PAPER

Controlled-exposure Experiments to Determine the Effects of Noise on Marine Mammals

AUTHORS

Peter Tyack
Woods Hole Oceanographic Institution

Jonathan Gordon
David Thompson
Sea Mammal Research Unit,
Gatty Marine Laboratory,
University of St. Andrews

ABSTRACT

Controlled exposure experiments or CEEs are an important technique for determining the responses of animals to signals that are not part of their own communicative repertoire. CEEs are useful for establishing the relationship between acoustic dosage and behavioral response, a critical element of risk assessment, similar to dose:response studies for exposure to chemicals. CEEs share some properties with "playback" experiments; the main difference between playbacks and CEEs is that CEEs involve the careful titration of acoustic exposure to the point where specific responses are observed. Most CEEs are applied research designed to answer questions related to wildlife conservation. The utility and power of CEEs lies in providing a sensitive measure of causal relationships between behavioral responses and particular stimuli. We review design features and experimental methods for CEEs, limiting our scope for this paper to studying the effects of underwater noise on wild marine mammals.

INTRODUCTION

Controlled exposure experiments or CEEs have emerged as an important technique for determining the responses of animals to signals that are not part of their own communicative repertoire. CEEs share some properties with "playback" experiments: a technique that ethologists have long used to investigate animal behavior (especially communication in birds, amphibians and some land mammals). The primary focus of playback experiments involves a natural signal from the animal's own vocal repertoire, or perhaps the signal from an important predator or prey species. Many playback experiments also use synthetic signals as reference or comparison stimuli. Much of the methodological experience of research groups that have conducted playback experiments (e.g. McGregor, 1992) can be applied to CEEs.

The definition we propose to use for CEEs in this paper is:

> *A field procedure in which controlled doses of an acoustic stimulus are applied to focal animals for the purposes of assessing their behavioral and/or physiological responses. The stimulus, which may either be generated by the noise producing object itself or reproduced electronically from recordings using an underwater sound projector, will be under the control of the experimenter.*

An important difference between playbacks and CEEs is that CEEs involve the careful titration of acoustic exposure to the point where specific responses are observed. In order for a playback experiment to work, the sound that is broadcast must be loud enough for the subject to hear it, but few playback experiments measure or estimate the received level at the animal. Part of the art of playback experiments involves adjusting acoustic parameters to obtain responses; in CEEs the emphasis is on evaluating the relationship between acoustic dosage and behavioral response. This means that CEEs are likely to require more explicit acoustic measurements and analyses than playbacks.

We will limit our scope for this paper to studying the effects of underwater noise on wild marine mammals.

What Questions Can CEEs be used to Answer?

While most playback experiments are designed to answer basic questions about animal communication, CEEs usually have an applied research goal—understanding the effects of anthropogenic noise on wild animals. Such investigations should be driven by specific management objectives, which will change for different circumstances and regions. The utility and power of CEEs lies in providing a sensitive measure of causal relationships between behavioral responses and particular stimuli. They are principally useful for investigating short-term behavioral and physiological responses to sounds, although the duration and spatial extent of CEEs can be increased with appropriate observational techniques. They should often be more powerful for investigating responses than non-experimental approaches as they allow control over many factors such as age, sex, individual experience, location, time, and season.

Identifying Management Objectives

Much legislation and regulation concerning the impact of human activities on marine mammals is couched in terms of conservation, usually the maintenance of viable populations and functioning ecosystems. Animal welfare considerations are often a matter for public concern and may also be a focus of legislation in some countries. Different regulatory concerns may require studies of impact to focus on different time and spatial scales.

The U.S. Marine Mammal Protection Act (MMPA) is based upon the goal of protecting and conserving marine mammal populations and their habitats; section 2.1.6 states that "the primary objective of their management should be to maintain the health and stability of the marine ecosystem." For several international treaties such as ACCOBAMS (Agreement on the Conservation of Cetaceans of the Black Sea, Mediterranean Sea and Contiguous Atlantic Area), it is adverse impacts on cetaceans that trigger stricter standards for regulation. The conservation plan of the ACCOBAMS treaty specifies the need to "regulate the discharge at sea of, and adopt within the framework of other appropriate legal instruments stricter standards for, pollutants believed to have adverse effects on cetaceans." This special status of pollutants believed to have adverse effects highlights the critical element of effects studies. Several national laws involve similar efforts to prevent adverse impacts to marine mammal populations. This regulatory perspective suggests a focus on whole populations over periods of several generations.

Other regulatory considerations may also suggest the need to investigate at smaller time and spatial scales. The core of the U.S. MMPA is a moratorium on the taking and importing of marine mammals, where "take" means "to harass, hunt, capture, or kill … any marine mammal." Harassment has been defined by amendments to the MMPA made in 1994, and includes two levels: a level A involving a potential to injure a marine mammal, as well as a level B defined as "potential to disturb a marine mammal or marine mammal stock in the wild by causing disruption of behavioral patterns." Under the MMPA, human activities that may disrupt the behavior of marine mammals require authorization for harassment "takes" independent of the potential impact on marine mammal populations. This requirement puts a priority on research that can specify the conditions under which there is a potential for harassment takes. The focus on harassment suggests studies of how individual animals respond to specific stimuli over time scales of seconds to hours.

The U.S. MMPA contains special provisions not only to protect marine mammals, but also to protect human uses of marine mammals such as subsistence whaling. Both consumptive (e.g. whaling) and non-consumptive (e.g. whale watching) use of marine mammals could be affected if the distribution of animals was changed or if individuals became harder to approach as a result of noise exposure. For example, the effects of oil industry activities on bowhead whales offshore of the North Slope of Alaska have been studied because of concern that native hunters might have to go farther offshore to hunt bowheads. The critical issue here is not only whether oil industry activities may themselves have an adverse impact on marine mammals, but also whether they change the distribution of whales in a way that makes it more difficult and dangerous for native hunters. This concern suggests studying how oil industry activities may affect the distribution of whales over scales of tens of kilometers and durations of days or more.

Some activities may operate under restrictions to have the least practicable impact on wild animals. If this is taken to include reducing harassment takes, then this requirement may call for sensitive behavioral studies. A wide range of parties, from whale watchers, consumers of a resource, producers of industrial noise, to regulators and environmental NGOs, subscribe to a perspective of environmental responsibility, often beyond regulatory requirements. If this is taken to demand the avoidance of unnecessary disturbance and the identification and adoption of less disruptive options wherever possible, then these goals also call for sensitive behavioral studies of impacts of activities on marine mammals.

Relevance of small-scale studies to larger scales

Most CEEs conducted to date have involved the monitoring of relatively short-term responses to short-term exposures of sound (several hours in both cases). Especially if the experiment involves following a specific individual, there is usually a tradeoff between the duration of each experiment and how many experiments can be conducted. In addition, many of the current methods for following specific animals, whether involving tagging or visual and/or acoustic monitoring, are not capable of following animals for much more than several days and are seldom suitable for investigating responses to long-term exposures. Satellite tags can follow animals for longer periods of time, but seldom provide sufficient detail about behavioral responses for CEEs. However, several CEEs have been designed to study changes in the distribution of animals over periods of several days of transmission. For example, the ATOC Marine Mammal Research Program operated a stationary sound source for oceanographic research on a schedule designed to evaluate shifts in distribution of whales over periods of several days. Aerial surveys were conducted after 1-3 days of transmission for 20 min every 4 hours and were compared to control distributions after at least 4 days when the source was silent (NRC, 2000).

