bottom panel) demonstrated either a complete cessation of vocalizations or, more likely, a temporary avoidance of the sonar use area on the range.

The AUTECH-BRS provided both a unique example of integrating complementary measurements on different time and space scales, and key lessons for subsequent cetacean sonar CEEs. By analyzing behavior and potential responses at the individual level at very high resolution but short duration, as well as coarser resolution but longer duration for

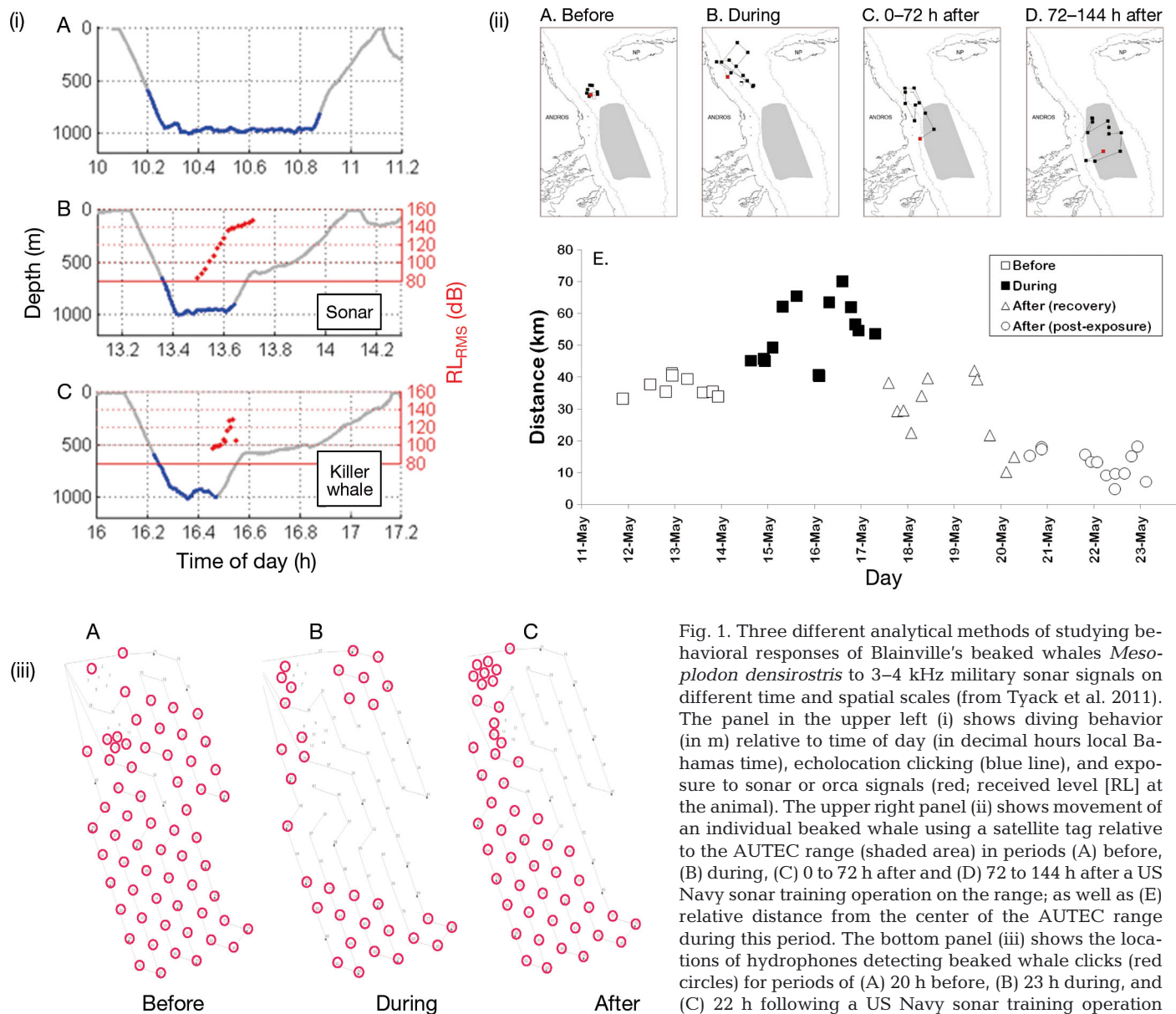


Fig. 1. Three different analytical methods of studying behavioral responses of Blainville's beaked whales *Mesoplodon densirostris* to 3–4 kHz military sonar signals on different time and spatial scales (from Tyack et al. 2011). The panel in the upper left (i) shows diving behavior (in m) relative to time of day (in decimal hours local Bahamas time), echolocation clicking (blue line), and exposure to sonar or orca signals (red; received level [RL] at the animal). The upper right panel (ii) shows movement of an individual beaked whale using a satellite tag relative to the AUTECH range (shaded area) in periods (A) before, (B) during, (C) 0 to 72 h after and (D) 72 to 144 h after a US Navy sonar training operation on the range; as well as (E) relative distance from the center of the AUTECH range during this period. The bottom panel (iii) shows the locations of hydrophones detecting beaked whale clicks (red circles) for periods of (A) 20 h before, (B) 23 h during, and (C) 22 h following a US Navy sonar training operation

fewer data types, it is evident that the responses of Blainville's beaked whales to sonar include avoidance of the sound source as well as changes in vocal and diving behavior. By comparing vocal activity in a large area, analyses on the local population level also suggested a general avoidance of sonar use areas. Each approach has strengths and limitations, but by integrating them, Tyack et al. (2011) were able to demonstrate the general nature of behavioral responses that occurred and the kinds of conditions under which they were likely to begin.

Several additional findings from the AUTECH-BRS provide important insight into other aspects of

cetacean behavior. DeRuiter et al. (2013b) studied potential vocal responses of delphinids and found little change in pilot whale whistle rates during simulated sonar CEEs but an increased probability of melon-headed whales producing sonar-like whistles. These findings demonstrated differences in responses across species and also context dependence of changes in vocal behavior during sound exposure. Additionally, oceanographic ecological variables (prey distribution and density) have been measured with scientific echosounders (Hazen et al. 2011), setting the stage for future integration of relevant environmental and ecological variables in CEEs with foraging animals.

MED-09

The AUTEK-BRS project conducted sonar CEEs with beaked whales within a naval range where animals were routinely exposed to sonar without documented stranding events and where a hydrophone array could detect clicking beaked whales in real time. This setting was selected to minimize the risk to animals while maximizing the probability of experimental success. Once this project successfully and safely conducted CEEs that measured beaked whale responses to sonar, there was a high priority for a similar study with other beaked whales in the Mediterranean Sea area where most sonar-related strandings had been documented. In order to develop research capabilities critical for such a study, the Sirena08 cruise tagged Cuvier's beaked whales *Ziphius cavirostris* in the Mediterranean and monitored whales in real time, using a hydrophone array towed from a very quiet ship (Pavan et al. 2010). This cruise succeeded in using the hydrophone array to detect and track Cuvier's beaked whales, link detections with visual sightings, and to tag an individual whale in the Alboran Sea (Haun et al. 2008). Given this success, the MED-09 study was designed to study Cuvier's beaked whale responses using high-resolution tags and simulated sonar CEEs. Given the high priority assigned to working with Cuvier's beaked whales, the team maintained a focus on this species, even when conditions made this difficult. During the 39 operational days at sea, this cruise was unable to tag Cuvier's beaked whales due to a combination of weather and difficulties in approaching whales closely enough. While a singular focus on beaked whales was maintained in MED-09 because of their high research priority, these challenges led to the evolution of a broader taxonomic approach in later studies, focusing on more species, and a mode of operation that was less dependent on 1 large research vessel (Southall et al. 2012).

The 3S experiments

The 3S project comprised an international collaboration of scientists from Norway, the UK, the USA, and the Netherlands to study the effects of sonar on cetaceans and fish in Norway. Across many NATO countries, new 1–2 kHz sonars were becoming operational (Kvadsheim et al. 2007), and there was concern that these new systems might have greater effects on some species than the higher frequency (6–8 kHz) sonars that had been primarily used up to

that point in European waters. Specific events involving killer whales in both Norway (Vestfjorden; WWF-Norway 2001) and the USA (Haro Strait; Southall & Gentry 2005) raised specific concerns of how killer whales might respond to sonar in ways that could adversely affect whale-watching of this iconic species. Within Norway, specific concern also existed that sonar might affect the presence of herring (WWF-Norway 2001), a key component of Norway's multi-billion dollar fishing industry.

The initial phase of 3S was thus focused on killer whales *Orcinus orca* and their prey, herring *Clupea harengus*, and a motivation to compare sensitivity of killer whales and herring to the new 1–2 kHz sonar relative to existing European naval sonars operating in the 6–8 kHz band. Measurements of hearing in killer whales suggested they have approximately 30 dB lower sensitivity in the 1–2 kHz frequency band than in the 6–7 kHz band (Hall & Johnson 1972, Szymanski et al. 1999). In contrast, herring were predicted to be sensitive to 1–2 kHz sounds but much less sensitive to 6–7 kHz (Enger 1967, Mann et al. 2005). While it is generally accepted that a sound must be audible to induce behavioral responses, the effect of exposure sensation level or perceived loudness on strength or probability of response remains poorly understood.

Initial 3S CEEs with killer whales and herring occurred in November 2006 when large numbers of herring and killer whales gathered in Vestfjorden, Norway. Observing herring with a bottom-mounted echosounder, 6 experiments were conducted over 9 nights comprising 42 exposure sessions. Results indicated that herring did not respond to either 1–2 or 6–7 kHz sonar signals, but did respond to playback of killer whale sounds (Doksæter et al. 2009). In contrast, just 2 controlled exposure experiments sessions (one using 6–7 kHz sonar at a maximum source level of 197 dB re 1 μ Pa at 1 m and the other using 1–2 kHz sonar at a maximum source level of 209 dB re 1 μ Pa at 1 m) were conducted with killer whales (Kvadsheim et al. 2007). A NATO fleet exercise (FLOTEX Silver 2006) took place in the same waters of Vestfjorden during the same period. A retrospective study of killer whale sightings from 2002 to 2008, including research surveys, concluded that declines in killer whale sightings were most strongly correlated with the decline in herring biomass, but that sonar usage in a low herring biomass year (2006) likely had some impact on killer whale distribution (Kuningas et al. 2013).

