



Monitoring population-level responses of marine mammals to human activities

ERICA FLEISHMAN,¹ John Muir Institute of the Environment, The Barn, One Shields Avenue, University of California, Davis, California 95616, U.S.A.; **DANIEL P. COSTA**, Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California 95060, U.S.A.; **JOHN HARWOOD**, Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews KY16 8LB, United Kingdom; **SCOTT KRAUS**, New England Aquarium, 1 Central Wharf, Boston, Massachusetts 02110, U.S.A.; **DAVID MORETTI**, Naval Undersea Warfare Center, 1176 Howell Street, Newport, Rhode Island 02841, U.S.A.; **LESLIE F. NEW**, Department of Mathematics, Washington State University, Vancouver, Washington 98686, U.S.A.; **ROBERT S. SCHICK**, Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews KY16 8LB, United Kingdom; **LISA K. SCHWARZ**, Institute of Marine Sciences, University of California, Santa Cruz, California 95060, U.S.A.; **SAMANTHA E. SIMMONS**, Marine Mammal Commission, 4340 East-West Highway, Suite 700, Bethesda, Maryland 20814, U.S.A.; **LEN THOMAS**, Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews KY16 8LB, United Kingdom; **RANDALL S. WELLS**, Chicago Zoological Society, % Mote Marine Laboratory, 1600 Ken Thompson Parkway, Sarasota, Florida 34236, U.S.A.

ABSTRACT

We provide guidance for monitoring whether human activities affect the physiology or behavior of marine mammals and, if so, whether those effects may lead to changes in survival and reproduction at the population level. We suggest that four elements be included in designing and implementing such a monitoring program. The first is development of a theory of change: a set of mechanistic hypotheses that outline why a given activity might be expected to have one or more measurable effects on individuals and populations, and ideally the magnitude, timing, and duration of the effects. The second element, definition of biologically meaningful effect sizes, ultimately facilitates the development of a monitoring program that can detect those magnitudes of effect with the desired levels of precision. The third element, selection of response variables for monitoring, allows inference to whether observed changes in the status of individuals or populations are attributable to a given activity. Visual observations, passive acoustic and tagging instruments, and direct physical measurements all can provide data that facilitate quantitative hypothesis testing. The fourth element is specification of the temporal sequence of monitoring. These elements also can be used to inform monitoring of the responses of other taxonomic groups to human activities.

¹Corresponding author (e-mail: efleishman@ucdavis.edu).

Key words: acoustics, environmental impact statements, Marine Mammal Protection Act, results chains, take, vital rates.

As the number and intensity of human activities in marine ecosystems have expanded, researchers, regulators, and stakeholders increasingly have aimed to monitor the responses of marine mammals to these activities. Biological monitoring (*i.e.*, detection of status and trend at any given level of biological organization; Busch and Trexler 2002) typically is designed to achieve one of four objectives (Noon 2002), which are not mutually exclusive. The first is to determine whether the biological effects of a given activity comply with laws and regulations. The second is to evaluate additional hypotheses—generally not restricted to addressing compliance—about the short-term or long-term biological effects of an activity. The third is to assess biological condition over time, whether in association with or independent from one or more activities. The fourth is to differentiate between biological responses to natural environmental variation and human activities. Here, we present four elements to be considered when monitoring whether one or more activities affect the physiology or behavior of marine mammals and, if so, whether those effects may lead to changes in vital rates (survival and reproduction) of a population. Although we focus on monitoring in the context of the laws and regulations of the United States, the methods are independent of jurisdiction.

The U.S. Marine Mammal Protection Act of 1972 (MMPA; 16 U.S.C. § 1361 *et seq.*) is one of the main laws to constrain human activities that may take (*i.e.*, harass, hunt, capture, or kill; or to attempt to harass, hunt, capture, or kill; 16 U.S.C. § 1362[13]) marine mammals in U.S. waters and by U.S. citizens on the high seas. When enacted in 1972, the MMPA was unique with respect to the protections it gave to a group of organisms in relation to their environment. The MMPA mandates that marine mammals be maintained as significant functioning elements of their ecosystems and that optimum sustainable populations be maintained. An optimum sustainable population is defined as “the number of animals which will result in the maximum productivity of the population or the species, keeping in mind the carrying capacity of the habitat and the health of the ecosystem of which they form a constituent element” [16 U.S.C. § 1362(9)]. Implementing regulations elaborate on this definition (50 C.F.R. § 216.3). To achieve these mandates, the MMPA prohibits the unauthorized taking of marine mammals [16 U.S.C. § 1362(13)].

Only those takings and the environmental changes that affect survival or reproduction in the wild are relevant to quantitative assessment of population-level effects (NRC 2013). Much recent attention has focused on assessing whether individual-level responses of marine mammals to anthropogenic underwater sounds may have population-level effects (*e.g.*, Christiansen *et al.* 2013*a, b*; Costa 2012; Costa *et al.* 2013; New *et al.* 2013, 2014; Pirota *et al.* 2014). These sounds not only may cause temporary or permanent loss of hearing or shifts in hearing thresholds (reviewed in Southall *et al.* 2007) but may reduce the ability of marine mammals to communicate, navigate, and forage (Richardson *et al.* 1995). Human activities that produce underwater sounds include boating and shipping, all phases of oil and gas exploration and production, construction and operation of energy facilities or coastal infrastructure (*e.g.*, pile driving, demolition), military sonar operations, and other military training and operations.

A panel of scientists previously outlined a series of transfer functions to link individual-level responses to disturbance to vital rates (fig. 1 in NRC 2005) and gauged the extent to which data and statistical methods were sufficient to parameterize those

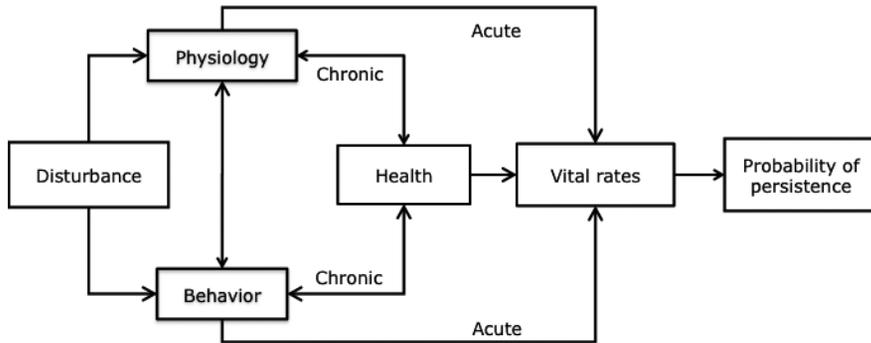


Figure 1. Conceptual model of the process by which physiological or behavioral responses of individuals to disturbance might propagate to the population level, as mediated by health. Modified from New *et al.* 2014.

functions. New *et al.* (2014) developed conceptual models of these transfer functions, translated the transfer functions into a formal mathematical structure, and used empirical data to parameterize exploratory models of the transfer functions (Fig. 1). These conceptual and quantitative models can accommodate physiological or behavioral responses of marine mammals to acute or chronic disturbances (*e.g.*, Moretti *et al.* 2014), including but not limited to sound. The models that account for chronic disturbance link individual-level vital rates to short-term changes in the behavior and physiology of individuals *via* health (all internal factors that affect homeostasis).