It can be argued that, when conservation considerations are pre-eminent, population level effects are of most concern and short-term responses may have little impact on population processes. In which case, improving understanding of short-term reactions may not seem to be a key knowledge requirement. This flags a fundamental dilemma, which is worth exploring here. Given the difficulty of defining marine mammal populations and assessing abundance, population level effects will be very difficult to detect unless they are very dramatic (Taylor et al., 2000). If one waits until the adverse impacts of a human activity become obvious at the population level, it may be too late to reverse the effects.

Furthermore, any natural population is subject to so many different natural and manmade effects that proving cause and effect, i.e. showing that a population decline was really caused by a noise source and not by some other factor, will be extremely difficult. Certainly, any team of scientists motivated to argue a contrary position would have a wealth of alternative hypotheses to work with. In several parts of the world, marine mammals may have left areas where increasing levels of industrial development degraded their habitat (e.g. Bryant et al., 1984). However, it has proved difficult to integrate studies conducted over long time periods using different methods, and has proven nearly impossible to demonstrate causation between industrial development and changes in distribution (e.g. Richardson et al., 1987), especially when observations are not continued after a decline in industrial activities. It may also be the case that during a period when populations were increasing for other reasons, the deleterious effects of a human activity, which could be significant when conditions were less favorable, might be obscured.

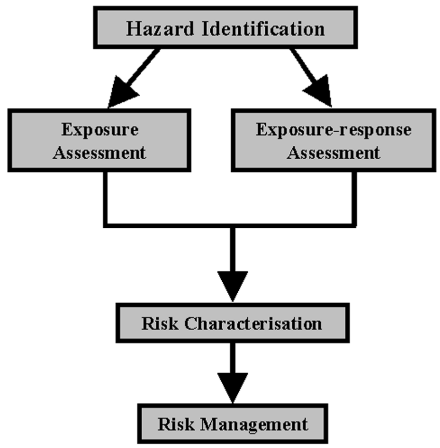
Most fundamentally though, population level responses in marine mammals are long-term effects that, by definition, will take a long time to detect, and cannot be observed until they have occurred. It is neither good management, nor ethically defensible, to allow population level effects to occur before identifying and addressing potential problems. From this perspective, when biologists plan CEEs to study short-term responses of individuals, it is important that they design the short-term study in ways that allow the assessment of longer-term impacts to the individual and cumulative impacts to all of the individuals in a population. Generally, it will be most appropriate to use a suite of complementary research approaches including short-term investigations of individual response, studies involving intermediate spatial and temporal scales, and longer-term monitoring of populations and the activities that may affect them.

Identifying the Key Knowledge Gaps: the place of CEEs in the context of risk assessment

We know very little about the effects of noise on marine mammals. An approach that has proved very useful for protecting humans from environmental risks in the face of uncertainty involves risk assessment coupled with decision analysis. Harwood (2000) encourages the use of a risk assessment framework for wildlife conservation management and provides examples of how this might be applied to reducing the by-catch of small cetaceans. Figure 1 illustrates the risk assessment framework developed for assessing the effects of chemical pollutants on human health by the US Environmental Protection Agency (NRC, 1983, 1994a; EPA, 1992).

The first stage of the risk assessment developed by NRC (1983, 1994a) and EPA (1992) involves identifying a hazard, such as epidemiological data suggesting a link between exposure to a pollutant and health problems. The next stage involves two steps that can be taken in parallel: (1) determining the probability of exposure in the population and (2) experiments to define the dose:response relationship between exposure and the hazard. Once these steps have been completed, it is possible to calculate the total effect of the summed exposure to the hazard for the population. A final stage involves comparing the benefits of different strategies to manage the risk. Harwood (2000) points out that the initial NRC (1983) report characterized the risk in terms of probability of death or serious injury to individuals at various distances from the source. This kind of characterization may be useful for estimating the number of "takes" expected under the MMPA, as is required for incidental take authorizations. EPA (1992) broadened the goal of risk characterization to effects on populations. This is closer to the basic objective of most conservation legislation, and this broader goal will often be appropriate for evaluating effects of noise on marine mammals.

FIGURE 1
Schematic of risk assessment procedures suggested by EPA (1992).



A key feature of risk assessment is the development of models that include an assessment of the consequences of uncertainty about the various parameters, about the models used, and about the effectiveness of different management approaches. Harwood (2000) advocates Bayesian decision analysis for this problem with an iterative approach to estimating prior distributions. He specifically advocates that Bayesian analyses may offer a scientific basis for employing the precautionary principle, which forms the basis of many laws relevant to the problems discussed here. When following a precautionary approach, conservative (precautionary) values should be incorporated when uncertainty exists. However, these pessimistic values can be challenged with empirical data and if appropriate, revised. The risk assessment process will help to identify those parameters that contribute most uncertainty, and direct research accordingly. The expense, difficulty, and long time-frame required for much of this research strongly argues for careful risk assessment to direct research efforts to most effectively reduce uncertainty. By an iterative process, risk assessment models can be repeatedly challenged by new data and refined to reassess policy. One attractive aspect of this framework is that noise producers may be motivated to fund research to reduce uncertainty around parameters that they consider overly pessimistic.

Harwood proposed in Gordon et al. (2003) that the following steps could be followed in considering a new threat to wildlife.

- Assemble stakeholders to agree on available information, aims of management, acceptable levels of risk
- Identify potential management scenarios
- Develop models of processes and challenge them with data
- Incorporate uncertainty
- Assess performance of different management procedures

Harwood (2000) suggests that conservation biologists already successfully use many of the tools he advocates. These tools are seldom called risk assessment, but are used in population viability analyses, which help identify situations where populations are at risk of extinction. If the approach is accepted for problems as drastic as extinction, it would be surprising if it could not be applied to risk of harassment. Harwood (2000) points to the problem of fisheries by-catch of harbor porpoise as a successful example of risk assessment. The hazard was identified by the combination of information on the by-catch of porpoise in fishing nets with suggestions that the populations might be declining. Intensive monitoring programs by the U.S. NMFS estimated the rate of by-catch, along with population status. Caswell et al. (1998) incorporated uncertainty in porpoise demography to characterize the risk of population decline. Harwood (1999) took a risk management approach in evaluating the effectiveness of two different management strategies for solving this problem.

The technical and quantitative approach advocated by Harwood (2000) is not common in conservation biology. Several recent controversies suggest the importance of all parties involved agreeing up front on the approach to be used. For example, the development of the research project to study effects of the SURTASS LFA sonar involved a similar process to that listed above (Tyack 1998a). Stakeholders met many times, but unfortunately did not agree on the definitions of zones for different levels of risk. In the end, these conflicts were addressed by litigation and legal judgments. When political and legal processes get involved in selecting management options, the end result is often very different from what most conservation scientists would judge to be best for the populations involved. The jury is out on whether decisions are better made following the risk assessment approach. At the very least, disagreements could perhaps have been better addressed initially through the process of developing and testing models of risk; at worst, such a process would encourage clear thinking and help identify research priorities.

Whether or not a formal risk assessment approach is used to make management decisions, the structure of risk assessment illustrated in Figure 1 highlights the role of Controlled Exposure Experiments. CEEs are used to assess the relationship between acoustic exposure and behavioral response, as part of the exposure-response assessment. Dose: response studies are familiar tests of the effects of chemicals on the physiology and biochemistry that underlie health in animals and humans. Similar lab studies may be used to study the effects of sound on physical injury or hearing in animals. For example, just as toxicological studies aim to find the dosage that kills 50% of the subjects, studies on the effects of underwater explosions have defined exposure that yield 50% probabilities that animals will be killed from the blast (Yelverton, 1981; as described in Richardson et al., 1995). Injury has been studied in similar ways; Ketten (1995) reviews studies that estimated exposures that yield 50 % probabilities that the eardrum will be ruptured. Just as lethal studies designed to protect humans use other species as test subjects, so these blast studies submerged terrestrial mammals such as rats and sheep to test the impact of underwater blasts. The use of surrogates always poses problems; these are highlighted by the use of eardrum rupture as an indicator of injury since the external ear canal is not even attached to the eardrum in cetaceans.