The 3S study period was changed from winter to summer in 2008 to maximize cetacean sightings and

available daylight. As killer whales are more dispersed in summer, sperm whales *Physeter macrocephalus* and long-finned pilot whales *Globicephala melas* were added as subject species. In the second phase of the 3S study (3S2), only the 1–2 kHz sonar was studied further, and species studied included the humpback whale *Megaptera novaeangliae*, minke whale *Balaenoptera acutorostrata*, and northern bottlenose whale *Hyperoodon ampullatus*.

The core experimental protocol was to escalate received exposure dose (quantified as received level) during 30 to 60 min of controlled exposure to a sonar signals from a towable, operational military sonar source (Socrates, TNO). Following a pre-exposure baseline, an exposure session was initiated with the source vessel positioned 8 to 10 km from the tagged whale. Transmission source levels started at approximately 150 dB re 1 μ Pa at 1 m, and were increased to the full source level of 209 dB re 1 μ Pa at 1 m over 5 to 20 min. Except for the single northern bottlenose whale tested (Miller et al. 2015), additional escalation of the sound level received by the whale was achieved by moving the source vessel toward the whale position, continuously updated by sightings from a separate observation vessel. The source vessel heading was fixed and no further maneuvers were made once the vessel reached a distance of 1 km from the whale, allowing avoiding animals to move away from the oncoming vessel. Using this method, received levels varied from very low levels of <90 dB re 1 μ Pa to as high as 180 dB re 1 μ Pa (Wensveen et al. 2015).

Experiments initially consisted of both 1–2 and 6–7 kHz hyperbolic upswEEP sonar signals. Paired presentations within each CEE sought to control for expected between-whale variation in response, while accounting for presentation order by alternating which signal was presented in the first session. Additional control and sonar exposures included no-sonar control sessions to account for the possible effect of the approaching vessel itself, 1–2 kHz hyperbolic downswEEP exposures, and playback of killer whale sounds at natural source levels (Miller et al. 2011). No-sonar control sessions were identical to exposure sessions, except that no sonar signal was transmitted. Initially, no-sonar control sessions were conducted randomly within the exposure schedule for each experiment (Miller et al. 2011). However, in 3S2, a no-sonar session was conducted as the first session for each experiment, except for the bottlenose whale, to test the effect of vessel approach alone before there was any possibility of the whales becoming sensitized to the signature of a vessel that had recently transmitted sonar.

Two broad classes of analytical procedures were used. Inspired by dose-escalation ('titration') procedures used in phase I clinical trials in human medicine (Simon et al. 1997), diverse analyses identified whether responses occurred during each session, and the received acoustic level associated with the response was used as a 'response threshold' data-point. Thus, each exposure session was evaluated on a case-by-case basis using quantitative multivariate time-series break-point analyses (e.g. Mahalanobis distance; Antunes et al. 2014, Miller et al. 2014, 2015) as well as expert evaluation of the complex datasets recorded for each exposure session (Miller et al. 2012, Sivle et al. 2015). The broad range of received levels from the dose-escalation exposure sessions was important as it enabled identification of behavioral response onset, and lack of any response during a session could be treated using right-censored data methods, which are also standard in medical trials. Relevant covariates for each session (e.g. subject identity or behavioral state, sonar frequency, session order) enabled explicit evaluation of how such contextual variables influenced threshold and quantification of between- and within-animal variability in response thresholds.

Using this response threshold identification approach, it was possible to quantify the probability of a particular response occurring relative to a specific received sound level (Fig. 2) for these experimental contexts. Initial probabilistic dose-response curves were reported for avoidance in killer whales (Miller et al. 2014) and long-finned pilot whales (Antunes et al. 2014). For killer whales, response thresholds were generally lower for 6–7 kHz than for 1–2 kHz sonar, but no order effect was observed. However, the sonar frequency effect was not statistically significant, so a single dose-response curve for both frequencies was derived, with considerable uncertainty resulting from the relatively high levels of between and within animal variability in response. A broader analysis applying expert-identified behavioral response severity was used to derive dose-response functions for low, moderate, and severe responses using Cox proportional hazard models (Harris et al. 2015). In the model selection, species, sonar frequency, and behavioral context prior to exposure were included, so that more specific dose-response functions could justifiably be derived (Fig. 3). Killer whales were found to be more likely to respond to sonar in these exposure contexts at lower received levels than sperm whales or long-finned pilot whales.

In addition to the session-by-session analyses used to derive dose-response functions, several analyses

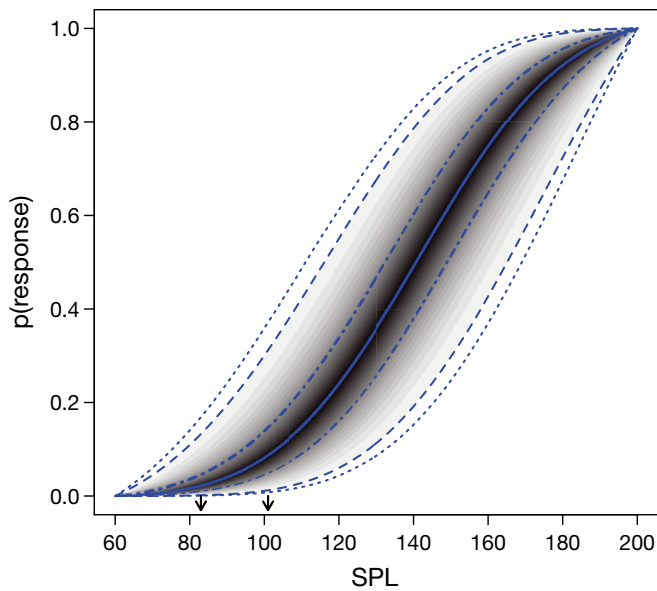


Fig. 2. Dose-response function for the onset of avoidance of sonar by killer whales *Orcinus orca*, as a function of sonar received level (sound pressure level [SPL]; from Miller et al. 2014). The solid central line represents the mean, followed by 50%, 95%, and 99% credible interval lines. The dose-response model assumes the signal is audible, but the limited data on the hearing threshold are marked in the figure with small arrows for 1 (left) and 2 kHz (right) signals, respectively. Reproduced from Miller PJO, Antunes RN, Wensveen PJ, Samarra FI, Alves AC, Tyack PL, Thomas L (2014) Dose-response relationships for the onset of avoidance of sonar by free-ranging killer whales. *J Acoust Soc Am* 135:975–993, with the permission of the Acoustical Society of America

have treated the exposure period as a block in a more traditional before-during-after design. Analysis of expert-identified behavioral responses indicated significantly more severe responses to actual sonar exposures than no-sonar control approach sessions (Miller et al. 2012). A hidden-state analysis of sperm whale behavior identified increases in non-foraging, non-resting 'silent-active' behavior of sperm whales in response to 1–2 kHz sonar and playback of killer whale sounds, but no such change for 6–7 kHz sonar playbacks, no-sonar vessel controls, or a sonar incidentally recorded at lower levels on 3 non-experiment tag records (Fig. 4; from Isojunno et al. 2016).

The 3S2 experiments conducted 1–2 kHz sonar CEEs with 11 humpback whales, as well as a single experiment for each of the difficult-to-tag minke and northern bottlenose whales. Severity scores of expert-identified responses indicated that both the minke and bottlenose whales showed high-severity behavioral responses to 1–2 kHz sonar, while humpbacks demonstrated less severe responses for higher

received levels (Sivle et al. 2015). The individual northern bottlenose whale switched from typical surface behavior between deeper foraging dives to avoidance behavior and conducted the deepest and longest dive recorded for the species, during which no feeding-related echolocation sounds were produced (Miller et al. 2015). Acoustic and visual detections of other whales in the area decreased substantially after the CEE, suggesting that other individuals were also likely affected.

Humpback whales were studied using similar methods, but the protocols were specifically designed to enable measurement of low sonar levels received by the whale to test how behavior was potentially affected by adding a 'ramp-up' procedure prior to full-level transmissions. Results indicated a non-significant effect overall, but that levels were more strongly reduced (–6 dB for maximum sound pressure level, SPL_{max}) for whales that avoided the approaching sonar source (Wensveen 2016). Consistent with theoretical modeling (von Benda-Beckmann et al. 2014), the effectiveness of ramp-up to reduce received sound levels depended critically upon responsiveness. This humpback study demonstrated that experimental tests of mitigation procedures could be a useful application of future CEEs and that ramp-up can reduce exposure levels in some conditions.

SOCAL-BRS

The Southern California Behavioral Response Study (SOCAL-BRS) began in 2010 and evolved from the AUTEK-BRS and MED-09 studies, sharing similar methodological approaches and objectives. Like the other marine mammal and sonar CEEs, SOCAL-BRS involves an interdisciplinary, multi-team collaboration of researchers, although specific measures were implemented to reduce the overall size and cost of the effort while increasing sample sizes of experimental subjects for some species (see Southall et al. 2012 and the following paragraphs). SOCAL-BRS was designed to increase understanding of marine mammal reactions to sound and provide a more robust scientific basis for estimating the impact of specific naval active sonar systems operating in the approximately 3–4 kHz range, systems which are used by a number of navies. The overall species focus was 3-fold: particularly sensitive species (beaked whales), endangered species (baleen whales and sperm whales), and common species (delphinids). Species from all 3 groups were successfully tagged and included in CEEs, providing considerable base-