As a proof of concept of the models, New *et al.* (2014) used field data on diving patterns and body composition of southern elephant seals (*Mirounga leonina*) to assess how a hypothetical disturbance that reduced foraging efficiency might affect the survival of 1-yr-old animals. Foraging efficiency affects the condition of female southern elephant seals as measured by body mass. Body mass is associated with pup mass at weaning (Arnbom *et al.* 1993) and subsequent pup survival (McMahon *et al.* 2000, 2003). By modeling the effect on vital rates of a gradient of reduction in foraging, and assuming that foraging by all individuals in the population was reduced, New *et al.* (2014) predicted the likelihood that disturbances of varying duration, either in 1 yr or over 30 yr, would lead to changes in population size and in an individual's number of descendants relative to that of other individuals in the population. A similar analysis of data on northern elephant seals (*Mirounga angustirostris*) included a transfer function that related reproduction to reductions in foraging (Costa *et al.* 2016). Costa *et al.* (2016) estimated increases in fat mass at daily resolution on the basis of empirical data on changes in the elephant seals' vertical drift rate during dives; vertical drift rate was modeled as a function of buoyancy from fat mass (Schick *et al.* 2013a). Their model projections suggested that an inability to forage over extensive areas (in this case, 7,850 km²) during a foraging trip would not affect vital rates (Costa *et al.* 2016).

Ideally, models for other species in which measures of health are used to relate disturbance to vital rates also will be parameterized with robust data on physiology, behavior, and long-term abundance and movements (*e.g.*, data from mark-recapture studies). For example, similar measurements of changes in buoyancy are possible for other marine mammals that make drift dives, especially large whales (Beck *et al.* 2000, Williams *et al.* 2000, Nowacek *et al.* 2001, Miller *et al.* 2004,

Nousek-McGregor *et al.* 2014). However, such data are not universally available and sometimes are difficult to collect, especially over long time periods. A high proportion of behavioral measurements that commonly are incorporated into models rely on attachment of electronic transmitters or data-collection devices to individual animals. Sample sizes may be small, and uncertainties about the meaning or population-level effects of changes in behavior that are based on remotely sensed data may be high. Accordingly, we aim to provide guidance for monitoring whether one or more human activities affect the physiology or behavior of marine mammals and, if so, whether those effects may lead to changes in survival and reproduction. We suggest four elements to include in designing such a monitoring program: development of a theory of change, definition of biologically meaningful effect sizes (magnitudes of effect), selection of response variables for monitoring, and specification of the temporal sequence of monitoring.

We introduce these elements in the context of the requirements of the MMPA, and illustrate their implementation with examples. Nevertheless, the elements easily can be adapted to effects assessments or the design of monitoring programs to inform decision-making and management under other statutes (*e.g.*, the U.S. National Environmental Policy Act or the U.S. Department of Agriculture Forest Service's requirement to maintain viable populations of species of conservation concern [36 C.F.R. § 219]). Furthermore, although we emphasize marine mammals, the guidance is transferable among species, ecosystems, and natural and anthropogenic environmental changes.

Development of a Theory of Change

We suggest that monitoring of population-level effects of human activities be informed by a theory of change—a set of hypotheses about the mechanisms by which a given activity might be expected to have a given, measurable biological effect, whether negative or positive (*e.g.*, Stem *et al.* 2005). Use of theories of change is well-established in the field of program evaluation (the systematic assessment of the implementation or results of a program) (*e.g.*, Bonner 2003, Auspos and Kubisch 2004, Blamey and Mackenzie 2007). *De facto* theories of change have been developed to explain why a given human activity or class of activities, especially sound-producing activities, may have a negative effect on a population of marine mammals (*e.g.*, NRC 2005, Parks *et al.* 2007, Miller *et al.* 2015). Theories of change also have been developed in the context of management actions intended to reduce the undesirable effects of human activities on marine mammals. For example, the broadcasting of neutral or warning sounds prior to loud, impulsive sounds may allow animals to move away from the sound sources while the level of sound is still relatively low (Richardson *et al.* 1995, Jefferson and Curry 1996, Miller *et al.* 2009).

Theories of change can be communicated in many ways, including the use of conceptual models and results chains (Margoluis *et al.* 2013). A conceptual model presents mechanistic hypotheses about the relations between a response variable and the anthropogenic or natural covariates that directly and indirectly affect the response variable. A results chain presents hypotheses about the mechanisms by which a given intervention (*i.e.*, an action intended to improve a particular situation) will ameliorate the undesirable effects of a human activity and thereby improve the status of the response variable (Margoluis *et al.* 2013). Even if no intervention is planned, and thus development of a results chain is not warranted, theories of change are applicable to the design of monitoring to assess population-level effects of human activities.

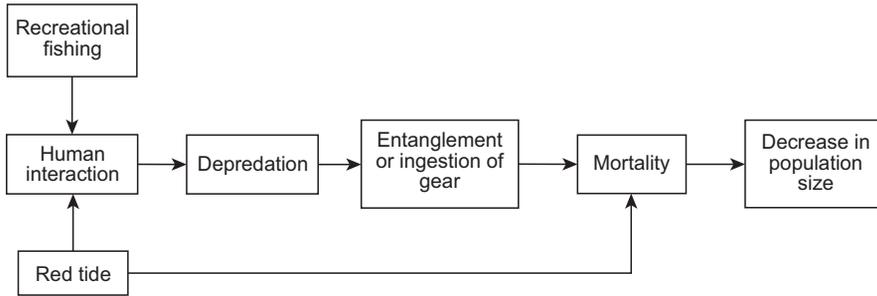


Figure 2. Conceptual model of the mechanisms by which predation by common bottlenose dolphins (*Tursiops truncatus*) of bait or catch from recreational fishing gear in Sarasota Bay, Florida could have population-level effects (see Powell and Wells 2011).

Conceptual models and results chains are included in a set of open-source, standardized methods for project design, management, and monitoring that were developed by the Conservation Measures Partnership (<http://www.conservationmeasures.org>), a group of more than 20 international conservation organizations. Results chains often are more explicit than conceptual models (Margoluis *et al.* 2013). For example, results chains specify not only the hypothesized indirect and direct effects of a disturbance and which effects are linked, but the hypothesized direction of each intermediate effect (*e.g.*, an increase or decrease), and ideally, the hypothesized magnitude and timing of the effect. In some cases, one can use simulation modeling to identify variables or relations in the theory of change that will reduce uncertainty to the greatest extent. One also can apply value-of-information analyses to identify which uncertainties are most relevant to decision-making (Runge *et al.* 2011, Moore and Runge 2012).