Few marine mammals are held in lab settings. Ethical and practical considerations prevent most if not all researchers from considering studies in which their animals might be injured. During the past 10 years, several laboratories have developed techniques to study what sound exposures cause harmless temporary changes in hearing called temporary threshold shifts (Kastak et al, 1999; Schlundt et al., 2000). These experiments help to define for marine mammals the lowest exposures that cause physiological changes which mark an early signpost for risk of injury. Since the auditory system functions in similar ways in pools or pens as in the field, there is little reason for concern that research with trained animals on the physiology of hearing is not valid for the same species in the wild. Field CEEs (as defined here) are not an appropriate way of investigating physical injury or hearing damage; in addition to ethical considerations, it would prove very difficult to expose wild animals to sufficiently intense sounds for prolonged periods and to subsequently assess any damage caused. (However, playback experiments could be used to assess pre-existing hearing impairment in individual animals by studying the received levels or signal to noise ratios at which animals react to exposure to a meaningful sound.)

Most classic dose:response studies involve lab animals in carefully controlled settings. Given that the results are usually extrapolated to an entirely different species—humans—other differences between the lab animals in factors such as health, diet, etc. are usually viewed as second order effects. CEEs are used to study the relationships between an acoustic stimulus and the behavioral response it evokes. Behavior is usually highly context dependent. This raises concern about ecological validity: are the experiments conducted in settings close enough to those for which the results will be applied? Some experiments such as Schlundt et al. (2000) not only measured temporary threshold shifts, but also noted exposure levels at which alterations in behavior were observed. In general however, issues of ecological validity suggest that CEEs should be conducted in the field if the results are designed to regulate exposure of wild animals.

Advantages and Shortcomings of a Controlled Exposure Experimental Approach

A question posed in several workshops on the appropriate use of CEEs (e.g. Gordon et al., 2003) was: "When should one use CEE versus other research approaches?" The short answer is that CEEs should be considered when they promise to offer the most effective methodology for providing the key information that is required for effective management. Without prejudging the results of any risk assessment, we can safely assume that effective management will require information on the hearing abilities of marine mammals, the potential for physical and auditory damage, short-term behavioral effects, the potential for masking and long-term consequences for populations. CEEs will generally only be useful for answering a subset of these questions. In evaluating whether to conduct a CEE, a cost-benefit analysis should be conducted comparing any potential risk to the subjects to the benefits of the results to the population. CEEs will entail some additional acoustic disturbance, and therefore CEEs designed as purely applied research should only be considered where research is clearly focused on providing information needed for management.

Experiments Versus Correlation Studies

There has been considerable discussion among biologists concerned about the effects of noise concerning whether experimental approaches or correlational studies (looking for responses that were correlated with noise exposures that were not controlled by the research team) are more appropriate for investigating effects of anthropogenic noise. The most important point to make here is that these are not mutually exclusive. The strengths of each approach often correct for the weaknesses of the other. For example, many factors lead CEEs to examine short-term responses to short-term exposures. Where existing sources transmit for long periods of time, correlational studies often must study long-term responses to long-term exposures. In most cases, research involving both approaches will be required. In addition, careful modelling studies may also sharpen the focus on the most critical uncertainties, if not resolve them. Further, there is no one right answer, the effectiveness of each approach will depend on the specific circumstances, including the species being investigated, the noise source and the ease with which it can be reproduced, and the research location.

Some general points to consider when decisions have to be made between these two include:

- Experiments represent a particularly powerful method for establishing cause and effect relationships.
- CEEs may be more expensive to conduct, but if they are more powerful in a statistical sense, less research will be required and the answer may be found more quickly. How these costs and benefits balance will vary from case to case.
- When knowledge of responses to a new sound source is required before that source comes into general operation (in the cases of a new sonar or experimental oceanographic sound source for example), then experiments are the only option.
- CEEs may also allow investigation of responses in naive animals (animals that haven't experienced the sound before) and may allow work to be conducted in an area where the sound source does not operate normally, but in which research conditions are particularly favorable, e.g. access to animals, good weather, a known and studied population.
- Correlational studies examine effects of existing noise pollution and do not require additional noise inputs into the marine environment. Controlled exposure experiments will always involve the production of extra noise in the environment. The extent to which this is an issue will depend on the power of the sound source being used, the duration of the experiment, the sensitivity of the animals exposed and their population status. The use of a sound source in experimental conditions in which transmission will be controlled, the responses of animals will be monitored and the experiment can be halted if any deleterious responses are observed should carry fewer attendant risks than "normal" uncontrolled exposures. Ultimately, any risks must be weighed against the advantages of speedily gaining an understanding of a potential environmental threat. Clearly, suitable risk minimization measures must be rigorously applied.
- The costs and logistical difficulty of broadcasting underwater sound for CEEs will vary depending on source type. This can become an important consideration where it is necessary to re produce very powerful signals such as those from seismic air gun arrays or military sonar. These may be very expensive and logistically difficult to produce, favoring studies other than experimental ones. Another approach involves modelling of sound source characteristics, acoustic propagation, physiology, audition, and vocalizations. Such models may need to estimate poorly known parameters, and the modelling exercise can help determine the most critical un knowns. For example Erbe (1997) has used this kind of modelling to estimate how vessel noise could mask communication in beluga whales, and she then tested masking with captive belugas to test her predictions.

Experimental Design and Methods: how can CEEs be most effectively carried out

Design

CEEs have many features in common with ethological playback experiments and much can be learned and applied from this better developed field. A NATO Advanced Study Institute on the design of playback experiments was held in 1992. This addressed issues of design and methodology that are relevant to CEE. The institute highlighted the four following important considerations:

1. The selection of subjects and stimuli for experiments should be appropriate to the hypothesis being tested.

A particular criticism of early playback efforts was that in some cases an insufficient number and diversity of signal types and experimental subjects had been tested to allow general inferences to be made about the responses of members of a population or species to a call type. CEEs focused on the impact of a particular sound may not have the same concern about diversity of sound types. However, if broad statements are to be made about the effects of general noise types on members of a species then a sufficient number of representative experimental stimuli must be tested. The concern about appropriate diversity of subjects applies equally to CEEs as to playbacks. These should represent an adequate sample of animals from both sexes covering the full range of ages and motivational states. There are often practical limitations to which species can be tested, and to the number of subjects. If only a limited range of subjects can be tested, then this should include animals thought to be most sensitive to acoustic disruption.

2. Experiments should exhibit "external reference" or "ecological validity."

This point refers to the extent to which the experimental procedures reflect situations that are typical of the real world. We have already discussed how this argues for most CEEs to be conducted with a representative sample of wild animals in their natural environment. Similarly, noise exposures should reflect the full qualitative and quantitative range of sounds likely to be experienced. The critical goal is that experiments evaluate responses of animals that experience sound fields that are similar to the actual sources as used at sea.