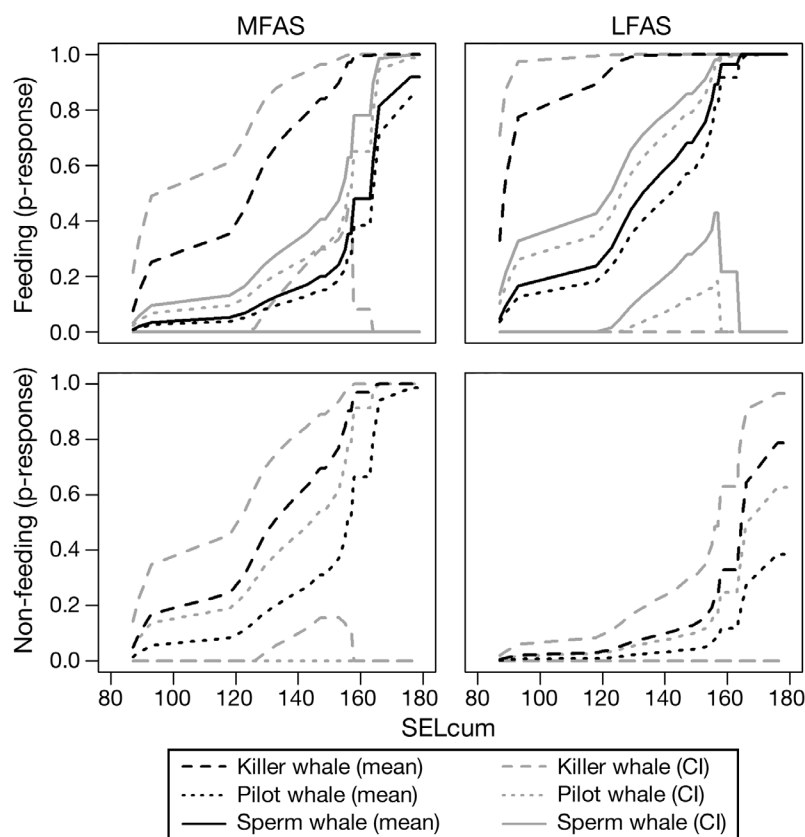


Fig. 3. The probability of a moderate-severity response occurring in (dashed lines) killer whales *Orcinus orca*, (dotted lines) long-finned pilot whales *Globicephala melas* and (solid lines) sperm whales *Physeter macrocephalus* versus received acoustic energy (cumulative sound exposure level [SELcum]) at a signal of (left panels) 6–7 kHz sonar (mid-frequency active sonar, MFAS) and (right panels) 1–2 kHz sonar (low-frequency active sonar, LFAS) and behavioral states of (top panels) feeding and (bottom panels) non-feeding. Mean probabilities are all shown in black, 95% confidence intervals are shown in grey (from Harris et al. 2015)

line information on undisturbed behavior, including a number of species that had never been previously studied with high-resolution tag sensors. SOCAL-BRS initially involved exposures with signals simulating active sonar systems operating in the 2.5–4 kHz frequency band as well as pseudorandom noise, but evolved to include the first ever CEEs conducted using very similar signals from full-scale operational AN/SQS-53C sonars transmitting from actual naval vessels. The overall results of SOCAL-BRS to date are considered briefly here in terms of methodological advances, baseline behavior, and CEE results.

The AUTEK-BRS was successful in demonstrating the nature of Blainville's beaked whale responses to simulated and actual military sonar using complementary methods on very different time and space scales. However, the total sample size was low relative to the total effort and cost of the CEE aspects of

the study, and results for other species were limited. The US Navy and NOAA still had a pressing interest in obtaining data on additional marine mammals, including endangered species that are regularly exposed to sonar as well as Cuvier's beaked whales *Z. cavirostris*, the species most commonly represented in sonar-associated strandings. SOCAL-BRS was planned for areas off southern California with both a wide diversity and abundance of marine mammals and regular sonar operations. As described in greater detail by Southall et al. (2012), the overall approach was to apply similar sonar CEE methods to new species, but to streamline field teams and maximize flexibility by adjusting the species focus based on *in situ* distribution of animals and weather conditions.

A key enabler was the development of a small, hand-deployable vertical line array sound source capable of generating high source levels (up to 212 dB re 1 μ Pa at 1 m) but not requiring a large, oceanographic research ship. SOCAL-BRS built on the Tyack et al. (2011) approach, as well as successes and lessons learned in MED-09 and 3S which helped define goals and objectives. The MED-09 project demonstrated the risks of focusing exclusively on a high priority species

such as *Ziphius* that requires a rare combination of special conditions for tagging to succeed. The 3S project demonstrated the benefits of combining a focus on such high priority species that were difficult to tag with lower priority species that were easier to tag. These insights and capabilities enabled the use of a smaller centralized research platform supporting fast, independent tag boats capable of covering large areas to locate, tag, and track several different species. This evolution was extremely effective in enabling SOCAL-BRS to use smaller, more agile teams able to work adaptively with more species, resulting in sample sizes of dozens for many focal species as opposed to much smaller sample sizes in earlier projects. Sample sizes for beaked whales remained small given the inherent difficulty in locating and tagging them; however, this adaptive approach enabled focus on these species when con-

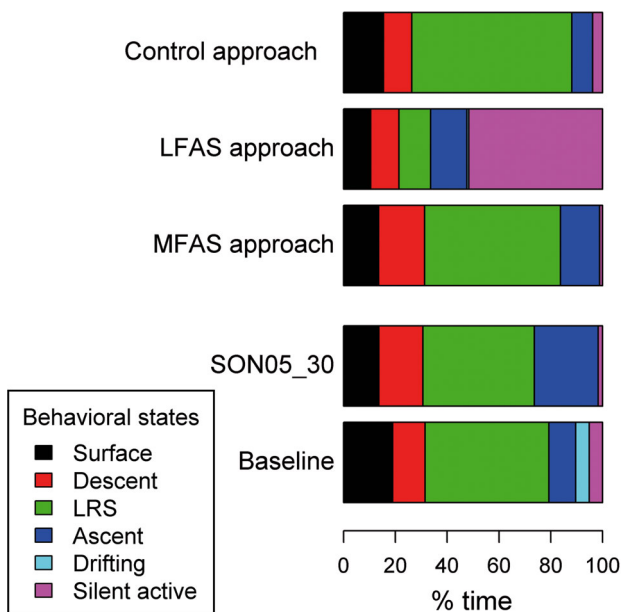


Fig. 4. Time budgets of sperm whales *Physeter macrocephalus* during baseline, incidental sonar exposures (SON05_30), 6–7 kHz mid-frequency active sonar (MFAS) and 1–2 kHz low-frequency active sonar (LFAS) sessions, and no-sonar control sessions (from Isojunno et al. 2016). Note the strong increase in the proportion of time spent in the ‘silent active’ state during 1–2 kHz LFAS sessions and the corresponding decrease in layer restricted search (LRS) which is considered to represent foraging. Time budgets for other exposure types were more similar to time budgets observed in baseline periods

ditions allowed but alternative options when they did not. Additional methodological advances included using active fisheries acoustics to measure prey distribution for feeding baleen whales, resulting in key insights into baseline foraging ecology (e.g. Goldbogen et al. 2013) and the context-dependent nature of behavioral responses (Friedlaender et al. 2016).

While some SOCAL-BRS data were still being collected and analyzed at the time of this review, many results were already published, notably for blue and beaked whales. Goldbogen et al. (2013) provided the first analyses for blue whales, demonstrating that for a subset of individuals, changes in a suite of diving behaviors occurred during some CEEs with both simulated sonar and pseudorandom noise, but that individual behavioral state was critically important in determining response probability. As shown in the example in Fig. 5 for simulated sonar, the composite dive response (a suite of associated dive metrics within the principal component analysis) changed as a function of sonar exposure for animals engaged in deep (>50 m) feeding behavior much more than for those in shallow or non-feeding conditions.

Goldbogen et al. (2013) did not analyze individual responses quantitatively. However, specific examples demonstrated the nature of response associated with the general response patterns described above. Some individuals engaged in shallow feeding behavior demonstrated no clear changes in diving or movement even for quite high (~160 dB re 1 μ Pa) received levels for exposures to 3–4 kHz sonar signals (Fig. 6A). Other individuals exhibited clear responses, including altered dive behavior and increased speed at exposure onset with much lower (~100 to 140 dB re 1 μ Pa) received levels of both pseudorandom noise (Fig. 6B) and sonar (Fig. 6C) signals.

Subsequent analyses have further illustrated the importance of measuring and directly incorporating key contextual variables in response analyses. Friedlaender et al. (2016) demonstrated a 5-fold increase in the ability to quantify variability in blue whale diving behavior when prey distribution and density variables were incorporated into across-individual analyses. These results provided the first integration of direct measures of prey distribution into the analyses of potential responses to sound exposure in feeding marine mammals and illustrated that responses evaluated without such measurements for foraging animals may be misleading. Additional studies are expanding on the findings of Goldbogen et al. (2013) and Friedlaender et al. (2016) which illustrate the context-dependent nature of response probability. Ongoing analyses are considering the importance of

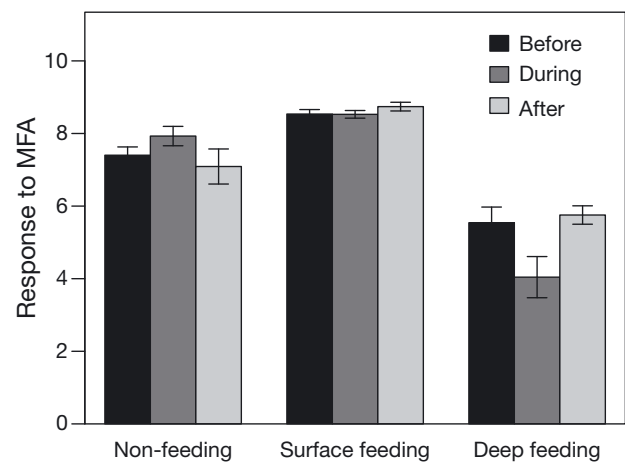


Fig. 5. Scaled dive response values for blue whales *Balaenoptera musculus* exposed to simulated 3–4 kHz mid-frequency active sonar (MFAS) controlled exposure experiments (CEEs) using principal component analyses and generalized mixed models (from Goldbogen et al. 2013). Results are from sequential 30 min periods before, during, and after CEEs for a total of 16 blue whales. Error bars represent 1 SD across individuals