To illustrate how theories of change could inform development of a monitoring program, we developed a conceptual model of the physiological and behavioral responses of long-term resident common bottlenose dolphins in Sarasota Bay, Florida, to recreational fishing and red tides (Fig. 2). These two disturbances can affect survival of the dolphins either indirectly, by reducing health, or directly; interactions with boaters, anglers, fishing vessels, or piers can lead to entanglement in, or ingestion of, recreational fishing gear. The yearly proportion of instances in which dolphins in this population engaged in patrolling, begging, scavenging, actual or attempted depredation of bait and fish, and provisioning (collectively referenced as *human interactions*) increased by about 400% (from about 0.075 to about 0.325) from 2000 through 2007 (Powell and Wells 2011). During this period, the cumulative proportion of the population that engaged in human interactions or were entangled at least once increased by about 700% (from about 2% to about 14%) (Powell and Wells 2011). In 2006, deaths of five dolphins that were attributed to entanglement or ingestion led to a 2% decrease in the size of the resident population of dolphins in Sarasota Bay (Powell and Wells 2011). Human interactions may become more prevalent during or shortly following red tides (blooms of the toxic dinoflagellate *Karenia brevis*), which deplete the dolphins' prey base (Powell and Wells 2011). Red tides also increase mortality of dolphins in Sarasota Bay regardless of human interactions.

Anglers in Florida currently are required to immediately release fish that are below a certain size or that are not in season (*i.e.*, regulatory discards). We developed a

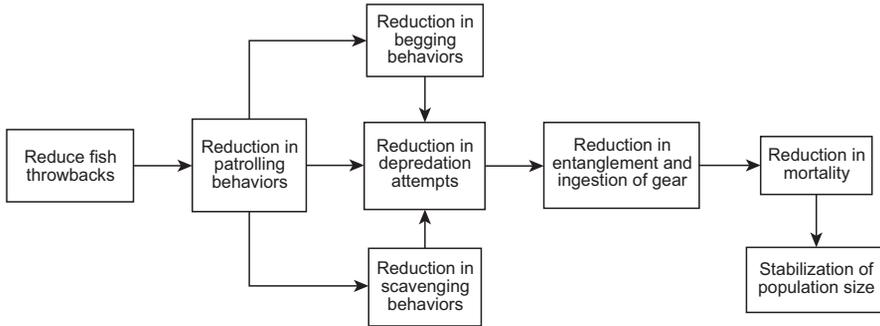


Figure 3. Results chain illustrating the hypothesized effect of minimizing the potential for predation of bait or catch from recreational fishing gear by common bottlenose dolphins (*Tursiops truncatus*) in Sarasota Bay, Florida (see Powell and Wells 2011).

results chain to present a hypothetical scenario in which modified laws permit anglers to maintain fish temporarily in a live well or other device rather than immediately throwing fish back into the open water, and then to release the fish when dolphins are not present (Fig. 3). If humans no longer serve as sources of food, then dolphins might gradually spend less time patrolling and scavenging. If humans do not intentionally feed the dolphins, an activity that is illegal under the MMPA, the animals might spend less time begging. The incentive for depredation also might decrease. These hypothetical changes in dolphin behavior might reduce entanglement, ingestion of gear, boat strikes, other serious injuries, and mortality. In years with red tides and therefore a reduction in prey abundance, human interactions might increase relative to years without red tides, but to a lesser extent than if the reduction in regulatory discards had not been implemented. As described below, reliable methods exist for monitoring each step in the results chain.

Definition of Biologically Meaningful Effect Sizes

The second element in monitoring population-level effects of human activities is the definition of biologically meaningful effect sizes (magnitudes of responses). Documenting a physiological or behavioral response of one or more individuals to an activity precedes understanding the activity's potential effect on a population. But simply documenting a response is insufficient to conclude that the response is biologically meaningful (NRC 2013).

The dearth of explicitly defined levels of biological effects that will trigger changes in human activities (Nie and Schultz 2012) complicates implementation of monitoring and adaptive management for virtually all taxonomic groups and situations. Regulations rarely include unambiguous definitions of biologically meaningful effect sizes. For example, the National Oceanic and Atmospheric Administration's National Marine Fisheries Service (NOAA Fisheries; for cetaceans and pinnipeds) and the U.S. Fish and Wildlife Service (USFWS; for walrus, polar bears, sea otters, and manatees) may authorize an activity that will kill or harass marine mammals if the activity will have a "negligible" effect on rates of recruitment or survival and on subsistence uses by Alaskan natives (16 U.S.C. § 1361 *et seq.*). Yet negligible is not defined quantitatively in the MMPA as, for example, a given magnitude of change in a population's growth rate or vital rates. Federal regulations define *negligible* as either not reasonably

expected or unlikely to have an adverse effect on annual recruitment or survival (50 C.F.R. 216 Section 103), but *adverse* is not quantified.

The basis for quantitatively defining concepts such as negligible impact, adverse impact, recovery, or endangered usually is an acceptable level of risk (*i.e.*, a probability), also referenced as a risk characterization. Society may consider some level of change in vital rates to be tolerable. For example, member states of the European Union suggested that an average annual decrease in population size of $>1\%$ over 12 yr would lead to unfavorable conservation status, a potential violation of the European Union's habitats directive (92/43/EEC). The latter implies that an annual decrease in population size of $\leq 1\%$ over 12 yr is tolerable. Although science can inform selection of the acceptable level of risk, the acceptable level fundamentally is a societal value that is predicated on ethical judgments or personal policy preferences (Wilhere 2012).

Although common, it is not best practice to measure change and then decide whether the level of change is acceptable. Collection of data on physiological or behavioral responses to disturbance will yield more insight if biologically meaningful effects are defined before monitoring begins, and those definitions used to specify allowable levels of take (Murphy and Weiland 2010). Models of population growth, which are parameterized with estimates of demographic rates, allow one to gauge how variation in survival and reproduction affect a population's probability of persistence over a given period of time. Once an allowable level of take has been specified, outputs from a population model also can inform the design of a monitoring program with the capacity to detect the associated effect size (the defined magnitude of the biological effect) with desired levels of precision (Noon 2002). For example, one could apply a population model to assess the extent to which a 2% annual decrease—or a positive or negative change of any other magnitude—in the size of the resident population of dolphins in Sarasota Bay (Powell and Wells 2011) is likely to affect the population's probability of extinction over 20, 50, or 100 yr.

Monitoring of the distribution, abundance, and vital rates of many populations of marine mammals is hampered by long periods in which animals are not available for detection, extremely low detection probabilities (the probability of detecting an organism given its presence; Borchers *et al.* 2002), and limited feasibility of handling. If the true abundance or density of the species is relatively low (*e.g.*, when the species is endangered), then statistical power or certainty is likely to remain low even if sampling effort is quite high (Taylor and Gerrodette 1993). One way to address low statistical power is to include a level of uncertainty in the decision rule itself and quantify uncertainty throughout the monitoring process. Thus, if societal values or regulations warrant, decision rules for populations with low statistical power (high levels of uncertainty) can be biased toward conferral of protection (Regan *et al.* 2013).

It generally will be necessary to identify acceptable levels of population-level change on a case-by-case basis. As the guidelines for implementation of the California Environmental Quality Act recognize, "An ironclad definition of significant effect is not always possible because the significance of [the effect of] an activity may vary with the setting" (Title 14 California Code of Regulations § 15000 *et seq.*). Additionally, an activity may not have the same effects under different circumstances (Ellison *et al.* 2012). For example, responses of marine mammals to sound are affected by the spatial and temporal relations between the sound source and the receiving animal, the animal's behavior given its previous experience, and the similarity of the sound to other biologically meaningful sounds (Ellison *et al.* 2012).