3. Experiments should be designed to have biological relevance and test biologically significant responses.

In the context of CEEs the issue of biological relevance relates directly to the management goal of the research. These may change depending on the interpretation of regulations. For example, research designed to evaluate what exposures are required to lead to harassment takes under the MMPA need evaluate the threshold for "disruption of behavior patterns." This has in the past been interpreted to involve any detectable change in behavior (Swartz and Hofman, 1991). However, two committees of the National Academy of Sciences (NRC, 1994b, 2000) have argued for a distinction between detectable changes in behavior and disruption of behavior that may have an adverse impact. NRC (2000) advocated changing the definition of harassment to meet a criterion of "disruption of biologically significant activities" and the definition of harassment has recently been changed for military activities (Kaiser, 2003). Just as EPA (1992) broadened the NRC (1983) view of risk assessment from estimating numbers of individual "takes" to estimating impact on whole populations, so this change in the new definition of harassment requires extra work to relate changes in behavior to potential effects on growth, reproduction, and survival. Where the management objective involves maintaining the health of populations, the criterion for "biological relevance" may require going beyond the risk characterization for individuals to estimate the overall expected impact on the population. While this broader issue involves all of the steps delineated in Figure 1, it specifically influences the design of CEEs. It is a challenging scientific problem to attempt to relate short-term behavioral disturbance to effects on demography, for the models and data for these two areas differ in scale by orders of magnitude. One promising area for linking the two involves energetics. If a stimulus reliably disrupts foraging behavior, one can evaluated the disruption in terms of reduced foraging rate and potentially increased energetic cost. Several studies have been able to relate variation in the economics of foraging with reproduction in female mammals.

4. Experiments should be designed to avoid bias from observer effects.

There are many subtle ways in which an observer/experimenter's knowledge of the exposure at any particular time could unwittingly influence both the way that data are collected and scored and the way that experimental protocols are applied. For this reason it is desirable that procedures such as CEEs should be conducted "double blind" with those conducting the experiment and those scoring behavioral responses in the data collection team being unaware of the experimental treatment. There are a variety of obstacles to double blind procedures in CEEs. Some ingenuity may be required to achieve a double blind protocol on a small boat, especially when it is important to monitor hydrophones for other reasons, for example to monitor vocal responses. A double blind procedure also requires that the observer cannot determine when an exposure begins—this is to avoid inadvertent cueing to particular behaviors. While the reasons for requiring this are clear, it may not always be desirable in a CEE. It may often be necessary to start an exposure at a particular time in a behavioral cycle, when an animal is at the surface and within a certain range for example—or when other conditions are favorable for collecting a certain type of data. A compromise may be for the experimenter to decide approximately when the exposure should occur (based on a predetermined set of conditions) but for there to be a random time delay before its onset so that inadvertent cuing of playback on significant behaviors could be avoided. In effect, the experimenter does not have control of the broadcast start time at the temporal scale at which responses will be analysed. It may also be necessary for some of the team to know playback conditions in real time so that experiments can be halted or modified if adverse responses are observed. In this case, ideally, the mitigation monitoring team would be isolated from the experimental monitoring team.

Double blind procedures are particularly important wherever there is a subjective component to the recording method, or where the observation process itself may influence the behavior. Where studies use telemetry data or where automated acoustic detection programs are used to quantify vocal responses, observer bias or influence during the behavioral categorization stage is unlikely to be an issue. Where human scoring of behavior is necessary, but where it is difficult to keep observers blind while the experiment is underway, there are other solutions to ensure a double blind design. Behavioral data recorded on video and/or audiotape can be analysed independently by researchers with no knowledge of exposure conditions. (Making a backup record of CEEs in this way is good practice anyway.) Where underwater sound recordings need to be analysed, for example, to assess changes in vocal behavior, the experimental condition will often be obvious to the analyst because the playback sound will be audible on the recording. A potential solution here will be to mix samples of playback sound on top of recordings for both the exposure and non-exposure recordings before they are analysed.

Additional Methodological Considerations

Choosing behavioral and physiological parameters to measure responses

The importance of having an in-depth appreciation of the biology of the study animals and a good knowledge of the particular population being observed cannot be overemphasized. Where possible and acceptable, it is advantageous to conduct research with populations that have been subject to long-term studies and work in close collaboration with those that carried out that research. This can often create opportunities to interpret short-term responses from a perspective that includes knowledge of naturally occurring variation in behavior. For example, if a CEE causes animals to move away from a habitat they have consistently inhabited for decades, this response would be judged more significant than if the animals would have been likely to have left under control conditions anyway, or even if prior distribution and movement patterns were unknown.

As far as possible, researchers should measure effects on behaviors whose functions are well understood. If the experiment is designed for conservation research, behavioral parameters should be selected so that disruption of the behavior can be linked to long-term consequences for survival, growth or reproduction. Other relevant considerations are that statistical power will be higher when a behavior that is predictable and has low inherent variability is measured and that measured behaviors should be those that can be easily and reliably observed using methods which are not susceptible to bias.

The most appropriate experimental design to apply at any particular time may depend on the animal's behavioral state. For example, a protocol that can be easily applied when animals are travelling may not be appropriate when animals are socializing. The experimenter may thus require several alternative protocols, which may involve measuring different behavioral cues, with the animal's behavioral state determining which is applied.

Prioritizing species and individuals as focal animals for exposure experiments

Ideally a research project would study all species inhabiting the affected habitats in all of the relevant behavioral states, but this will rarely be feasible. A priority should then be to study a representative selection, and within this prioritize:

- Species or populations that are believed, on the basis of their behavior, auditory characteristics or other biological features, to be particularly sensitive.
- Species or populations with poor conservation status (though consider using a surrogate species where possible if it is thought to be equally acoustically sensitive).
- Species or populations that on the basis of their behavior or distribution are likely to be most heavily exposed to sound being investigated.

All three of these factors affect the vulnerability of a population. Vulnerability of individuals is a combination of their sensitivity and exposure while the threat to the population is also a function of its conservation status. For some of the most endangered species, even the risk of the experiment may be of concern and the use of a proxy might be considered. For example, the southern right whale might be a good substitute in experiments for the biologically similar, but much more endangered, North Atlantic or North Pacific right whales.

Within species, studies should be conducted preferentially on populations for which long-term data are already available. Within populations, it would be ideal to study both sexes and all age classes and behavioral states. Given that there are always limitations on meeting this goal, those classes thought likely to be most vulnerable to short-term and long-term disruption (e.g. calves, mothers, breeding animals) should be given highest priority. At first sight this may not appear to be a very precautionary course of action, but research which does not measure effects on the most vulnerable individuals may seriously underestimate the degree to which they may be affected. To the extent that results of CEEs are used to regulate noise exposure, the precautionary course is to select the most sensitive animals for study.

Sources of variation in individual response

It is important to recognize that several factors may affect how any particular individual responds when exposed to a noise. These include:

Auditory sensitivity

There is a great range in auditory sensitivity of different animal species. The auditory systems of most fish have relatively low sensitivity to sound, which declines rapidly above several hundred Hz (Tyack 1998b). Some fish species have specialized auditory systems that are more sensitive, and capable of detecting higher frequency sounds. Recent data suggest that a few taxa of fish are capable of detecting intense high frequency pulses in the 25-150 kHz range (Mann et al., 1998). The inner ear of mammals is tuned to specific frequency ranges, and different taxa may be sensitive to different ranges. For example, delphinid toothed whales have ears specialized for high frequencies, with best hearing in the 50-80 kHz range, while the ears of baleen whales are specialized for low frequency hearing below several kHz (Tyack 1998b). These differences in auditory sensitivity are critical for determining which taxa are likely to be affected by manmade sound sources that operate in specific frequency bands.

The auditory system is sensitive to illness, injury, and exposure to loud sounds. Few individuals of one species have "normal" hearing if this is defined as the best hearing possible for members of the species. For this reason there is a great deal of variability in the auditory sensitivity of different individuals of the same species. Animals with specific injuries may have impaired hearing at particular frequencies. As animals, especially males, age, they have more diffuse hearing loss, especially in higher frequencies. All of these differences in auditory sensitivity will affect how an animal responds to noise.