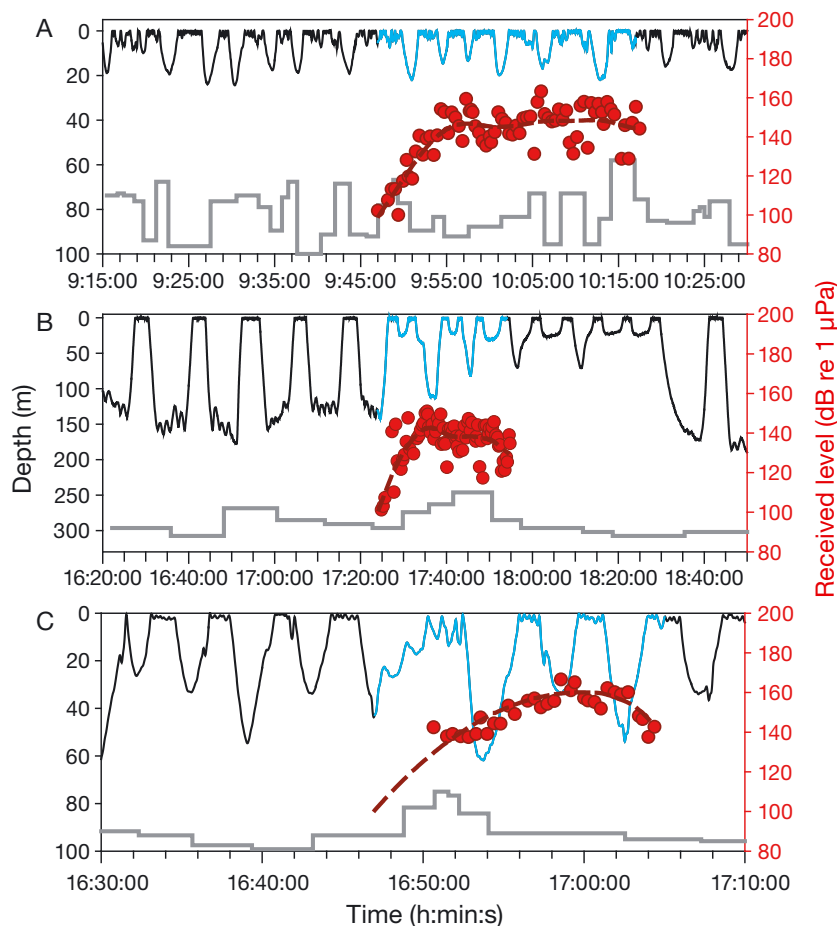


Fig. 6. Individual blue whale *Balaenoptera musculus* behavior during SOCAL-BRS controlled exposure experiments (CEEs) (from Goldbogen et al. 2013). Black lines indicate animal depth (in m, left axis) before and after sound exposure, while blue segments indicate depth during sound exposure. Gray lines show relative animal speed estimated from recorded flow noise. Received levels for each signal transmission during CEEs are shown as red circles (in dB, right axis) and red-dashed lines are estimates of median values for the full exposure period, including some relatively low exposure level periods where they were masked by noise internal to the tag. Results are shown for CEEs with (A) a shallow feeding individual exposed to 3–4 kHz sonar, (B) a deep feeding individual exposed to pseudorandom noise, and (C) a non-feeding individual exposed to 3–4 kHz sonar

behavioral state, environmental context, and spatial orientation and movement of sound sources relative to animals, particularly source–receiver range. These include studies of behavioral state-switching in blue whales using hidden Markov models (DeRuiter et al. in press), as well as individual response analyses using structured expert severity scoring and quantitative change point statistical methods.

Beaked whale CEEs from SOCAL-BRS expanded our understanding of the nature of behavioral response in this group of particularly sensitive species, adding to the findings of Tyack et al. (2011) and Miller et al. (2015) described in the previous sections.

DeRuiter et al. (2013a) documented pronounced and sustained behavioral changes in 2 Cuvier's beaked whales, marking the first ever CEEs for this high priority species. Behavioral responses for 2 subjects tested in 2010 and 2011 (Fig. 7) involved clear changes in diving. These included prolonged ascent from deep dives, termination of echolocation-based foraging, and marked changes in response intensity using multivariate cluster analysis-based Mahalanobis distance analysis (see DeRuiter et al. 2013a for additional details) that exceeded baseline conditions for hours following exposures. Stimpert et al. (2014) observed similar patterns of behavioral response in a Baird's beaked whale to simulated 3–4 kHz sonar CEEs, although response intensity abated more quickly following exposure in that individual compared to the *Ziphius* response measured by DeRuiter et al. (2013a). These combined results are consistent with the findings of Tyack et al. (2011) and Miller et al. (2015) in demonstrating clear, strong, and pronounced behavioral changes, including sustained avoidance with associated energetic swimming and cessation of feeding behavior at quite low received levels (~100 to 135 dB re 1 μ Pa) for exposures with relatively nearby (2 to 5 km) sound sources in simulated sonar CEEs.

In terms of exposure to actual military sonar, limited beaked whale data from SOCAL-BRS represent the only high-resolution, multi-sensor measurements of detailed behavior in these instances presently available. As evident in Fig. 7, the 2011 subject was incidentally exposed to distant actual sonar at the onset of the tag deployment. However, no measurements were made before these exposures and this was an observational, uncontrolled exposure with key aspects (such as source-animal range and orientation) unknown, limiting the conclusions that may be drawn. However, the observed lack of response of this individual is similar to measured behavior of a Cuvier's beaked whale in the SOCAL-BRS CEE conducted with an actual US Navy AN/SQS-53C sonar system (see Southall et al. 2014). In

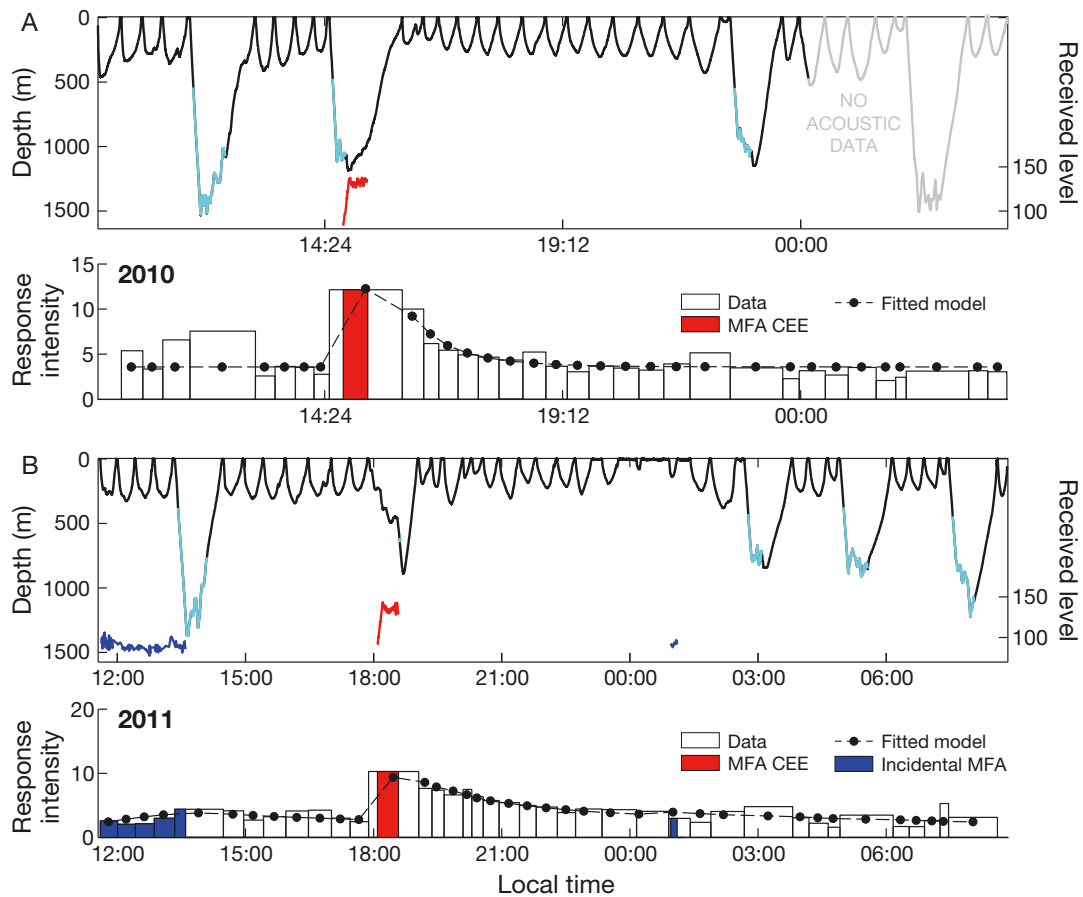


Fig. 7. Individual Cuvier's beaked whale *Ziphius cavirostris* behavior during SOCAL-BRS controlled exposure experiments (CEEs) (from DeRuiter et al. 2013a). Results are shown for individuals tested in (A) 2010 and (B) 2011. For each individual, the top panel shows animal depth (left axis) and received levels (dB re 1 μ Pa) (red line) for each 3–4 kHz mid-frequency active (MFA) sonar transmission (right axes), with (light blue lines) echolocation clicking periods; lower panels show integrated response intensity using a Mahalanobis distance statistic that integrates multiple response variables to a univariate difference metric over time before, during, and after CEEs. The whale in 2011 (panel B) was also incidentally exposed to much lower levels of sonar (shown in dark blue) from other, uncontrolled sources during periods before and after received signals from the CEE (red)

this instance, received levels were controlled *in situ* by positioning the ship approximately 70 km away from the tagged individual, based on sound propagation modeling, such that received levels were slightly lower (100 to 120 dB re 1 μ Pa) than those observed by DeRuiter et al. (2013a) to induce strong behavioral responses. Although the tag detached from this individual at the end of the CEE, which precluded post-exposure behavioral analysis, behavior measured before and during this CEE did not reveal similarly strong detectable behavioral changes. Further analyses of these results are ongoing and additional CEE measurements with operational navy vessels are being obtained in the later phases of SOCAL-BRS. The limited observational and CEE results to date with operational navy vessels suggest that additional contextual variables (e.g. exposure range) may also mediate the probability of response, possibly along

with received level. This is one of a number of key questions regarding factors influencing response probability that require considerable additional focused research (discussed below).

CONCLUSIONS, IMPLICATIONS, AND FUTURE RESEARCH

General conclusions from recent sonar-related CEEs

The above studies represent significant progress in experimental methods to measure behavioral responses (or lack thereof) of various cetaceans to different sonar signals on very fine scales in a causative rather than correlative manner. This is possible because of the use of a controlled, experimental approach with high-resolution measurement tags. No-

sonar control sequences have been particularly useful to evaluate potential changes relative to natural variability in behavioral patterns and to test for potential experimental effects such as source movement. Escalating received levels through sound source movement and/or source level escalation has been important for achieving a broad range of received levels needed to derive robust exposure-response curves. However, each mode of dose escalation may differentially affect whale response. The stationary source form of dose escalation does not duplicate all features of exposure to sonar from a moving naval vessel, but may simulate some aspects of an approach given the rapidly increasing levels used. The escalation methodology using an approaching source is more realistic, and likely increases the relative probability of response compared to situations where the source is stationary or moving away. The resulting sample sizes of active sonar CEEs with cetaceans have been relatively small for some, but not all, species. However, many useful and very timely results with practical management implications have been generated, and several broad conclusions are evident.