The numerous studies of behavioral responses to sound have employed diverse measures, from feeding activity to group size and composition (e.g., Curé *et al.* 2015, Samarra and Miller 2016, Harris *et al.* 2016). Harris and Thomas (2015) reviewed the status of research on behavioral responses to sonar, including controlled experiments on the exposure of both captive animals and free-ranging animals to either simulated sources of military sonar or real sources of military sonar, observational studies (*via* tags, passive acoustic monitoring, and visual observations) of responses to sonar and vessels, and predator playback studies. Collectively, these studies addressed the possibility that behavioral responses to sound could affect foraging, blood and tissue condition, intraspecific communication, and movement. A review of methods for collection of data and analysis of acoustic sequences (Kershenbaum *et al.* 2016) also illustrates the range of metrics available for evaluation of behavioral responses, the associated uncertainties, and, by extension, the necessary investments.

Selection of Variables to Monitor

Selection of response variables and covariates is the third element in monitoring whether changes in biological status are attributable to a given activity. Ultimately, assessment of population-level effects of human activities requires monitoring of population status—the end-points in a conceptual model or results chain. However, monitoring population-level responses to human activities also requires time-series data on, at minimum, the intensity, duration and frequency of the activity within a given space; population size; and the proportion and attributes (e.g., life stages) of the population that may be exposed to the activity over its duration or over a generation. The most useful response variables are those that are both responsive to the activity and strongly related to the vital rates of the species (Noon 2002) (Table 1). The biology of the species, the hypothesized response to the activity, and existing data or monitoring programs can inform selection of variables.

The cost of monitoring a given variable can vary over orders of magnitude among species. For example, visual observations or visual assessments of health of Atlantic harbor seals (*Phoca vitulina*) that haul out at low tide in New England are likely to be far cheaper and more straightforward than visual observations or assessments of beaked whales in the Caribbean, which are deep divers that occur far from the shore and are difficult to detect. Moreover, cost encompasses data processing and storage. Remote collection of data initially may appear cheaper than field operations, but it may require considerable time and expertise to, say, identify vocalizations and the behaviors with which the vocalizations may be associated. The frequency of monitoring that will yield the greatest gain in information depends in part on when and where the species is available to be detected, its detection probability, and connectivity among populations. For data that are collected with attached tags, the possible and ideal frequency of monitoring also is affected by which tags are suitable for a given species and by trade-offs between the temporal resolution of recording or transmission and the lifetime of the tag. For a given species and scenario, simulations (or, in some cases, analytic calculations) can be used to estimate the frequency and duration of monitoring necessary to obtain a given statistical power or level of certainty, and the associated costs.

In and of themselves, data on the number of animals in a given population or area, or the number of animals that might be exposed to a given activity over the short term, are not sufficient to estimate vital rates. Nevertheless, collection of these data is common, especially when the primary intent of monitoring is to meet regulatory

Table 1. Examples of variables that can be monitored to test hypotheses about the mechanisms by which marine mammals may respond to human activities at the individual level and population level (see Fig. 1). These variables are believed to be among the most responsive to human activities or the most strongly related to the vital rates of marine mammals. Population growth rate can be estimated and projected on the basis of survival rate, reproductive rate, length of offspring dependence, age structure, and data obtained from recovered carcasses (*e.g.*, Easterling *et al.* 2000, Caswell 2001, Borchers *et al.* 2002, Newman *et al.* 2014).

Method of data collection	Variables that can be measured directly or derived	Class of variable
Visual observation	Population size	Persistence
	Population density by age and sex in space and time	Behavior
	Age structure	Vital rate
Mark-resight	Population growth rate	Persistence
	Survival (differentiated by age or age class)	Vital rate
	Reproduction (differentiated by age or age class)	Vital rate
	Duration of offspring dependence on adults	Vital rate
	Age structure	Vital rate
Attached tags	Population growth rate	Persistence
	Population size	Persistence
	Travel distance and speed	Behavior
	Drift rate	Health
	Characteristics of dives and fine-resolution movements (some of which may be associated with foraging)	Behavior
	Interactions with humans	Behavior
	Vocalizations (some of which may be associated with foraging)	Behavior
	Conspecific interactions	Behavior
Population density in space and time (differentiated by age and sex)	Behavior	
Visual assessment of health	Body fat	Health
	Skin condition	Health
	Parasite load	Health
	Rake marks	Health
Hands-on assessment of health	Injury	Physiology
	Total mass	Health
	Lipid mass	Health
	Metabolic rate	Physiology
	Hormone levels in hair, skin, blubber, or excretions	Physiology
	Blubber thickness	Health
	Blood chemistry	Physiology
	Pregnancy status	Health
	Hearing	Physiology
	Percent lipid in blubber	Health
Carcass recovery	Composition of prey consumed	Behavior
	Cause of death	Vital rate
Rescue and rehabilitation	All data obtained from a hands-on assessment of the premortality health of a dead animal	Health, physiology
	All data obtained from a hands-on assessment of the health of an animal that is injured or ill	Health, physiology
Behavioral observations	Activity budget	Behavior
	Interactions with humans	Behavior
	Conspecific interactions	Behavior

criteria. For example, NOAA Fisheries currently estimates take from anthropogenic sound as the number of animals that may be exposed to levels of received sound that the agency associates with behavioral disruption or with physical injuries. The precision of estimates of the abundance or density of marine mammals often is constrained by low detection probabilities. The animals generally spend much more time below the surface of the water than at the surface (Taylor *et al.* 2007). Furthermore, density arguably has limited value for population-level inference unless the entire range of the species is sampled given the difficulty in identifying potential source, sink, and pseudosink populations (in a pseudosink, immigration maintains artificially high densities of organisms, but the population will not be extirpated in the absence of immigration) (*e.g.*, Pulliam 1988, Donovan 1995, Watkinson and Sutherland 1995). Still, information on density across a population's range over time may be helpful for estimating the relation between the activity and behavioral change.

For species that use acoustic signaling (*e.g.*, most if not all cetaceans), and for which the spectral and contextual characteristics of vocalizations are known, it may be possible to collect passive acoustic data on behaviors that likely have a functional relation to health (Clark *et al.* 2007, McDonald *et al.* 2012) (Table 1). For example, echolocation clicks associated with the start of deep foraging dives by Blainville's beaked whales (*Mesoplodon densirostris*) were detected before, during, and after mid-frequency active sonar operations in the Bahamas (McCarthy *et al.* 2011; Moretti *et al.* 2010, 2014). Passive acoustic data also are being used to estimate the daily and seasonal abundance and distribution of Blainville's beaked whales (Marques *et al.* 2009) and other species (Marques *et al.* 2013). If tagging of species that use acoustic signaling is feasible, then data on movements, vocalizations, and environmental conditions can be collected when the animals are beyond visual monitoring range (*e.g.*, Nowacek *et al.* 2001, Block *et al.* 2011, Costa *et al.* 2012, Goldbogen *et al.* 2014, Schorr *et al.* 2014) (Table 1). Some of the many variables that can be derived from tagging data include the incidence and rate of foraging, drift rate, levels of exposure to sound and the duration of exposure, and responses to stimuli (Harris and Thomas 2015).