Noise exposure histories and experience of subjects

A basic feature of animal learning is that animals may cease responding or habituate to stimuli that are not associated with any reinforcement, while they may become very responsive (sensitized) to stimuli that are associated with either positive or negative reinforcements. These learning processes may modulate the way an animal responds to a novel stimulus vs. after it has heard it many times. Habituation and sensitization may occur as the result of multiple exposures of individuals to a particular stimulus. If a CEE studies responses of animals to a new sound stimulus that they have not heard before, the results may differ from how the animals would respond after repeated exposure. For example, Cox et al. (2001) found that harbor porpoises avoided a pinger, but that the avoidance distance reduced by half in just four days. Where possible, experimenters should keep track of the exposure history of individuals and test for habituation/sensitization effects. This is only possible for sound sources that have not been introduced to the study animals except under the experimenter's control. Even for these sound sources, few studies have tracked changes in responsiveness of individual animals over repeated exposures. For sound sources that have already been introduced into the ocean, different results may be obtained with CEEs conducted with more or less naive populations. The expected experience of animals or populations depends upon the spatial and temporal patterns of previous and ongoing noise exposure, which is often difficult to quantify. Experimenters should be careful to take these effects into account as they design and interpret their CEE experiments. For example, animals would be expected to habituate to stimuli that are unlikely to be associated with reinforcement, so CEEs studying the extent of behavioral disruption to initial exposures are likely to be a conservative first approximation. However, studies of how effective pingers are in keeping porpoises away from nets may exaggerate the effectiveness if they only focus on avoidance responses to initial exposures (Cox et al., 2001).

Behavior and context

CEEs conducted in different contexts (e.g. in offshore/inshore (Buck and Tyack, 2000), confined/exposed locations (Richardson et al., 1995, pp 284-286), in the presence/absence of predators, with subjects in social groups/alone) may produce very different responses. Studies that test such variability may identify situations that result in heightened sensitivity. It is recommended that experiments should be repeated with animals in a range of different behavioral contexts. Unfortunately, it is seldom possible to study all contexts. Often knowledge about the behavioral ecology of the study species may help identify the contexts most relevant to the issue the experiment is designed to clarify. In general, it will usually be important to select the context in which the animals are most likely to show the greatest disruption of behavior. From a conservation perspective it may be most appropriate to investigate the consequences of disruption at times in the life cycle when this might have the most significant biological effect, when energy budgets are low for example.

Variation in Received Signals

Some manmade sounds (e.g. sonar) are very predictable at the source, indeed some are specifically designed to be so. Others, such as shipping or drilling, may be more variable. In addition, environmental conditions, such as water depth, animal depth, and propagation conditions will all affect the characteristics of received sound. For example, impulses from air guns have fast rise times and durations typically measured in tens of msec near the source, but frequency-dependent multipath may modify air gun impulses at ranges of tens of km to frequency modulated chirps with apparent durations of hundreds of msec (Greene and Richardson, 1988). Accounting for such variability requires experiments that cover the whole range of signals. However, this generality can only be achieved by reducing the sample size within each exposure type and therefore sacrificing statistical power.

Conclusion

This section mentions a variety of sources in the variability of behavioral responses of individual animals. Because of this variability, the most sensitive tests for responses will involve experimental designs that compare matched observations of experimental exposures and control observations of the same individuals, occurring close in time and under similar conditions. For example, the duration of the songs of humpback whales changes over time (Payne et al., 1983) and even at one time the songs of different individuals vary. By comparing the duration of songs from the same individual whales before, during, and after exposure to a low-frequency sonar, Miller et al. (2000) were able to show a significant difference in data from just 6 individuals. If data were pooled without knowledge of which whale produced which song, the test would have had lower power.

Controlling for the effects of the observation and playback vessel

When experiments are conducted in the open ocean, both a playback vessel (carrying a sound source) and an observation vessel may be involved. Potentially, either or both of these platforms on their own could have an effect on the subjects. Clearly, it is important to choose vessels that have a minimal impact; generally these boats should be as quiet as possible and keep as far away from the subjects as is consistent with the collection of good data. In addition, their effects should be controlled for by having them present, and behaving in exactly the same way, during both experimental and control conditions. Employing double-blind procedures will help in achieving this. In some circumstances it may be necessary to test whether the presence of the vessels may influence the responses of the subjects to CEE stimuli, but this usually would be a secondary effect.

Measuring Masking with CEE

One of the potentially deleterious effects of noise in the marine environment is the masking of biologically significant signals. The first warning in the literature about effects of noise on marine mammals concerned the possibility that shipping noise might mask the vocalizations of baleen whales, drastically reducing their effective range (Payne and Webb, 1971). On the other hand, animals might be able to modify their vocalizations in order to maintain the ability to communicate over the typical separation of source and receiver. The best known such compensation mechanism involves increasing the source level of the vocalization to compensate for the increased noise (known as the Lombard effect in humans). Playbacks of conspecific songs show that birds respond by increasing the source level of their own songs (Brumm & Todt, 2004), but in the more than three decades since alarm was sounded over this issue for marine mammals, we are not aware of any studies on changes in source level of vocalizations as a response to increased noise. However, other compensation mechanisms have been studied. For example, Au et al. (1985) report that when a beluga whale was moved to a location with raised levels of background noise in a specific frequency band, the beluga changed the frequencies of its echolocation signals to avoid the noise. Miller et al. (2000) suggest that humpback whales increase the redundancy of their song in order to compensate for noise from a low frequency sonar.

Field playbacks of both signals and masking noise could be used to investigate these processes. One indication that masking may be occurring would be changes in the characteristics of vocalizations, used for echolocation or communication, that seem to be adaptive in reducing their vulnerability to masking. During a CEE experiment, a potentially masking noise could be broadcast to investigate whether animals adapted their acoustic signals in response to noise exposure. These modifications might include either increasing the source level of the vocalization, increasing the redundancy of the call, or shifting the signal outside of the dominant frequencies of the noise. Another approach would involve the playback of both signals that elicited a predictable response (these might be the calls of predators or, for breeding males, vocalizations of receptive females – Parks, 2003) and noise (Frankel et al., 1995). The received levels at which responses occurred for different levels of noise could then be investigated (e.g. Lesage et al., 1999). Noise could either be generated experimentally or CEEs could be conducted in different normally occurring noise fields.

Equipment and Practical Considerations

UW Sound Sources

Underwater sound production equipment can be difficult to obtain. Some sources that have used for CEEs are summarized in Gordon et al. (2003). The ease with which a particular sound can be produced during a CEE is an important consideration. For example, it is easy to produce simple high frequency signals such as those from "pingers" designed to reduce fisheries bycatch, but very large sound transmission equipment may be required to broadcast powerful low frequency signals. The only way to accurately reproduce the sound of a supertanker or a full-scale air gun array is to operate the object itself, often a logistically difficult and extremely expensive proposition. There is no single solution to the tradeoffs between using simple approximations of actual sources and using the actual sources. The experimenter must evaluate the importance of the differences between approximation and reality, and the practical issues involved in arranging to use the actual source.

Playing back recordings from an underwater loudspeaker can offer flexibility, but powerful sound projectors capable of reproducing low frequencies are specialized and expensive pieces of equipment. In any experiments involving electronic reproduction of sound transmitted underwater, it is important to ensure that only high quality recordings are used and that other cues, such as recording artifacts, which would not be present with the real sound source, are not broadcast. Underwater broadcasts should be monitored with calibrated hydrophone equipment to be sure that the intended signals are actually being broadcast without distortion or interference from other sounds. Experiments should ideally be designed with a control condition in which a "blank" recording is transmitted from the source rather than the experimental stimulus.