Clearly identifiable behavioral responses to sonar exposure have been measured in some, but not all, individuals of all species tested. While many, but not all, responses have been relatively mild and/or brief, it is important to note that these were short-term experiments intentionally designed not to harm subjects, but rather to identify the onset and nature of different behavioral responses. Common responses have included avoidance of the area of sonar exposure, either directed away from a stationary or often perpendicular to the track of an approaching vessel (Miller et al. 2012, Goldbogen et al. 2013). Cessation or modification of vocal behavior during sonar exposure has been shown in multiple studies (e.g. DeRuiter et al. 2013b, Miller et al. 2012, 2014, Alves et al. 2014). Additionally, cessation of foraging has been documented in multiple studies, based on acoustic (e.g. Miller et al. 2012, DeRuiter et al. 2013a) or kinematic (e.g. lunges) indicators of feeding (Goldbogen et al. 2013, Friedlaender et al. 2016). The responses observed during and after relatively short (tens of minutes) exposure periods are not thought to have produced significant or permanent adverse effects, an important aspect of the study design. For example, subject md07_245 from the Tyack et al. (2011) study, which showed a strong and prolonged response, has been sighted numerous times in subsequent years. However, as noted, these experiments were deliberately designed to demonstrate the onset

of response and not to produce adverse effects. Nevertheless, these kinds of sub-lethal behavioral changes may still have significant energetic and physiological consequences given sustained or repeated exposure to these kinds of stimuli, something which is certainly a common occurrence in some high-use sonar areas. The severity of behavioral responses has been considered within some studies (e.g. Miller et al. 2012, Sivle et al. 2015) using an adaptation of the Southall et al. (2007) response severity scaling.

Both the AUTEK and 3S programs conducted playback of killer whales as a positive control to sonar exposures. These killer whale playbacks were useful for demonstrating the ability of the observation systems to detect behavioral responses (i.e. Doksæter et al. 2009, Miller et al. 2012) and to provide a useful biologically relevant 'yardstick' against which to compare responses to sonar (Curé et al. 2012, 2013, 2015). Animals are predicted to have evolved behavioral responses to predators (cessation of foraging, avoidance) that may be costly in the short term provided they are effective at reducing the risk of being killed (Frid & Dill 2002). The killer whale playbacks made it possible to test the hypothesis that the response(s) documented to sonar exposure resemble(s) a predator threat response (Tyack et al. 2011, Miller et al. 2012). The overall nature of behavioral responses of sperm whales and humpback whales to sonar were quite similar to responses to playbacks of mammal-eating killer whale sounds recorded in places other than where they were played back (Miller et al. 2012, Curé et al. 2013, 2015, 2016, Sivle et al. 2015, Isojunno et al. 2016). Long-finned pilot whales responded strongly and consistently to playback of herring-feeding killer whale sounds, but responses indicated a mobbing response (Curé et al. 2012), which was not observed during sonar exposures. Responses similar to mobbing behavior have been documented for subsequent dedicated playback experiments using feeding-associated sounds produced by unfamiliar killer whales (P. J. O. Miller unpubl. data), indicating that this may be a common strategy of long-finned pilot whales responding to killer whale presence (De Stephanis et al. 2015). These results suggest that cetaceans respond to sonar in a manner that may be shaped by anti-predator adaptations and association, but that sonar sounds are not categorized as so intense a threat or competitor as actual killer whale sounds.

While the probability of response to sonar varied considerably within and across the marine mammals tested, one particular generalization across the various taxa tested appears clear. Namely, individuals of

all 4 beaked whale species tested—*Mesoplodon densirostris* (Tyack et al. 2011), *Ziphius cavirostris* (De Ruiter et al. 2013a), *Berardius bairdii* (Stimpert et al. 2014), and *Hyperoodon ampullatus* (Miller et al. 2015)—have shown particular sensitivity in terms of the probability of response (100% of individuals tested have responded strongly to simulated sonar), very low received levels at the onset of behavioral response, and the relative strength and sustained duration of individual responses, compared to some other cetaceans such as blue whales (Goldbogen et al. 2013) and pilot whales (Antunes et al. 2014). However, the responses of some species (notably Baird's beaked whales) were not markedly stronger than those documented in some other non-beaked whale cetaceans such as killer whales (Miller et al. 2014). Individuals of all 4 beaked whale species did respond with cessation of echolocation-based foraging and extended dive durations consistent with avoidance reactions. These collective results appear categorically different from those of most other cetacean species, where more varied responses and, in some cases, a lack of any apparent behavioral response even for relatively intense exposures have been observed. The strong and consistent nature of behavioral responses measured in beaked whales to date is consistent with earlier observations based on stranding events that they may be more sensitive to noise exposure (e.g. Cox et al. 2006). The strong response for a bottlenose whale measured by Miller et al. (2015) was particularly striking in that it resulted from a 1–2 kHz sonar signal, which is even further from the species-typical sound production than the 2–7 kHz sonars in the other studies, presented over 5 km from the source in an area where operational military sonars seldom occur. While the CEE results for beaked whales now span 4 different species in different parts of the world, these studies have all involved small sample sizes (1 to 2 individuals per species), and additional experimental work is clearly needed.

Some individuals of non-beaked whale species have also demonstrated relatively strong reactions to sonar (e.g. killer whale subject oo09_144a in Miller et al. 2012, 2014; the blue whale example in Goldbogen et al. 2013 in Fig. 6C). Interestingly, the 3S results (Miller et al. 2012, 2014) have demonstrated that some killer whales are also particularly sensitive to sonar compared to other species. This represents an interesting mismatch with the predation-risk theory posed and tested for other species in that they may treat sonar exposure as a general threat to which responses have been shaped by natural responses to the threats posed by predation.

Among the non-beaked whale CEE results, there is considerable variability within and between species, even in the same studies, in terms of the probability and nature of behavioral response. The combined results indicate diversity in the probability of responses within species, some of which may vary because of variation in the received levels and because of differences in context (as suggested by Ellison et al. 2012). Unfortunately, it is not possible to control or measure the large variety of potentially relevant contextual factors. However, contextual variables have been increasingly identified and successfully described within sonar CEEs in terms of their relative contributions to response probability. For instance, within the 3S studies, a Bayesian dose-response analysis quantified a high degree of within and between animal variability in the onset of behavioral response (Antunes et al. 2014, Miller et al. 2014). Behavioral state at the start of exposure also appears a relevant consideration. Blue whales feeding on deep, dispersed prey were more likely to change diving behavior and avoid simulated sonar sources than whales feeding at shallow depths on highly concentrated prey (Goldbogen et al. 2013). These findings demonstrate the importance of identifying and accounting for both intrinsic and extrinsic behavioral and exposure variables to fully predict responses to sonar.

Key questions remain regarding potential sources of context-dependent variability that may be important in predicting the probability of response, notably the relative importance of behavioral state and spatial relationships between sound sources and animals. However, significant progress has been made in this regard, guiding future studies and future observational monitoring. There is increasing recognition and quantification of these issues within CEEs, including the recognition that not every complex context interaction must be directly studied. Provided that CEEs involve subjects sharing the same context with the areas, seasons, and populations with which sonar operations typically occur, results can be used to generate probabilistic risk functions that can estimate the responses of these populations without the requirement to quantify all of the context-specific causes for the variability. Studies on focal animals where context can be quantified provide a common basis for comparison and analysis not possible with uncontrolled observational methods. Experimental studies can identify behavioral changes caused by sonar. Once these responses have been defined, they can be used to inform more targeted monitoring programs designed to identify when and where such responses occur during actual sonar exercises.

Applications for assessing potential effects of sonar

The combined BRS results obtained over the past decade have significantly improved our understanding of the types and probabilities of response in marine mammals to simulated and actual naval sonar sources; recent efforts in SOCAL-BRS have begun to explore the responses of cetaceans to controlled exposures of actual naval sonar systems in realistic scenarios, including use of the most intense sources at long ranges. In many conditions, animals do not respond to sonar in any detectable way. However, various behavioral responses have been observed in some conditions; these have been quantified in fine-scale resolution and can be experimentally demonstrated to result from sonar exposure. There is considerable within- and between-species variability in terms of the probability of response, likely related to both intrinsic (e.g. individual differences in threshold for response, behavioral state during exposure) and extrinsic (e.g. spatial orientation of source and receiver) exposure contextual factors. Within the experimental contexts tested, received sound level has been associated with the probability of responses in an experimentally controlled dose-escalation context to generate preliminary, probabilistic dose-response functions (e.g. Miller et al. 2014). In some conditions, individual-specific factors such as behavioral state may interact with exposure variables such as received level to influence the probability of response (e.g. Goldbogen et al. 2013). However, received level remains an important exposure variable for consideration and one that has been central within the regulatory context. Sonar BRS results that include exposure levels and other relevant identifiable contextual variables are clearly needed by navies and regulatory agencies. Current predictions of response probability based on the available, rapidly increasing data provide vastly superior estimates relative to earlier response functions based upon anecdotal, uncontrolled observations, recreations of strandings, and/or captive studies with little relevance to free-ranging animals (Department of the Navy 2008).

Many of the regulatory compliance questions and data requirements facing navies that operate active sonars require sufficient detail on the probability of behavioral response to sonar exposures. This is most directly accomplished through the use of experiments to establish causal links between exposure and responses, with high-resolution measurements with the potential to detect subtle aspects of individual response behavior. Once such responses have been identified, targeted observational methods building

on experimental results can be developed to monitor and measure the effects of realistic sonar exercises over the associated predicted ranges. Then monitoring can be developed to measure effects of actual sonar exercises over broader time and space scales for local populations. An excellent example of this comes from Moretti et al. (2014), who used passive acoustic monitoring of echolocation-based foraging behavior in beaked whales. Using these methods coupled with sound propagation modeling, Moretti et al. (2014) derived a received level exposure-response function for the initiation of foraging dives as a function of exposure to actual sonar exercises. The results have associated caveats as the units of analysis involve an unknown mixture of responses from different individuals, repeated measurements from the same individuals, and may not represent the most sensitive individuals because they are sampled from a range where sonar is commonly used and sensitive individuals may be less likely to be in the range than in other areas of their habitat. However, the clear benefits of this kind of approach is that it builds on both experimental and observational studies within a species (Tyack et al. 2011) to investigate the probability of a particular behavioral response over more representative time and space scales in the context of real naval operations. This combination of high-resolution experimental and coarser-resolution observational methods is a timely and important element of this field of research; we discuss this in greater detail in 'Directions for future research' below. Here we propose a focused and adaptive application of experimental sonar CEE approaches as the fundamental basis for approaching and understanding cetacean behavioral response to inform and direct more targeted and effective observational monitoring of local populations of animals during realistic operations.