If the biological effects of a given activity are believed to be chronic rather than acute, then a measure of health or condition is essential for assessing population-level effects (Table 1). If animals cannot readily be captured but can be encountered, repeated photographs of the same individuals may allow assessment of health trends (*e.g.*, Wells 2003, Pettis *et al.* 2004, Bradford *et al.* 2012, Hunt *et al.* 2013). For example, visual measures of the health of North Atlantic right whales (*Eubalaena glacialis*) include body fat, skin condition, parasite loads, evidence of vessel strikes or entanglement in fishing gear, and rake marks forward of the blowholes (Hamilton and Marx 2005, Schick *et al.* 2013b, Nousek-McGregor *et al.* 2014, van der Hoop *et al.* 2014). Evidence from elephant seals, right whales, and sperm whales suggests that changes in the drift phase of diving animals may provide insights into relative proportions of body fat and thus body condition (Nowacek *et al.* 2001, Miller *et al.* 2004, Schick *et al.* 2013a, Nousek-McGregor *et al.* 2014). If animals cannot be captured but their hair, skin, blubber or excretions can be sampled, then it may be possible to extract hormones that sometimes reflect reproductive status or chronic stress (*e.g.*, glucocorticoids, progesterone) (Rolland *et al.* 2012, Dickens and Romero 2013, Kellar *et al.* 2015).

Recent advances in tags that include accelerometers may provide direct estimates of changes in animal mass or condition while at sea (Aoki *et al.* 2011, Miller *et al.* 2012). If animals can be captured or recaptured, then direct physical measurements of health also become tractable. For example, estimates of the health of long-term

resident bottlenose dolphins in Sarasota Bay, Florida, include body mass index (Hart *et al.* 2013), blubber thickness, percent lipid in blubber, blood chemistry and hematology, occurrence of lung disease, hormone concentrations, and fetal survival (Wells *et al.* 2004, Schwacke *et al.* 2010, 2012, 2014; Wells *et al.* 2014). Ongoing collection of these data, along with continuing, monthly photographic identification surveys and physical measurements by a local program that responds to marine mammal strandings, would make it feasible to monitor each step in the results chain described above (Fig. 3) (Wells 2009, Powell and Wells 2011). Recreational-fishing effort can be monitored at piers and jetties (Powell and Wells 2011), while red tides are monitored by a network of researchers and federal and state agencies throughout Florida. Human interactions can be monitored *via* focal follows of pairs of dolphins, one of which is known to have become entangled in fishing gear or to interact with humans and one of which is not so known (Powell and Wells 2011). Behavior of recreational fishers before and after implementation of management regulations could be collected *via* surveys or interviews.

Temporal Sequencing of Monitoring

The fourth element in monitoring population-level effects of human activities is specification of the temporal sequence of monitoring. The directions, magnitudes, and time periods of effects hypothesized in theories of change suggest a sequence of measures that can inform the design of a monitoring program. The time period over which physiology or behavior is expected to change, and over which those changes affect health and vital rates, may be uncertain, especially with respect to proposed mitigation actions. For example, there is considerable uncertainty in the time period over which behaviors of bottlenose dolphins in Sarasota Bay might change in response to minimization of regulatory discards, and over which the behavioral changes might affect body condition and mortality (Fig. 3). In this case, the lag time between policy changes and biological changes would depend in part on the extent of human compliance. Hypothesized scenarios of human compliance could be used to estimate upper and lower bounds on these time lags.

Assessment of population-level effects of a given activity is most feasible if one can estimate the range of values of demographic parameters (*e.g.*, survival of different age classes or life stages, probability of reproduction, age at first reproduction) for many years—ideally over decades—in the absence of the activity. In some situations, it is feasible to monitor a population before and after an activity is implemented. The status of the population and its environment prior to the activity is considered to be the baseline, and reflects current and previous responses to diverse disturbances (McCarthy *et al.* 2010; Moretti *et al.* 2010, 2014; Murphy and Weiland 2010). Continued monitoring after initiation of the activity may allow one to quantify relations between the activity, behavioral or physiological changes in response to the activity, and health.

However, monitoring cannot always be initiated before an activity begins. Moreover, because the natural environment and human activities are dynamic, simply comparing state variables before and after a disturbance may not allow inference to causation (*e.g.*, Underwood 1991). Instead, evaluating whether a biological change is attributable to a given action requires assessment against a control, reference population, or counterfactual: what would have happened—potentially including other types of activities—if the activity in question had not been implemented (Ferraro 2009).

Assessment against a counterfactual requires statistical comparison of data from the group or population exposed to a given activity with data from another group or population that is ecologically similar and has a similar environment but is not exposed to the activity (*e.g.*, Kraus *et al.* 2001, Abadie and Imbens 2006, Ho *et al.* 2007, Nolte and Agrawal 2013). For example, Claridge (2013) applied visual mark-recapture density estimation methods to beaked whales in the Northwest Providence Channel off the coast of Abaco, Bahamas, and at the Atlantic Undersea Test and Evaluation Center (AUTEK) in the Tongue of the Ocean, approximately 100 km south of Abaco. The beaked whales at AUTEK were exposed to mid-frequency active sonar operations. Neither animals from Abaco nor from AUTEK were resighted outside the area in which they originally were observed. Samples of DNA and fatty-acid assays similarly suggested that the two groups of beaked whales are separate populations. Such observations are providing biological measurements including group size, gestation time, time to weaning, time to parturition, and the ratio of adult females to dependent juveniles and calves. These measurements, in turn, are informing the development of a model of energetics that can be used to examine the potential cumulative effects of sonar exposure (New *et al.* 2013). It may be possible to simulate a matched population if a comparable population does not exist or there are considerable operational impediments to measurement of the comparable populations.

The state variables that are monitored may change over time as one tests the assumptions of the theory of change. If the activity does not appear to cause changes in physiology or behavior, then measures of health or condition may not be necessary. If physiology or behavior appears to be changing in response to the activity but health is stable, then parameterization of transfer functions that relate health to vital rates may not be necessary. If health appears to be changing, then it may be advisable to collect data that test which mechanisms might be driving the changes.

Summary

Diverse regulators and actors seek practical ways of monitoring the biological effects of proliferating human activities in marine environments. We believe it is feasible to implement monitoring that assesses the chain of potential relations from initiation of a human activity to population dynamics—from physical and behavioral responses to the activity, to shifts in health, to changes in vital rates. Such monitoring data support development of population models, which aid interpretation of the biological meaning of a given change or difference in a demographic rate (*e.g.*, the maximum effect on the population growth rate of a given change in survival). Estimates of survival, reproduction, and physical growth can be used to parameterize models of the growth of closed or open populations. The many available methods for building population models (*e.g.*, Easterling *et al.* 2000, Caswell 2001, Borchers *et al.* 2002, Newman *et al.* 2014) provide options for species with diverse detection probabilities, abundances, demographic rates, life histories, and biological traits. Population models help elucidate both the potential effects of variation in demographic rates and the necessary level of confidence in estimates of demographic rates to infer population-level effects. The simulations and empirical parameterization that are made possible by monitoring along the chain noted above allow one not only to address the requirements of take authorizations (such as the in the U.S. MMPA) but to increase understanding of ecological mechanisms, reduce the likelihood of future undesirable effects, and facilitate maintenance of species as functioning elements of their ecosystems.