Other characteristics of the sound source—for example, whether or not it is moving, how it is oriented and moves with respect to the subject (approaching or receding)—may also be important considerations during a controlled exposure experiment. These characteristics should be carefully considered in the experimental design.

How powerful a source should be used?

A sound source should be loud enough to be capable of exposing an animal to sounds that span the range of received levels being investigated. If regulations set an acceptable upper limit of exposure then it will be important to test responses to exposures up to this level to have confidence that adverse responses do not occur. In practice, this may dictate using large and powerful sources. Smaller sound sources at close range may not adequately mimic the effects of larger sources at greater ranges because cues to range (such as reverberation) will not be present. Another problem with weak sources is that received levels will tend to drop off more rapidly as an animal moves away from a small source close by. By the same token, at short ranges, error in estimates of the length of the transmission path will result in greater errors in estimates of received levels. Unless one knows the precise location of the subject at all times, this makes it more difficult to predict the transmission loss to adjust the source level to expose the subject to a specific received level.

Some of the more powerful modern sources involve arrays of individual sources operated in sophisticated ways to make the source directional. Sources such as air gun arrays not only are highly directional, but the frequency content and rise time of the signal also varies considerably with location with respect to the source. It would be difficult if not impossible to mimic with an omnidirectional projector these three-dimensional patterns experienced by a marine mammal as it swims near this kind of source. Initial tests could use an individual air gun source, but more realistic experiments would likely require a full array of air guns.

Most of the above considerations suggest advantages of large powerful sources if these are the kind of source whose effect must be tested. One concern for CEEs using large powerful sources is that a much greater area will be ensonified so that more non-target animals may be exposed to elevated sound levels. This is of particular concern when available observational methods cannot detect animals out to the range at which they may be affected. Where a powerful novel source is being tested, CEEs themselves should be subjected to the same risk assessment analysis as suggested above. Best available knowledge should be used to generate suitable models. Adopting a precautionary approach, tests should first be carried out within the bounds of these models and the results obtained should be used to modify these limits (see below).

Methods for estimating received levels

The source level of the playback sound can usually be predicted based upon the manufacturer's specifications and this can also be readily measured. Assessing the level that the animal receives is less straightforward but is crucial for CEE. Both accurate measures of received levels (which may be retrieved after the CEE) and a real time assessment of the subject's sound field are necessary. One straightforward solution for measuring received levels is to record directly at the animal using an acoustic recording tag (Johnson and Tyack, 2003). Acoustic recording tags have been deployed on several species. Fletcher et al. (1996) deployed the first of these recording tags on elephant seals. If real time feedback on an animal's exposure is required for the experimental design, data on received level measured at the animal can be telemetered. Seawater is a conductive medium and is not well suited to many forms of telemetry. Acoustic telemetry is the most useful for real time data transfer over ranges of a kilometer or more. If acoustic telemetry is used, care must of course be taken to ensure that the telemetry signals are not audible to the subject.

Unless data on received levels measured on the animal are telemetered to the experimenters, a real time assessment of received levels requires knowledge of the animals' location and an acoustic propagation model. We will first discuss localization, and then discuss propagation modelling. The location of animals seen at the surface can be measured using a variety of methods (e.g. laser range finding binoculars, measuring declination from the horizon with binoculars or video). However, calculating received levels in these conditions, where the animal's ears are in an acoustically complex region close to the surface is likely to be unreliable. In addition, water noise and breathing noises may mask sounds when animals are close to the surface. When animals are diving well away from the surface, the acoustic environment may be more predictable but it is harder to determine the animal's location.

Animals can be localized underwater using an acoustic transponder (Watkins et al., 1993), although the transponder signals must be designed and tested to be inaudible to the subject (Watkins and Tyack, 1991). Hydrophone arrays can be used to locate animals swimming within or close to the array using localization techniques based on the arrival times of the subject's own vocalizations or sound from an attached transmitting device. Watkins & Schevill (1972) pioneered the use of such arrays for locating marine mammals by their vocalizations. These early trials required a great deal of post-processing. Nowadays it should be possible to calculate locations in near real time, but we are unaware of any group that has published on real time 3-D passive tracking. Real time, high-resolution 3-D tracking of animals fitted with pingers has been achieved for ice breeding seals (Wartzok et al., 1992; Harcourt et al., 2000) and tracking at sea is feasible with currently available equipment. Use of multipaths (e.g. Thode et al., 2002) can enhance the determination of range and depth using sparse arrays.

Several groups are working on high acuity active sonars to track marine mammals, and this can be linked with the use of a transponder on a focal animal. Ridoux et al. (1997) used multi-beam scanning sonar to investigate the UW behavior of bottlenose dolphins without transponders. Miller (this volume) reviews an active sonar designed to detect marine mammals. [**What is this reference?**]

The received level of a signal is the source level (the level in dB measured at a distance of 1 m from the source) minus the transmission loss from 1 meter out to the location of the receiver (Urick, 1983). Once an animal's relative location has been determined, computer models can be used to calculate transmission loss (also called propagation loss) for many environments. These require information on a variety of parameters including variation in sound velocity with depth, and information about the seafloor. Received levels predicted by the propagation model should be verified in tests using a calibrated hydrophone over the expected ranges and depths of animals in the CEE. This is particularly important for shallow water environments (<1000m deep) where interactions with the bottom and greater variability in the water column can make the model predictions less precise.

The problems of locating marine mammals in three dimensions combined with inaccuracies in the predictions of propagation models mean that it is seldom possible to estimate exposure of an animal to better than 3-5 dB during an experiment.

Measuring marine mammal responses to noise.

More than a decade ago, our techniques for monitoring the behavior of marine mammals were so crude that the governing policy for harassment was that if you could detect a change in behavior, then it must constitute harassment (Swartz and Hofman, 1991). Various telemetry techniques and data logging devices have since been developed to provide detailed and objective data that allows monitoring the reactions of tagged animals to transient stimuli. VHF, satellite linked UHF, and acoustic telemetry are now routinely used to collect time series of data on location, dive behavior, swim speed, stomach temperature and heart rate. By 1994, the NRC (1994b) pointed out that "As researchers develop more sophisticated methods for measuring the behavior and physiology of marine mammals in the field (e.g. via telemetry), it is likely that detectable reactions, however, minor and brief, will be documented at lower and lower received levels of human-made sound. … In that case, subtle and brief reactions are likely to have no effect on the well being of marine mammal individuals or populations." [p 28]

NRC (1994b and 2000) argue that technological advances in our ability to monitor behavioral responses of marine mammals highlight the need to select responses for study that are biologically significant, in the sense that their disruption can be linked to adverse impact or impediments to survival, growth or reproduction. For example, the development of tags that can record acoustic dosage as well as behavioral responses (Johnson and Tyack, 2003) creates a particularly powerful tool. Johnson and Tyack (2003) provide several examples of subtle reactions of tagged cetaceans to acoustic stimuli, reactions that would not be detectable without the tag. Miksis et al. (2001) use a tag to detect subtle changes in heart rate of captive dolphins in response to acoustic stimuli. Traditional visual and/or acoustic observations can also provide detailed data on changes in patterns of movements, respiration and vocalizations. Improved understanding of marine mammal behavioral ecology should facilitate the development of models to link dose:response studies to the potential for adverse impact. Data on avoidance responses can help to quantify whether exposure to noise degrades the quality of critical habitat, and reduces resources available to a population. However, simple statistical tests for detectable differences in rates of isolated behaviors seldom can meet the NRC recommendations unless the behaviors are selected as part of a model for interpreting the biological significance of disruption. As experimenters plan CEEs and technologists add new sensors to these tags, they should cooperate to increase our ability to measure those responses most likely to get at critical issues of the biological significance of responses