The combined results are useful for improving the assessment of exposure-response functions and the type and severity of potential adverse effects of sonar operations. Recent results are already being applied within the ongoing evolution of regulatory compliance processes. However, it remains challenging to assess the relative biological significance and broader ecological and population consequences of short-term experimental measurements of behavioral response onset relative to the much larger and more complex navy training and testing operations. The evaluation of measured response severity is one effort to try to assess the significance of behavioral responses documented in BRS studies. Longer duration of exposure may lead to more severe responses (i.e. longer disruption of behavior or sensitization to repeated exposures)

or might, alternatively, result in tolerance or habituation following some initial response. Relative motion and spatial orientation of sound exposure may be particularly important (e.g. Ellison et al. 2012): vessels that consistently approach a whale, as in the 3S protocol, may lead to different patterns of response than vessels moving obliquely or away from an animal (Miller et al. 2012). Clearly the sound sources and exposure context of most of these experiments differ in a number of fundamental ways (e.g. source level, directivity, range, spatial orientation) from the kinds of real-world sonar systems. Indeed, these experiments were carefully designed to minimize risk of stranding while still detecting responses that may be of regulatory and biological significance. Because these contextual factors differ between CEEs and actual sonar exercises in ways that may affect the probability and type of behavioral responses (Ellison et al. 2012), direct extrapolation and application of experimental results to full-scale operational systems needs to be considered a fundamental part of the experimental design in future studies, as was done in Tyack et al. (2011) and Moretti et al. (2014) for sonar exercises at AUTECH.

Teasing apart the relationship between physical range from the sonar source to the whale (proximity) and received level of the signal at the whale is a related issue that could be addressed by using sonar systems with a variety of source levels. The initial BRS experiments logically began, particularly with beaked whales where lethal consequences of sonar exposure were known, with lower power, stationary systems to minimize the potential risk of stranding (e.g. Southall et al. in press). Later BRS studies have used moderate-power sonar systems that have lower source levels than the most intense naval sonars. For example, the 3S experiments with killer whales used a source level ranging from 197 to 214 dB re 1 μ Pa at 1 m. These source levels are consistent with some European naval sonar systems, but are much lower than the most intense naval sonars used by a number of navies around the world. The average acoustic avoidance response threshold of 142 dB re 1 μ Pa was recorded at an average source–whale distance of 3.8 to 4.6 km (Miller et al. 2014). However, a 142 dB re 1 μ Pa received level from a 225 dB re 1 μ Pa at 1 m source level operational sonar source would occur over much larger distances, up to 60 \times the distances in these experiments. It remains unknown to what degree the distance from the whale to the source might influence the acoustic thresholds for behavioral response, but DeRuiter et al. (2013a) found that distance to the source was likely an important factor for Cuvier's beaked whales located near an opera-

tional navy range. Because of the importance of spatial context of sonar exposure in estimating how many animals may be behaviorally affected by sonar, future research to quantify effects of source distance and movement is strongly recommended. Given that we have now demonstrated that such studies can be safely and effectively conducted, including the ability to measure and quantify the significance of contextual covariates, more direct investigation of the relative significance of spatial context is possible. This is a key area of research within SOCAL-BRS, as described above, along with the deliberate integration of experimental approaches using the most powerful operational active naval sonar (e.g. AN/SQS-53C) systems in the context of actual navy operations.

The application of rapidly increasing data from cetacean sonar CEEs to evaluate potential impacts and mitigation efficacy for realistic sonar operations is clearly challenging. Many factors beyond the type and relevance of available data on response probability affect such assessments, such as differential regulatory requirements, paradigms, and distinctions in different jurisdictions and ecosystems. Regulatory assessments must account for operational factors such as the type and nature of active transmissions in terms of overall magnitude, distribution and use patterns, and differential source and signal types. They must also consider the population status of exposed animals and the quality of potential alternate habitat surrounding areas where temporary displacement may regularly occur.

The past decade of research has unquestionably increased the scientific basis for improving regulatory assessments of potential sonar risks to cetaceans. So, too, has it evidenced a number of key questions and considerations for ongoing and future research. Studies have identified numerous sources of variability in the probability of response, including species and individual differences and context dependencies for response. Research indicates the need to move beyond step-function thresholds based on received level to empirically based probabilistic exposure-response functions. If the variability of responses makes probabilistic exposure-response functions too broad for effective regulation, further evolution might consider different exposure-response functions for broad contextual categories (e.g. particularly sensitive species, behavioral state differences, or source–receiver range). These should remain clearly distinct from operational mitigation requirements, as it is clearly unrealistic to identify practical methods by which navies could effectively incorporate such contextual variables in their real-world use of sonar.

Directions for future research

One of the major positive outcomes of the substantial body of recent research discussed here is the inter-disciplinary and international collaboration of biologists, engineers, statisticians, and acousticians that has resulted. Active and future research will require sustained effort and significant financial support from various navies for teams that can tackle these complex problems using novel experimental and analytical methods. Many critical, current, and pressing applied questions require experimentally testing the correct taxa, locations, time periods, and exposure contexts in order to define how these animals respond to relevant exposures. The longer-term objective is to develop science-based methods from these relevant behavioral metrics to monitor critical responses during actual operations in order to understand and predict the potential severity of effects from active military sonar. A number of targeted near-term research needs have been discussed throughout this review, and are summarized here:

(1) Increase the realism of sonar exposures used within experimental studies. Considerable effort is underway to conduct CEEs using the most intense active sonar systems with increasingly realistic exposure conditions to successfully match received levels tested with simulated sources using full-power systems at much greater physical range. These comparisons may support the evaluation of the relative importance of proximity relative to received exposure level for operational sonars.

(2) Evaluate the relative familiarity of experimental subjects with sonar. Many, though not all, active military sonar transmissions around the world occur in relatively concentrated regions around training and testing areas where most monitoring and research on potential effect has logically occurred. However, animals in these areas may be less sensitive to sonar because of habituation, tolerance or because sensitive animals may leave. This means that predictions derived from areas where sonars frequently transmit may underestimate potential responses for animals in regions where sonar use is less common but still occurs. Additional studies are needed in areas of less regular sonar use.

(3) Select appropriate locations, time periods, and focal species. Species tested have either been thought to be particularly sensitive, endangered or threatened, or very common in areas of sonar operations. Harris et al. (2016) present statistical techniques for deciding when and how to pool different species, contexts, and signal types based upon empirical observations of re-

sponses to sonar, but this can be challenging given the large number of different species and observed species-, individual-, and context-mediated differences in response behavior. Subsequent research should evaluate patterns of social structure, susceptibility to predation risk, and key life history parameters to develop generalized models of functional sensitivity categories based on species groups (beaked whales) or behavior state (e.g. feeding vs. traveling). These analyses can help to direct decisions about how many subjects of which species in which states are most effective in reducing uncertainty in response functions for specific sites and sources.

(4) Expand spatial and temporal scales of experimental studies to overlap with monitoring methods. Experimental responses need to be considered within the context of broader monitoring observations (as in Moretti et al. 2014) and relative to increasingly detailed, longer-term movement and diving behavior from satellite tags (e.g. Schorr et al. 2014), particularly where sonar exposure is sufficiently documented. Specifically, there is a clear need to apply CEE methodologies in tandem with observational monitoring in conditions where longer-duration or sequential exposures over multiple days can be measured and where relatively high-resolution measurements of individual behavior are possible using measurement devices that can be attached for days to weeks. Evaluations are also needed of both individual fitness parameters (using metrics of health or body condition that can be measured with archival tags), as well as of foraging habitat metrics, in order to evaluate potential population-level effects.

We argue here for an integrated research strategy using complementary experimental and observational methods to further advance our understanding of the potential effects and biological significance of military sonar on cetaceans. Focused work is needed to (1) improve our understanding of the relationship between response probability and intensity to acoustic exposure, including key contextual variables (e.g. behavioral state, source–receiver proximity); (2) increase the temporal and spatial scales over which responses can be evaluated with both targeted experimental studies and increasingly informed observational monitoring; and (3) identify the linkages of short- and medium-term behavioral responses to changes in individual vital rates. Given the substantial progress made in the past decade thanks to the efforts and collaborations described here, targeted studies in these areas over the next decade will fundamentally change the way we understand, manage, and mitigate the effects of active sonar on cetaceans.

Acknowledgements. The authors would like to acknowledge the many dedicated and talented people that have done so much hard work in the AUTEK, MED-09, 3S, and SOCAL-BRS projects. While they are too numerous to name individually here, many are represented in the authorship and acknowledgments of the referenced publications from these projects. All studies described were conducted under requisite local, federal, and institutional permits and animal care and use protocols. Funding was provided by a variety of military and governmental funding sources from several nations acknowledged within referenced publications, notably the US Office of Naval Research, US Navy Living Marine Resources Program, and the navies of the USA, Norway, and the Netherlands. P.L.T. acknowledges the support of the MASTS pooling initiative (The Marine Alliance for Science and Technology for Scotland) in the completion of this study. MASTS is funded by the Scottish Funding Council (grant reference HR09011) and contributing institutions. We appreciate the constructive suggestions of 2 anonymous reviewers and guest editor Dr. Wendy Dow Piniak.