ACKNOWLEDGMENTS

This work was supported by an award from the Office of Naval Research (N00014-12-1-0274-0) to the University of California, Davis. Our work also benefitted from discussions of a working group on establishment of a monitoring system to identify and evaluate potential adverse impacts caused by the Compliance Offset Protocol for U.S. Forest Projects, which was supported by a contract from the California Air Resources Board to M.W. Schwartz and E. Fleishman.

LITERATURE CITED

- Abadie, A., and G. W. Imbens. 2006. Large sample properties of matching estimators for average treatment effects. *Econometrica* 74:235–267.
- Aoki, K., Y. Y. Watanabe, D. E. Crocker, *et al.* 2011. Northern elephant seals adjust gliding and stroking patterns with changes in buoyancy: Validation of at-sea metrics of body density. *Journal of Experimental Biology* 214:2973–2987.
- Arnbom, T., M. A. Fedak, I. L. Boyd and B. J. McConnell. 1993. Variation in weaning mass of pups in relation to maternal mass, postweaning fast duration, and weaned pup behaviour in southern elephant seals (*Mirounga leonina*) at South Georgia. *Canadian Journal of Zoology* 71:1772–1781.
- Auspos, P., and A. Kubisch. 2004. Building knowledge about community change: Moving beyond evaluations. The Aspen Institute, Washington, DC. Available at <http://www.aspeninstitute.org/sites/default/files/content/docs/rcc/BUILDINGKNOWELDGE.pdf>.
- Beck, C. A., W. D. Bowen and S. J. Iverson. 2000. Seasonal changes in buoyancy and diving behaviour of adult grey seals. *Journal of Experimental Biology* 203:2323–2330.
- Blamey, A., and M. Mackenzie. 2007. Theories of change and realistic evaluation. *Evaluation* 13:439–455.
- Block, B. A., I. D. Jonsen, S. J. Jorgensen, *et al.* 2011. Tracking apex marine predator movements in a dynamic ocean. *Nature* 475:86–90.
- Bonner, L. 2003. Using theory-based evaluations to build evidence-based health and social care policy and practice. *Critical Public Health* 13:77–92.
- Borchers, D. L., W. E. Stephens and W. Zucchini. 2002. Estimating animal abundance: Closed populations. Springer, London, U.K.
- Bradford, A. L., D. W. Weller, A. E. Punt, Y. V. Ivashchenko, A. M. Burdin, G. R. VanBlaricom and R. L. Brownell, Jr. 2012. Leaner leviathans: Body condition variation in a critically endangered whale population. *Journal of Mammalogy* 93:251–266.
- Busch, D. E., and J. C. Trexler. 2002. The importance of monitoring in regional ecosystem initiatives. Pages 1–23 *in* D. E. Busch and J. C. Trexler, eds. *Monitoring ecosystems: Interdisciplinary approaches for evaluating ecoregional initiatives*. Island Press, Washington, DC.
- Caswell, H. 2001. *Matrix population models: Construction, analysis, and interpretation*. Sinauer, Sunderland, MA.
- Christiansen, F., M. H. Rasmussen and D. Lusseau. 2013*a*. Whale watching disrupts feeding activities of minke whales on a feeding ground. *Marine Ecology Progress Series* 478:239–251.
- Christiansen, F., M. H. Rasmussen and D. Lusseau. 2013*b*. Inferring activity budgets in wild animals to estimate the consequences of disturbances. *Behavioral Ecology* 24:1415–1425.
- Claridge, D. E. 2013. Population ecology of Blainville's beaked whales (*Mesoplodon densirostris*). Ph.D. dissertation, University of St Andrews, St Andrews, Scotland. 296 pp.
- Clark, C. W., D. Gillespie, D. P. Nowacek and S. E. Parks. 2007. Listening to their world: Acoustics for monitoring and protecting right whales in an urbanized ocean. Pages

- 333–357 in S. D. Kraus and R. M. Rolland, eds. *The urban whale*. Harvard University Press, Cambridge, MA.
- Costa, D. P. 2012. A bioenergetics approach to developing the PCAD model. Pages 423–426 in A. N. Popper and A. Hawkins, eds. *The effects of noise on aquatic life*. *Advances in Experimental Medicine and Biology*. Springer, New York, NY.
- Costa, D. P., G. A. Breed and P. W. Robinson. 2012. New insights into pelagic migrations: Implications for ecology and conservation. *Annual Review of Ecology, Evolution, and Systematics* 43:73–96.
- Costa, D. P., L. K. Schwarz, J. Maresh, P. W. Robinson and D. E. Crocker. 2013. A bioenergetics approach to understanding the population consequences of natural and anthropogenic disturbance. *Integrative and Comparative Biology* 53:E41–E41.
- Costa, D. P., L. Schwarz, P. W. Robinson, *et al.* 2016. A bioenergetics approach to understanding the population consequences of disturbance: Elephant seals as a model system. Pages 161–170 in A. N. Popper and A. Hawkins, eds. *Effects of noise on aquatic life II*. Springer, New York, NY.
- Curé, C., L. D. Sivle, F. Visser, *et al.* 2015. Predator sound playbacks reveal strong avoidance responses in a fight strategist baleen whale. *Marine Ecology Progress Series* 526:267–282.
- Dickens, M. J., and L. M. Romero. 2013. A consensus endocrine profile for chronically stressed wild animals does not exist. *General and Comparative Endocrinology* 191:177–189.
- Donovan, T. M., F. R. Thompson III, J. Faaborg and J. R. Probst. 1995. Reproductive success of migratory birds in habitat sources and sinks. *Conservation Biology* 9:1380–1395.
- Easterling, M. R., S. P. Ellner and P. M. Dixon. 2000. Size-specific sensitivity: Applying a new structured population model. *Ecology* 81:694–708.
- Ellison, W. T., B. L. Southall, C. W. Clark and A. S. Frankel. 2012. A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Conservation Biology* 26:21–28.
- Ferraro, P. J. 2009. Counterfactual thinking and impact evaluation in environmental policy. *New Directions for Evaluation* 122:75–85.
- Goldbogen, J. A., A. K. Stimpert, S. L. DeRuiter, *et al.* 2014. Using accelerometers to determine the calling behavior of tagged baleen whales. *Journal of Experimental Biology* 217:2449–2455.
- Hamilton, P. K., and M. K. Marx. 2005. Skin lesions on North Atlantic right whales: Categories, prevalence and change in occurrence in the 1990s. *Diseases of Aquatic Organisms* 68:71–82.
- Harris, C. M., and L. Thomas. 2015. Status and future research on the behavioral responses of marine mammals to U.S. Navy sonar. CREEM Technical Report 2015-3, University of St Andrews, Scotland. Available at <http://hdl.handle.net/10023/7741>.
- Harris, C. M., L. Thomas, D. Sadykova, *et al.* 2016. The challenges of analyzing behavioural response study data: An overview of the MOCHA (Multi-study Ocean acoustics Human effects Analysis) project. Pages 399–407 in A. N. Popper and A. Hawkins, eds. *Effects of noise on aquatic life II*. Springer, New York, NY.
- Hart, L. B., R. S. Wells and L. H. Schwacke. 2013. Reference ranges for body condition in wild bottlenose dolphins (*Tursiops truncatus*). *Aquatic Biology* 18:63–68.
- Ho, D., K. Imai, G. King and E. Stuart. 2007. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis* 15:199–236.
- Hunt, K. E., M. J. Moore, R. M. Rolland, *et al.* 2013. Overcoming the challenges of studying conservation physiology in large whales: A review of available methods. *Conservation Physiology* 1: doi:10.1093/conphys/cot006.
- Jefferson, T. A., and B. E. Curry. 1996. Acoustic methods of reducing or eliminating marine mammal–fishery interactions: Do they work? *Ocean & Coastal Management* 31:41–70.