NRC (1994b) distinguishes between short-term responses, which because of their brevity may be unlikely to have adverse impact, and long-term consequences that due to their extended duration are inherently more worrisome. This suggests the importance of considering the appropriate time scale when planning CEEs. Archival tags are attached for short periods with researchers tracking the animal to retrieve the tag when they detach. Such tags can store large amounts of data at high rates and this is simply downloaded after recovery of the tag. Longer-term attachments will usually require telemetry to offload data and most systems available for data telemetry have very restrictive bandwidths. While such telemetry tags may have more limited data recording capabilities, their ability to provide data as animals move over larger areas and for longer time periods is very important. Demonstration of causation between stimulus and response will usually be simpler with smaller scale CEEs, so when the policy issues focus at larger scales, a combination of small scale and larger scale CEEs will often be useful. Reaching the full potential of longer-term CEEs will require additional development of data telemetry from long-term tags.

Guidelines for Minimizing Risks Posed by CEEs

As CEEs are often used to explore the potential for noise to have harmful effects on marine mammals, the risk that the experiments themselves may harm their subjects needs to be assessed and minimized. As suggested above, the use of CEEs with novel sources should be subject to the same type of risk assessment exercise as any other acoustic emission. Without prejudging the results of such an exercise, there are aspects of CEE work that should always be considered.

CEE experiments will usually aim to expose animals to levels that elicit behavioral responses. It may be that such behavioral responses would have deleterious effects for individuals if exposures continue over long periods of time, but short-term CEEs may be able to evaluate the potential for long-term effects without incurring those risks. CEEs are not appropriate for investigating the tissue and hearing damage effects that may be caused by exposure to very high sound exposure levels, so no animal should ever be intentionally exposed to levels in the wild where injury is possible. In some cases, the sound sources employed will simply be incapable of transmitting sounds powerful enough to have these effects. Where the source level is sufficiently high to present a risk of injury, then sound levels at the subjects must be carefully controlled to minimize the risk.

The following procedures are recommended so that CEEs minimize any potential risks of exposure:

- Experiments should start by exposing animals to low received levels and increase levels gradually until a response is ob served. (A potential problem with this approach is that it may result in habituation, leading to a bias in the results with higher response thresholds being observed. This can be addressed once a response is observed at a particular level by repeating the experiment with new subjects starting just below this threshold.)
- Experimenters should be aware of the location and behavior of the closest animals during an experiment and minimize the possibility that there are other undetected animals closer than this. If species other than the subjects might be as sensitive, special attention should focus on preventing inadvertent exposure to these species.
- Work should only be conducted in good conditions of weather and with equipment such that the possibility of not detecting animals close to the source will be minimized.
- Only consider CEE when all factors—weather, equipment, team etc.—are of the highest quality and funding is adequate. This will maximize the useful data collected from each exposure and mini mize the number of exposures required to answer a particular question.
- The potential for CEEs to disrupt other scientific research or commercial activities such as whale watching should be considered and minimized. As noted earlier, there are great advantages to working in collaboration with teams conducting long-term research programs on populations

Political Considerations for Investigating Effects of Noise Exposure

In many cases there will be strong vested interests and entrenched views on both sides of issues being investigated with CEE. In particular, many industrial and military users could be severely inconvenienced if results of studies show harmful effects of manmade noise, while other groups may already be committed to assuming that a particular impact is occurring. These considerations will apply to those working on noise issues whatever research methods they use, and indeed, this is a problem common to any scientific investigation that affects powerful vested interests. One issue specific to CEEs is the intentional exposure of animal subjects to manmade noise. Some groups may oppose intentional exposure even when this is for the purpose of obtaining data critical for establishing regulations to protect animals from noise. Therefore, CEEs may be subject to pressures from all sides of the political spectrum.

Several approaches are available to minimize influence on the scientific teams doing the work and maximize the possibility that their research will faithfully reflect the real situation:

- Researchers must be independent, buffered from the influence of vested interests and should be free to publish their results speedily, openly, and ultimately in peer-reviewed journals. Powerful vested interests may exist on "both sides" of any particular issue.
- "Polluters" should pay for necessary research, but funds should be routed through a structure that ensures that the choice of researchers, the conduct of research, and analysis or dissemination of results is made based upon selection of the best science as opposed to meeting the political objectives of a special interest.

- Research should be focused on providing the answers needed for management. The existence of a research project should not be used as an excuse for a delay in taking conservation action.
- Different research groups interested in the effects of a particular noise source on a particular group of animals should coordinate to avoid unnecessary duplication. This is important not just to minimize the potential risk to subjects, but also to organize the strongest research effort. Of course replication of results is a critical part of science, so there are circumstances in which parallel or sequential efforts are appropriate.
- Where there is a great deal of uncertainty, as is the case here, a precautionary approach should be adopted, though this is difficult in situations where many noise producing activities are already well established, widely accepted, and considered essential (e.g. shipping).

Some groups may hope to rely on a strong interpretation of the precautionary principle to protect marine mammals given our ignorance of the effects of many noise sources, but as the last bullet mentions, this stance is not, in practice, very practical. Under a truly precautionary stance, no motorized ship would ever leave the harbor. The *status quo* is that marine mammals are exposed to hundreds of noise sources with unknown impact. Some may be trivial, others may pose an acute threat to animals that get too close, yet others may act in concert to chronically affect the health of populations. The critical role for CEEs is ultimately to help define where the real risks lie, and to direct regulation in this direction.

References

Au, W.W.L., Carder, D.A., Penner R.H. and Scronce, B.L. 1985. Demonstration of adaptation in beluga whale (*Delphinapterus leucas*) echolocation signals. J. Acoust. Soc. Am. 77: 726-730.

Brumm, H and Todt, D. 2004 Male-male vocal interactions and the adjustment of song amplitude in a territorial bird. Anim. Behav. 67:281-286.

Bryant, P.J., Lafferty, C.M. and Lafferty, S.K. 1984. Reoccupation of Laguna Guerrero Negro, Baja California, Mexico, by gray whales. In: The Gray Whale **Eschrichtius robustus,** eds. M.L. Jones et al., pp. 375-386, Orlando, FL: Academic Press.

Buck, J.R. and Tyack, P.L. 2000 Responses of gray whales to low-frequency sounds. J. Acoust. Soc. Am. 107:2774.

Caswell, H., Brault, S., Read, A. and Smith, T. 1998. Harbor porpoise and Fisheries: an uncertainty analysis of incidental mortality. Ecological Applications 8:1226-1238.

Cox, T.M., Read, A.J., Solow, A. and Tregenza, N. 2001. Will harbour porpoises *(Phocoena phocoena*) habituate to pingers? J. Cetacean Res. Manage. 3(1):81–86.

EPA 1992. Framework for Ecological Risk Assessment. U.S. Environmental Protection Agency, Risk Assessment Forum, Washington DC.

Erbe, C. 1997. *The Masking of Beluga Whale Vocalizations by Icebreaker Noise*. Ph D. Thesis, University of British Columbia, Vancouver, British Columbia, Canada

Fletcher, S., Le Boeuf, B.J., Costa, D.P., Tyack P.L. and Blackwell S.B. 1996. Onboard acoustic recording from diving northern elephant seals J. Acoust. Soc. Am. 100: 2531-2539.