LITERATURE CITED

- Altmann J (1974) Observational study of behavior: sampling methods. *Behaviour* 49:227–267
- Alves A, Antunes R, Bird A, Tyack PL, Miller PJ, Lam FPA, Kvadsheim PH (2014) Vocal matching of naval sonar signals by long-finned pilot whales (*Globicephala melas*). *Mar Mamm Sci* 3:1248–1257
- Antunes R, Kvadsheim PH, Lam FPA, Tyack PL, Thomas L, Wensveen PJ, Miller PJO (2014) High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*). *Mar Pollut Bull* 83:165–180
- Arranz P, DeRuiter SL, Stimpert AK, Neves S and others (2016) Discrimination of fast click series produced by tagged Risso's dolphins (*Grampus griseus*) for echolocation or communication. *J Exp Biol* 219:2898–2907
- Balcomb KCI, Claridge DE (2001) A mass stranding of cetaceans caused by naval sonar in the Bahamas. *Baham J Sci* 2:2–12
- Bradbury JW, Vehrencamp SL (1998) Principles of animal communication. Sinauer Press, Sunderland, MA
- Brownell RL, Yamada T, Mead JG, Van Helden AL (2004) Mass strandings of Cuvier's beaked whales in Japan: US Naval acoustic link? Paper SC/56/E37 presented to the IWC Scientific Committee, July 2004, Sorrento. Available from the Office of the Journal of Cetacean Research and Management
- Cato DH, Noad MJ, Dunlop RA, McCauley RD and others (2013) A study of the behavioural response of whales to the noise of seismic air guns: design, methods and progress. *Acoust Aust* 41:88–97
- Cheney DL, Seyfarth RM, Silk, JB (1995). The role of grunts in reconciling opponents and facilitating interactions among adult female baboons. *Anim Behav* 50: 249–257
- Costa DP, Robinson PW, Arnould JPY, Harrison AL and others (2010) Accuracy of ARGOS locations of pinnipeds at-sea estimated using fastloc GPS. *PLOS ONE* 5:e8677
- Cox TM, Ragen TJ, Read AJ, Vos E and others (2006) Understanding the impacts of anthropogenic sound on beaked whales. *J Cetacean Res Manag* 7:177–187
- Croll DA, Clark CW, Calambokidis J, Ellison WT, Tershy BR (2001) Effect of anthropogenic low-frequency noise on the foraging ecology of *Balaenoptera* whales. *Anim Conserv* 4:13–27
- Curé C, Antunes R, Samarra F, Alves AC, Visser F, Kvadsheim PH, Miller PJ (2012) Pilot whales attracted to killer whale sounds: acoustically-mediated interspecific interactions in cetaceans. *PLOS ONE* 7:e52201
- Curé C, Antunes R, Alves AC, Visser F, Kvadsheim PH, Miller PJO (2013) Responses of male sperm whales (*Physeter macrocephalus*) to killer whale sounds: implications for anti-predator strategies. *Sci Rep* 3:1579
- Curé C, Sivle LD, Visser F, Wensveen PJ and others (2015) Predator sound playbacks reveal strong avoidance responses in a fight strategist baleen whale. *Mar Ecol Prog Ser* 526:267–282
- Curé C, Isojunno S, Visser F, Wensveen PJ and others (2016) Biological significance of sperm whale responses to sonar: comparison with anti-predator responses. *Endang Species Res* 31:89–102
- D'Amico A, Gisiner RC, Ketten DR, Hammock JA, Johnson C, Tyack PL, Mead J (2009) Beaked whale strandings and naval exercises. *Aquat Mamm* 35:452
- De Stephanis R, Giménez J, Esteban R, Gauffier P, Garcia-Tiscar S, Sinding MHS, Verborgh P (2015) Mobbing-like behavior by pilot whales towards killer whales: a response to resource competition or perceived predation risk? *Acta Ethol* 18:69–78
- Department of the Navy (2008) Atlantic fleet active sonar training: final environmental impact statement/overseas environmental impact statement (FEIS/OEIS). Department of the Navy, Washington, DC
- DeRuiter SL, Southall BL, Calambokidis J, Zimmer WMX and others (2013a) First direct measurements of behavioral responses by Cuvier's beaked whales to mid-frequency active sonar. *Biol Lett* 9:20130223
- DeRuiter SL, Boyd IL, Claridge DE, Clark CW, Gagnon C, Southall BL, Tyack PL (2013b) Delphinid whistle production and call matching during playback of simulated military sonar. *Mar Mamm Sci* 29:E46–E59
- DeRuiter SL, Langrock R, Skirbutas T, Goldbogen JA, Calambokidis J, Friedlaender AS, Southall BL (in press) A multivariate mixed hidden Markov model to analyze blue whale diving behaviour during controlled sound exposures. *J R Stat Soc*. <https://arxiv.org/abs/1602.06570>
- Doksæter L, Godø OR, Handegard NO, Kvadsheim PH, Lam FPA, Donovan C, Miller PJ (2009) Behavioral responses of herring (*Clupea harengus*) to 1–2 and 6–7 kHz sonar signals and killer whale feeding sounds. *J Acoust Soc Am* 125:554–564
- Dunlop RA, Noad MJ, Cato DH, Kniest E, Miller PJO, Smith JN, Stokes MD (2013) Multivariate analysis of behavioral response experiments in humpback whales (*Megaptera novaeangliae*). *J Exp Biol* 216:759–770
- Dunlop RA, Noad MJ, McCauley RD, Kniest E, Paton D, Cato DH (2015) The behavioural response of humpback whales (*Megaptera novaeangliae*) to a 20 cubic inch air gun. *Aquat Mamm* 41:412–433
- Ellison WT, Southall BL, Clark CW, Frankel AS (2012) A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Conserv Biol* 26:21–28
- Enger PS (1967) Hearing in herring. *Comp Biochem Physiol* 22:527–538
- Evans DL, England GR (2001) Joint interim report: Bahamas marine mammal stranding event of 15–16 March 2000. US Department of Commerce, NOAA, National Marine

- Fisheries Service. www.nmfs.noaa.gov/pr/pdfs/health/stranding_bahamas2000.pdf (accessed 29 Oct 2016)
- Falls JB (1992) Playback: a historical perspective. In: McGregor PK (ed) *Playback and studies of animal communication*. Plenum Press, New York, NY, p 11–23
- Falls JB, Krebs JR, McGregor PK (1982) Song matching in the great tit (*Parus major*): the effect of similarity and familiarity. *Anim Behav* 30:997–1009
- Fernández A, Edwards JF, Rodríguez F, Espinosa de los Monteros A and others (2005) Gas and fat embolic syndrome involving a mass stranding of beaked whales (family *Ziphiidae*) exposed to anthropogenic sonar signals. *Vet Pathol* 42:446–457
- Filadelfo R, Mintz J, Michlovich E, D'Amico A, Tyack PL (2009) Correlating military sonar use with beaked whale mass strandings: What do the historical data show? *Aquat Mamm* 35:435–444
- Finneran JJ (2015) Noise-induced hearing loss in marine mammals: A review of temporary threshold shift studies from 1996 to 2015. *J Acoust Soc Am* 138:1702–1726
- Frankel AS, Clark CW (2000) Behavioral responses of humpback whales (*Megaptera novaeangliae*) to full-scale ATOC signals. *J Acoust Soc Am* 108:1930–1937
- Frantzis A (1998) Does acoustic testing strand whales? *Nature* 392:29
- Frid A, Dill LM (2002) Human-caused disturbance stimuli as a form of predation risk. *Conserv Ecol* 6:11
- Friedlaender AS, Goldbogen JA, Hazen EL, Calambokidis JA, Southall BL (2015) Feeding performance of sympatric blue and fin whales exploiting a common prey resource. *Mar Mamm Sci* 31:345–354
- Friedlaender AS, Hazen EL, Goldbogen JA, Stimpert AK, Calambokidis J, Southall BL (2016). Prey-mediated behavioral responses of feeding blue whales in controlled sound exposure experiments. *Ecol Appl* 26: 1075–1085
- Fristrup KM, Hatch LT, Clark CW (2003) Variation in humpback whale (*Megaptera novaeangliae*) song length in relation to low-frequency sound broadcasts. *J Acoust Soc Am* 113:3411–3424
- Goldbogen JA, Southall BL, DeRuiter SL, Calambokidis J and others (2013) Blue whales respond to simulated mid-frequency military sonar. *Proc R Soc Lond B Biol Sci* 280: 20130657
- Goldbogen JA, Stimpert AK, DeRuiter SL, Calambokidis J and others (2014) Using accelerometers to determine the calling behavior of tagged baleen whales. *J Exp Biol* 217: 2449–2455
- Goldbogen JA, Hazen EL, Friedlaender AS, Calambokidis J, DeRuiter SL, Stimpert AK, Southall BL (2015) Prey density and distribution drive the three-dimensional foraging strategies of the largest filter feeder. *Funct Ecol* 29:951–961
- Hall JD, Johnson CS (1972) Auditory thresholds of a killer whale (*Orcinus orca Linnaeus*). *J Acoust Soc Am* 51: 515–517
- Harris CM, Sadykova D, DeRuiter SL, Tyack PL and others (2015) Dose response severity functions for acoustic disturbance in cetaceans using recurrent event survival analysis. *Ecosphere* 6:236
- Harris CM, Thomas L, Sadykova D, De Ruiter SL and others (2016) The challenges of analyzing behavioral response study data: an overview of the MOCHA (Multi-study Ocean acoustics Human effects Analysis) project. In: Popper AN, Hawkins A (eds) *The effects of noise on aquatic life II. Advances in experimental medicine and biology* 875. Springer Science and Business Media, New York, NY p 399–407
- Haun J, Ryan K, Portunato N (2008) Marine mammal risk mitigation using passive acoustic technologies. In: *New trends for environmental monitoring using passive systems*. IEEE Workshop PASSIVE 2008, October 14–17, Hyères, p 91–95
- Hazen EL, Nowacek DP, St Laurent L, Halpin PN, Moretti DJ (2011) The relationship among oceanography, prey fields, and beaked whale foraging habitat in the Tongue of the Ocean. *PLOS ONE* 6:e19269
- Hazen EL, Friedlaender AS, Goldbogen JA (2015) Blue whales (*Balaenoptera musculus*) optimize foraging efficiency by balancing oxygen use and energy gain as a function of prey density. *Sci Adv* 1:e1500469
- Holt MM, Southall BL, Insley SJ, Schusterman RJ (2010) Call directionality and its behavioural significance in male northern elephant seals, *Mirounga angustirostris*. *Anim Behav* 80:351–361
- Hopp SL, Morton ES (1997) Sound playback studies. In: Hopp SL, Owren MJ, Evans CS (eds) *Animal acoustic communication: sound analysis and research methods*. Springer-Verlag, Berlin, p 323–352
- IACMST (Inter-Agency Committee on Marine Science and Technology) (2006) Report of the IACMST Working Group on underwater sound and marine life. IACMST Working Group Report No. 6. National Oceanography Centre, Southampton
- ICES (International Council for the Exploration of the Sea) (2005) Report of the ad-hoc group on impacts of sonar on cetaceans and fish (AGISC). ICES AGISC 2005, CM 2005/ACE:06. http://ec.europa.eu/environment/nature/conservation/species/whales_dolphins/docs/ices_second_report.pdf
- Isojunno S, Curé C, Kvadsheim PK, Lam FPA, Tyack PL, Wensveen P, Miller PJO (2016) Sperm whales reduce foraging effort during exposure to 1–2 kHz sonar and killer whale sounds. *Ecol Appl* 26:77–93
- Janik VM, Sayigh LS, Wells RS (2006) Signature whistle shape conveys identity information to bottlenose dolphins. *Proc Natl Acad Sci USA* 103:8293–8297
- Jepson PD, Arbelo M, Deaville R, Patterson IAP and others (2003) Gas-bubble lesions in stranded cetaceans: Was sonar responsible for a spate of whale deaths after an Atlantic military exercise? *Nature* 425:575–576
- Johnson MP, Tyack PL (2003) A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE J Oceanic Eng* 28:3–12
- Johnson M, Madsen PT, Zimmer WMX, Aguilar de Soto N, Tyack PL (2006) Foraging Blainville's beaked whales (*Mesoplodon densirostris*) produce distinct click types matched to different phases of echolocation. *J Exp Biol* 209:5038–5050
- Kuningas S, Kvadsheim PH, Lam FPA, Miller PJ (2013) Killer whale presence in relation to naval sonar activity and prey abundance in northern Norway. *ICES J Mar Sci* 70: 1287–1293
- Kvadsheim P, Benders F, Miller PJM, Doksæter L and others (2007) Herring (sild), killer whales (spekkhogger) and sonar—the 3S-2006 cruise report with preliminary results. FFI report 2007/01189. <http://rapporter.ffi.no/rapporter/2007/01189.pdf> (accessed 30 Dec 2015)
- Lorenz K (1937) On the formation of the concept of instinct. *Naturwissenschaften* 25:289–300