- Kellar, N. M., K. Catelani, M. Robbins, M. Trego, C. Allen, K. Danil and S. J. Chivers. 2015. Blubber cortisol: A potential tool for assessing stress response in free-ranging dolphins without effects due to sampling. *PLOS ONE* 10(2):e0115257.
- Kershenbaum, A., D. T. Blumstein, M. A. Roch, *et al.* 2016. Acoustic sequences in non-human animals: A tutorial review and prospectus. *Biological Reviews* 91:13–52.
- Kraus, S. D., P. K. Hamilton, R. D. Kenney, A. R. Knowlton and C. K. Slay. 2001. Reproductive parameters of the North Atlantic Right Whale. *Journal of Cetacean Research and Management* 2:231–236.
- Margoluis, R., C. Stem, V. Swaminathan, *et al.* 2013. Results chains: A tool for conservation action design, management, and evaluation. *Ecology and Society* 18(3):22. Available at <http://dx.doi.org/10.5751/ES-05610-180322>.
- Marques, T. A., L. Thomas, J. Ward, N. DiMarzio and P. L. Tyack. 2009. Estimating cetacean population density using fixed passive acoustic sensors: An example with beaked whales. *Journal of the Acoustical Society of America* 125:1982–1994.
- Marques, T. A., L. Thomas, S. W. Martin, *et al.* 2013. Estimating animal population density using passive acoustics. *Biological Reviews* 88:287–309.
- McCarthy, E., D. Moretti, L. Thomas, *et al.* 2011. Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar. *Marine Mammal Science* 27: E206–E226.
- McDonald, T. L., W. J. Richardson, C. R. Greene, Jr., S. B. Blackwell, C. S. Nations, R. M. Nielson and B. Streever. 2012. Detecting changes in the distribution of calling bowhead whales exposed to fluctuating anthropogenic sounds. *Journal of Cetacean Research and Management* 12:91–106.
- McMahon, C. R., H. R. Burton and M. N. Bester. 2000. Weaning mass and the future survival of juvenile southern elephant seals, *Mirounga leonina*, at Macquarie Island. *Antarctic Science* 12:149–153.
- McMahon, C. R., H. R. Burton and M. N. Bester. 2003. A demographic comparison of two elephant seal populations. *Journal of Animal Ecology* 72:61–74.
- Miller, P. J., M. P. Johnson, P. L. Tyack and E. A. Terray. 2004. Swimming gaits, passive drag and buoyancy of diving sperm whales *Physeter macrocephalus*. *Journal of Experimental Biology* 207:1953–1967.
- Miller, P. J. O., M. P. Johnson, P. T. Madsen, N. Biassoni, M. Quero and P. L. Tyack. 2009. Using at-sea experiments to study the effects of airguns on the foraging behavior of sperm whales in the Gulf of Mexico. *Deep Sea Research Part I: Oceanographic Research Papers* 56:1168–1181.
- Miller, P. J. O., M. Biuw, Y. Y. Watanabe, D. Thompson and M. A. Fedak. 2012. Sink fast and swim harder! Round-trip cost-of-transport for buoyant divers. *Journal of Experimental Biology* 215:3622–3630.
- Miller, P. J. O., P. H. Kvadsheim, F. P. A. Lam, *et al.* 2015. First indications that northern bottlenose whales are sensitive to behavioural disturbance from anthropogenic noise. *Royal Society Open Science* 2:140484. Available at <http://dx.doi.org/10.1098/rsos.140484>.
- Moore, J. L., and M. C. Runge. 2012. Combining structured decision making and value-of-information analyses to identify robust management strategies. *Conservation Biology* 26:810–820.
- Moretti, D., T. A. Marques, L. Thomas, *et al.* 2010. A dive counting density estimation method for Blainville's beaked whale (*Mesoplodon densirostris*) using a bottom-mounted hydrophone field as applied to a mid-frequency active (MFA) sonar operation. *Applied Acoustics* 71:1036–1042.
- Moretti, D., L. Thomas, T. Marques, *et al.* 2014. A risk function for behavioral disruption of Blainville's beaked whales (*Mesoplodon densirostris*) from mid-frequency active sonar. *PLOS ONE* 9(1):e85064.