Frankel, A.S., Mobley, J.R. Jr. and Herman, L.M. 1995. Estimation of auditory response thresholds in humpback whales using biologically meaningful sounds. In Sensory Systems of Aquatic Mammals, eds. R.A. Kastelein, J.A. Thomas and P.E. Nachtigal, pp. 55-70, Woerden, The Netherlands: De Spil Publishers.

Gordon, J, Thompson, D. and Tyack, P. 2003. The use of controlled-exposure experiments to determine the effects of noise on marine mammals: scientific, methodological and practical considerations. Report of a workshop on Controlled Exposure Experiments at the Conference of the European Cetacean Society, Rome, 10 May 2001.

Greene, C.R. Jr. and Richardson, W.J. 1988. Characteristics of marine seismic survey sounds in the Beaufort Sea. J. Acoust. Soc. Am. 83:2246-2254.

Harcourt, R.G., Hindell, M.A., Bell, D.G. and Waas, J.R. 2000. Three-dimensional dive profiles of free-ranging Weddell seals. Polar Biol. 23: 479-487

Harwood, J. 1999. A risk assessment framework for the reduction of cetacean by-catches. Aquatic Conservation: marine and freshwater ecosystems 9:593-599.

Harwood, J. 2000. Risk assessment and decision analysis in conservation. Biological Conservation. 95:219-226.

Johnson, M. and Tyack, P.L. 2003. A Digital Acoustic Recording Tag for Measuring the Response of Wild Marine Mammals to Sound. IEEE Journal of Oceanic Engineering. 28:3-12.

Kaiser, J. 2003. **Military Wins Changes That May Ease Research**. Science. 302:1487-1488. (in News of the Week)

Kastak, D., Schusterman, R.J., Southall, B.L. and Reichmuth, C.J. 1999. Underwater temporary threshold shift induced by octave-band noise in three species of pinniped. J. Acoust. Soc. Am. 106:1142-1148

Ketten, D.R. 1995. Estimates of blast injury and acoustic trauma zones for marine mammals from underwater explosions. In: Sensory systems of aquatic mammals, eds. R.A.

Kastelein, J.A. Thomas and P.E. Nachtigall. pp. 391–407. Woerden, The Netherlands: De Spil Publishers.

Lesage, V., Barrette, C., Kingsley, M.C.S. and Sjare, B. 1999. The effect of vessel noise on the vocal behavior of Belugas in the St. Lawrence River estuary, Canada. Mar Mam Sci. 15: 65-84.

Mann, D.A., Lu, Z., Hastings, M.C., Popper, A.N. 1998. Detection of ultrasonic tones and simulated dolphin echolocation clicks by a teleost fish, the American shad (*Alosa sapidissima*). J Acoust Soc Am. 104(1):562-8.
McGregor, P.K. 1992. Playback and Studies of Animal Communication. Plenum, New York. 227 pp.

Miller, P.J.O., Biassoni, N., Samuels, A. and Tyack, P.L. 2000. Whale songs lengthen in response to sonar. Nature. 405:903

Miksis J.L., Grund, M.D., Nowacek, D.P., Solow, A.R., Connor R.C. and Tyack, P.L. 2001. Cardiac Responses to Acoustic Playback Experiments in the Captive Bottlenose Dolphin, *Tursiops truncatus*. J. Comp. Psychol. 115:227-232.

NRC (National Research Council) 1983. Risk assessment in the Federal government: managing the process. Washington, D.C.: National Academy Press.

NRC (National Research Council) 1994a. Science and Judgment in Risk Assessment. Washington, D.C.: National Academy Press.

NRC (National Research Council) 1994b. Low-frequency sound and marine mammals: current knowledge and research needs. Washington, D.C.: National Academy Press.

NRC (National Research Council) 2000. Marine mammals and low-frequency sound. Washington, D.C.: National Academy Press.

Parks, S.E. 2003. Response of North Atlantic right whales (*Eubalaena glacialis*) to playback of calls recorded from surface active groups in both the North and South Atlantic. Mar Mam Sci. 19(3):563-580.

Payne, R. and Webb, D. 1971. Orientation by means of long range acoustic signalling in baleen whales. Annals of the New York Academy of Sciences. 188:110-141.

Payne, K.B., Tyack, P. and Payne, R.S. 1983. Progressive changes in the songs of humpback whales. In: Communication and behavior of whales. ed. Payne R., pp. 9-59, Boulder: Westview Press.

Richardson, W.J., Davis, R.A., Evans, C.R., Ljungblad, D.K. and Norton P. 1987. Summer distribution of bowhead whales, *Balaena mysticetus*, relative to oil industry activities in the Canadian Beaufort Sea, 1980-1984. Arctic. 40:93-104.

Richardson, W.J., Greene, C.R. Jr., Malme, C.I., and Thomson D.H. 1995. Marine Mammals and Noise. New York: Academic Press

Ridoux, V., Guinet, C., Liret, C., Creton, P.,

Steenstrup, R. and Beauplet G. 1997. A video sonar as a new tool to study marine mammals in the wild: Measurements of dolphin swimming speed. Mar Mam Sci. 13:196-206.

Schlundt, C.E., Finneran, J.J., Carder, D.A., and Ridgway, S.H. 2000. Temporary shift in masked hearing thresholds of bottlenose dolphins, *Tursiops truncatus*, and white whales, *Delphinapterus leucas*, after exposure to intense tones. J. Acoust. Soc. Am. 107:3496–3508.

Swartz, S.L. and Hofman, R.J. 1991. Marine mammal and habitat monitoring: requirements; principles, needs; and approaches. Report prepared for U. S. Marine Mammal Commission, August 1991, PB91-215046, Washington DC.

Taylor, B.L., Wade, P.R., DeMaster, D.P. and Barlow, J. 2000. Incorporating uncertainty into management models for marine mammals. Conservation Biology. 14:1243-1252.

Thode, A., Mellinger, D.K. and Martinez, A. 2002. Passive three-dimensional tracking of sperm whales using two towed arrays during the 2001 SWAMP cruise. J. Acoust. Soc. Am. 112:2399

Tyack, P. 1998a. Playback experiments of loud low frequency sound to singing humpback whales in Hawaiian waters. Whalewatcher. 37(1):3-12.

Tyack, P. 1998b. Acoustic communication under the sea. In: Animal acoustic communication: recent technical advances. eds. Hopp, S.L., Owren, M.J. and. Evans, C.S., pp 163-220, Heidelberg: Springer Verlag.

Urick, R.J. 1983. Principles of underwater sound. 3rd ed., NewYork: McGraw-Hill. 423 pp.

Wartzok, D, Sayegh, S., Stone, H., Barchak, J. and Barnes, W. 1992. Acoustic tracking system for monitoring under-ice movements of polar seals. J. Acoust. Soc. Am. 92:682-687

Watkins, W.A. and Schevill W.E. 1972. Sound source location by arrival-times on a non-rigid three- dimensional hydrophone array. Deep-Sea Research. 19: 691-706.

Watkins, W.A. and Tyack, P. 1991. Response of sperm whales (*Physeter catodon*) to tagging with implanted sonar transponder and radio tags. Mar Mam Sci. 7(4):409-413.

Watkins W.A., Daher M.A., Fristrup K.M., Howald T.J. and Notarbartolo G. di Sciara.1993. Sperm whales tagged with transponders and tracked underwater by sonar. Mar Mam Sci. 9(1):55-67.

Yelverton, J. T. 1981. Underwater explosion damage risk criteria for fish, birds, and mammals. 102nd Annual meeting of the Acoustical Society of America, Miami Beach, Florida.