- Mann DA, Popper AN, Wilson B (2005) Pacific herring hearing does not include ultrasound. *Biol Lett* 1:158–161
- McCarthy E, Moretti D, Thomas L, DiMarzio N and others (2011) Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar. *Mar Mamm Sci* 27:E206–E226
- McGregor PK (ed) (1992) Playback and studies of animal communication. Plenum Press, New York, NY
- McGregor PK (ed) (2013) Playback and studies of animal communication. (Vol 228) Springer Science & Business Media, New York, NY
- Miller PJO, Biassoni N, Samuels A, Tyack PL (2000) Whale songs lengthen in response to sonar. *Nature* 405:903
- Miller PJO, Johnson MP, Madsen PT, Biassoni N, Quero M, Tyack PL (2009) Using at-sea experiments to study the effects of airguns on the foraging behavior of sperm whales in the Gulf of Mexico. *Deep-Sea Res* 56: 1168–1181
- Miller PJO, Antunes R, Alves AC, Wensveen P and others (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 Tech Rep SOI-2011-001. <http://soi.st-andrews.ac.uk/documents/424.pdf> (accessed 10 Sep 2012)
- Miller PJO, Kvadsheim PH, Lam FPA, Wensveen PJ and others (2012) 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. *Aquat Mamm* 38:362–401
- Miller PJO, Antunes RN, Wensveen PJ, Samarra FI and others (2014) Dose-response relationships for the onset of avoidance of sonar by free-ranging killer whales. *J Acoust Soc Am* 135:975–993
- Miller PJO, Kvadsheim PH, Lam FPA, Tyack PL and others (2015) First indications that northern bottlenose whales are sensitive to behavioral disturbance from anthropogenic noise. *R Soc Open Sci* 2:140484
- Moretti D, Thomas L, Marques T, Harwood J and others (2014) A risk function for behavioral disruption of Blainville's beaked whales (*Mesoplodon densirostris*) from mid-frequency active sonar. *PLOS ONE* 9:e85064
- National Research Council (NRC) (2005) Marine mammal populations and ocean noise: determining when noise causes biologically significant events. National Academy Press, Washington, DC
- Norris KS, Prescott JH, Asa-Dorian PV, Perkins P (1961) An experimental demonstration of echo-location behavior in the porpoise, *Tursiops truncatus* (Montagu). *Biol Bull* 120:163–176
- Nowacek DP, Johnson MP, Tyack PL (2004) North Atlantic right whales (*Eubalaena glacialis*) ignore ships but respond to alerting stimuli. *Proc R Soc B* 271:227–231
- Nowacek DP, Thorne LH, Johnston DW, Tyack PL (2007) Responses of cetaceans to anthropogenic noise. *Mammal Rev* 37:81–115
- Pavan G, Fossati C, Francia C, Zardin F (2010) Acoustic surveys in the Alboran Sea to map the presence and distribution of Cuvier's beaked whales. *J Acoust Soc Am* 128: 2483
- Schorr GS, Falcone EA, Moretti DJ, Andrews RD (2014) First long-term behavioral records from Cuvier's beaked whales (*Ziphius cavirostris*) reveal record-breaking dives. *PLOS ONE* 9:e92633
- Simmonds MP, Lopez-Jurado LF (1991) Whales and the military. *Nature* 351:448
- Simon R, Freidlin B, Rubenstein L, Arbuck SG, Collins J, Christian MC (1997) Accelerated titration designs for phase I clinical trials in oncology. *J Natl Cancer Inst* 89: 1138–1147
- Sivle L, Kvadsheim PH, Curé C, Isojunno S and others (2015) Severity of expert-identified behavioural responses of humpback whale, minke whale and northern bottlenose whale to naval sonar. *Aquat Mamm* 41:469–502
- Southall BL, Gentry RL (2005) Assessment of acoustic exposures on marine mammals in conjunction with USS *Shoup* active sonar transmissions in the eastern Strait of Juan de Fuca and Haro Strait, Washington on 5 May 2003. Technical report released by the US National Oceanographic and Atmospheric Administration, National Marine Fisheries Service, Office of Protected Resources, Silver Spring, MD, January 21, 2005. www.nmfs.noaa.gov/pr/pdfs/acoustics/assessment.pdf (accessed 29 Oct 2016)
- Southall BL, Braun R, Gulland FMD, Heard AD, Baird RW, Wilkin SM, Rowles TK (2006) Hawaiian melon-headed whale (*Peponocephala electra*) mass stranding event of July 3–4, 2004. NOAA Tech Memo NMFS-OPR-31, NOAA, Silver Spring, MD
- Southall BL, Bowles AE, Ellison WT, Finneran JJ and others (2007) Marine mammal noise exposure criteria: Initial scientific recommendations. *Aquat Mamm* 33:411–522
- Southall BL, Berkson J, Bowen D, Brake R and others (2009) Addressing the effects of human-generated sound on marine life: an integrated research plan for US federal agencies. Interagency Task Force on Anthropogenic Sound and the Marine Environment of the Joint Subcommittee on Ocean Science and Technology, Washington, DC. www.whitehouse.gov/sites/default/files/microsites/ostp/oceans-mmnoise-IATF.pdf (accessed 29 Oct 2016)
- Southall BL, Moretti D, Abraham B, Calambokidis J, Tyack PL (2012) Marine mammal behavioral response studies in Southern California: Advances in technology and experimental methods. *Mar Technol Soc J* 46:48–59
- Southall BL, Calambokidis J, Barlow J, Moretti D and others (2014) Biological and behavioral response studies of marine mammals in Southern California, 2013 (SOCAL-13). Final project report, 20 April 2014. http://sea-inc.net/assets/pdf/SOCAL13_final_report.pdf (accessed 30 Dec 2015)
- Southall BL, Quick N, Hastie G, Tyack PL, Boyd I (in press) Mitigation of harm during a novel behavioral response study involving active sonar and wild cetaceans. *J Cetacean Res Manag*
- Stimpert AK, DeRuiter SL, Southall BL, Moretti DJ and others (2014) Acoustic and foraging behavior of a Baird's beaked whale, *Berardius bairdii*, exposed to simulated sonar. *Sci Rep* 4:7031
- Stimpert AK, DeRuiter SL, Falcone EA, Joseph J and others (2015) Sound production and associated behavior of tagged fin whales (*Balaenoptera physalus*) in the Southern California Bight. *Anim Biotelem* 3:23
- Szymanski MD, Bain DE, Kiehl K, Pennington S, Wong S, Henry KR (1999) Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioral audiograms. *J Acoust Soc Am* 106:1134–1141
- Tinbergen N (1963) On aims and methods of ethology. *Z Tierpsychol* 20:410–433

- Tyack PL (1983) Differential response of humpback whales, *Megaptera novaeangliae*, to playback of song or social sounds. *Behav Ecol Sociobiol* 13:49–55
- Tyack P (2009) Acoustic playback experiments to study behavioral responses of free-ranging marine animals to anthropogenic sound. *Mar Ecol Prog Ser* 395: 187–200
- Tyack PL, Gordon J, Thompson D (2003) Controlled exposure experiments to determine the effects of noise on large marine mammals. *Mar Technol Soc J* 37:41–53
- Tyack PL, Johnson M, Soto NA, Sturlese A, Madsen PT (2006) Extreme diving of beaked whales. *J Exp Biol* 209: 4238–4253
- Tyack PL, Zimmer WMX, Moretti D, Southall BL and others (2011) Beaked whales respond to simulated and actual navy sonar. *PLOS ONE* 6:e17009
- Visser F, Miller P, Antunes R, Oudejans M and others (2014) The social context of individual foraging behaviour in long finned pilot whales. *Behaviour* 151:1453–1477
- von Benda-Beckmann AM, Wensveen PJ, Kvadsheim PH, Lam FPA, Miller PJO, Tyack PL, Ainslie MA (2014) Modeling effectiveness of gradual increases in source level to mitigate effects of sonar on marine mammals. *Conserv Biol* 28:119–128
- Wensveen PJ (2016) Detecting, assessing, and mitigating the effects of naval sonar on cetaceans. PhD thesis, School of Biology, University of St Andrews. <http://hdl.handle.net/10023/8684> (accessed 29 Oct 2016)
- Wensveen PJ, von Benda-Beckmann AM, Ainslie MA, Lam FPA, Kvadsheim PH, Tyack PJ, Miller PJO (2015) How effectively do horizontal and vertical response strategies of long-finned pilot whales reduce sound exposure from naval sonar? *Mar Environ Res* 106:68–81
- WWF-Norway (2001) Lack of environmental assessment resulted in negative impact on fish and tourism. Letter from WWF-Norway to Norwegian Minister of Defense, MOD archive no 01/01309-1 (in Norwegian)
- Yack TM, Barlow J, Calambokidis J, Southall BL, Coates S (2013) Identification of previously unknown beaked whale habitat in the Southern California Bight using a towed hydrophone array. *J Acoust Soc Am* 134: 2589–2595

Editorial responsibility: Wendy Dow Piniak (Guest Editor), Gettysburg, USA

*Submitted: January 4, 2016; Accepted: August 22, 2016
Proofs received from author(s): October 31, 2016*