- Murphy, D. D., and P. S. Weiland. 2011. The route to best science in implementation of the Endangered Species Act's consultation mandate: The benefits of structured effects analysis. *Environmental Management* 47:161–172.
- NRC (National Research Council). 2005. Marine mammal populations and ocean noise: Determining when noise causes biologically significant effects. The National Academies Press, Washington, DC.
- NRC (National Research Council). 2013. Assessing risks to endangered and threatened species from pesticides. The National Academies Press, Washington, DC.
- New, L. F., D. J. Moretti, S. K. Hooker, D. P. Costa and S. E. Simmons. 2013. Using energetic models to investigate the survival and reproduction of beaked whales (family Ziphiidae). *PLOS ONE* 8:e68725.
- New, L. F., J. S. Clark, D. P. Costa, *et al.* 2014. Using short-term measures of behavior to estimate long-term fitness of southern elephant seals. *Marine Ecology Progress Series* 496:99–108.
- Newman, K., S. T. Buckland, B. Morgan, *et al.* 2014. Modeling population dynamics: Model formulation, fitting and assessment using state-space methods. Springer, New York, NY.
- Nie, M. A., and C. A. Schultz. 2012. Decision-making triggers in adaptive management. *Conservation Biology* 26:1137–1144.
- Nolte, C., and A. Agrawal. 2013. Linking management effectiveness indicators to observed effects of protected areas on fire occurrence in the Amazon rainforest. *Conservation Biology* 27:155–165.
- Noon, B. R. 2002. Conceptual issues in monitoring ecological resources. Pages 27–71 *in* D. E. Busch and J. C. Trexler, eds. *Monitoring ecosystems: Interdisciplinary approaches for evaluating ecoregional initiatives*. Island Press, Washington, DC.
- Nousek-McGregor, A. E., C. A. Miller, M. J. Moore and D. P. Nowacek. 2014. Effects of body condition on buoyancy in endangered North Atlantic right whales. *Physiological and Biochemical Zoology* 87:160–171.
- Nowacek, D. P., M. P. Johnson, P. L. Tyack, K. A. Shorter, W. A. McLellan and D. A. Pabst. 2001. Buoyant balaenids: The ups and downs of buoyancy in right whales. *Proceedings of the Royal Society of London B* 268:1811–1816.
- Parks, S. E., C. W. Clark and P. L. Tyack. 2007. Short- and long-term changes in right whale calling behavior: The potential effects of noise on acoustic communication. *Journal of the Acoustical Society of America* 122:3725–3731.
- Pettis, H. M., R. M. Rolland, P. K. Hamilton, S. Brault, A. R. Knowlton and S. D. Kraus. 2004. Visual health assessment of North Atlantic right whales (*Eubalaena glacialis*) using photographs. *Canadian Journal of Zoology* 82:8–19.
- Pirotta, E., L. New, J. Harwood and D. Lusseau. 2014. Activities, motivations and disturbance: An agent-based model of bottlenose dolphin behavioral dynamics and interactions with tourism in Doubtful Sound, New Zealand. *Ecological Modelling* 282:44–58.
- Powell, J. R., and R. S. Wells. 2011. Recreational fishing depredation and associated behaviors involving common bottlenose dolphins (*Tursiops truncatus*) in Sarasota Bay, Florida. *Marine Mammal Science* 27:111–129.
- Pulliam, H. R. 1988. Sources, sinks, and population regulation. *American Naturalist* 132:652–661.
- Regan, T. J., B. L. Taylor, G. G. Thompson, J. F. Cochrane, K. Ralls, M. C. Runge and R. Merrick. 2013. Testing decision rules for categorizing species' extinction risk to help develop quantitative listing criteria for the U.S. Endangered Species Act. *Conservation Biology* 27:821–831.
- Richardson, W. J., C. R. Greene, C. I. Malme and D. H. Thomson. 1995. *Marine mammals and noise*. Academic Press, New York, NY.
- Rolland, R. M., S. E. Parks, K. E. Hunt, *et al.* 2012. Evidence that ship noise increases stress in right whales. *Proceedings of the Royal Society of London B* 279:2363–2368.

- Runge, M. C., S. J. Converse and J. E. Lyons. 2011. Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program. *Biological Conservation* 144:1214–1223.
- Samarra, F. I. P., and P. J. O. Miller. 2016. Identifying variations in baseline behavior of killer whales (*Orcinus orca*) to contextualize their responses to anthropogenic noise. Pages 963–968 in A. N. Popper and A. Hawkins, eds. *Effects of noise on aquatic life II*. Springer, New York, NY.
- Schick, R. S., L. F. New, L. Thomas, *et al.* 2013a. Estimating resource acquisition and at-sea body condition of a marine predator. *Journal of Animal Ecology* 82:1300–1315.
- Schick, R. S., S. D. Kraus, R. M. Rolland, *et al.* 2013b. Using hierarchical Bayes to understand movement, health, and survival in the endangered North Atlantic right whale. *PLOS ONE* 8(6):e64166.
- Schorr, G. S., E. A. Falcone, D. J. Moretti and R. D. Andrews. 2014. First long-term behavioral records from Cuvier's beaked whales (*Ziphius cavirostris*) reveal record-breaking dives. *PLOS ONE* 9(3):e92633.
- Schwacke, L. H., M. J. Twiner, S. De Guise, *et al.* 2010. Eosinophilia and biotoxin exposure in bottlenose dolphins (*Tursiops truncatus*) from a coastal area impacted by repeated mortality events. *Environmental Research* 110:548–555.
- Schwacke, L. H., E. S. Zolman, B. C. Balmer, *et al.* 2012. Anemia, hypothyroidism, and immune suppression associated with polychlorinated biphenyl exposure in bottlenose dolphins (*Tursiops truncatus*). *Proceedings of the Royal Society B* 279:48–57.
- Schwacke, L. H., C. R. Smith, F. I. Townsend, *et al.* 2014. Health of common bottlenose dolphins (*Tursiops truncatus*) in the Gulf of Mexico following the Deepwater Horizon oil spill. *Environmental Science and Technology* 48:93–103.
- Southall, B. L., A. E. Bowles, W. T. Ellison, *et al.* 2007. Marine mammal noise exposure criteria: Initial scientific recommendations. *Aquatic Mammals* 33:411–521.
- Stem, C., R. Margoluis, N. Salafsky and M. Brown. 2005. Monitoring and evaluation in conservation: A review of trends and approaches. *Conservation Biology* 19:295–309.
- Stow, C. A., S. S. Qian and J. K. Craig. 2005. Declining threshold for hypoxia in the Gulf of Mexico. *Environmental Science & Technology* 39:716–723.
- Taylor, B. L., and T. Gerrodette. 1993. The uses of statistical power in conservation biology: The vaquita and northern spotted owl. *Conservation Biology* 7:489–500.
- Taylor, B. L., M. Martinez, T. Gerrodette, J. Barlow and Y. N. Hrovat. 2007. Lessons from monitoring trends in abundance of marine mammals. *Marine Mammal Science* 23:157–175.
- Underwood, A. J. 1991. Beyond BACI: Experimental designs for detecting human environmental impacts on temporal variations in natural populations. *Australian Journal of Marine and Freshwater Research* 42:569–587.
- van der Hoop, J., M. Moore, A. Fahlman, *et al.* 2014. Behavioral impacts of disentanglement of a right whale under sedation and the energetic cost of entanglement. *Marine Mammal Science* 30:282–307.
- Watkinson, A. R., and W. J. Sutherland. 1995. Sources, sinks and pseudo-sinks. *Journal of Animal Ecology* 64:126–130.
- Wells, R. S. 2003. Dolphin social complexity: Lessons from long-term study and life history. Pages 32–56 in F. B. M. de Waal and P. L. Tyack, eds. *Animal social complexity: Intelligence, culture, and individualized societies*. Harvard University Press, Cambridge, MA.
- Wells, R. S. 2009. Learning from nature: Bottlenose dolphin care and husbandry. *Zoo Biology* 28:1–17.
- Wells, R. S., H. L. Rhinehart, L. J. Hansen, *et al.* 2004. Bottlenose dolphins as marine ecosystem sentinels: Developing a health monitoring system. *EcoHealth* 1:246–254.
- Wells, R. S., C. R. Smith, J. C. Sweeney, *et al.* 2014. Fetal survival of common bottlenose dolphins (*Tursiops truncatus*) in Sarasota Bay, Florida. *Aquatic Mammals* 40:252–259.

- Wilhere, G. F. 2012. Inadvertent advocacy. *Conservation Biology* 26:39–46.
- Williams, T. M., R. W. Davis, L. A. Fuiman, *et al.* 2000. Sink or swim: Strategies for cost-efficient diving by marine mammals. *Science* 288:133–136.

Received: 14 May 2015
Accepted: 10 January 2016