

 Open access • Journal Article • DOI:10.1139/CJZ-2016-0098

A systematic review on the behavioural responses of wild marine mammals to noise: the disparity between science and policy — [Source link](#)

Catalina Gomez, Jack Lawson, Andrew J. Wright, Alejandro D. Buren ...+2 more authors

Institutions: Fisheries and Oceans Canada, George Mason University

Published on: 02 Nov 2016 - Canadian Journal of Zoology (NRC Research Press)

Topics: Noise

Related papers:

- [Evidence that ship noise increases stress in right whales](#)
- [Communication masking in marine mammals: A review and research strategy.](#)
- [Errata: Marine Mammal Noise Exposure Criteria: Updated Scientific Recommendations for Residual Hearing Effects](#)
- [Responses of cetaceans to anthropogenic noise](#)
- [Marine mammals and noise](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/a-systematic-review-on-the-behavioural-responses-of-wild-2l2kh2ksig>



A systematic review on the behavioural responses of wild marine mammals to noise: the disparity between science and policy

Journal:	<i>Canadian Journal of Zoology</i>
Manuscript ID	cjz-2016-0098.R2
Manuscript Type:	Review
Date Submitted by the Author:	08-Sep-2016
Complete List of Authors:	Gomez, Catalina; Fisheries and Oceans Canada Newfoundland and Labrador Region Lawson, Jack; Fisheries and Oceans Canada Newfoundland and Labrador Region Wright, Andrew; George Mason University Buren, Alejandro; Fisheries and Oceans Canada Newfoundland and Labrador Region Tollit, Dominic; SMRU Consulting North America Lesage, Veronique; Department of Fisheries and Oceans
Keyword:	WHALES < Taxon, dolphins, noise, BEHAVIOUR < Discipline, received level, regulations



A systematic review on the behavioural responses of wild marine mammals to noise: the disparity between science and policy

C. Gomez¹, J. W. Lawson², A. J. Wright³, A. Buren⁴, D. Tollit⁵, and V. Lesage⁶

¹ Department of Fisheries and Oceans (DFO), Marine Mammal Section, Northwest Atlantic Fisheries Centre, St. John's, Newfoundland and Labrador, Canada, A1C 5X1.

E-mail: Catalina.Gomez@dfo-mpo.gc.ca.

² DFO, Marine Mammal Section, Northwest Atlantic Fisheries Centre, St. John's, Newfoundland and Labrador, Canada, A1C 5X1. E-mail: Jack.Lawson@dfo-mpo.gc.ca

³ Department of Environmental Science and Policy, George Mason University, 4400 University Dr, Fairfax, VA 22030. E-mail: marinebrit@gmail.com

⁴ DFO, Marine Mammal Section, Northwest Atlantic Fisheries Centre, St. John's, Newfoundland and Labrador, Canada, A1C 5X1. E-mail: Alejandro.Buren@dfo-mpo.gc.ca

⁵ SMRU Consulting North America, 510-1529 6th Avenue West, Vancouver, British Columbia, Canada. E-mail: djt@smruconsulting.com

⁶ DFO, Maurice Lamontagne Institute, DFO, Mont-Joli, Quebec, Canada, G5H 3Z4. E-mail: Veronique.Lesage@dfo-mpo.gc.ca

Corresponding author: C. Gomez. Current mailing address: Department of Fisheries and Oceans, Ocean and Ecosystem Sciences Division, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada, A1C 5X1. Telephone: (902) 426-8810. E-mail: Catalina.Gomez@dfo-mpo.gc.ca.

A systematic review on the behavioural responses of wild marine mammals to noise: the disparity between science and policy

C. Gomez, J. W. Lawson, A. J. Wright, A. Buren, D. Tollit, and V. Lesage

Abstract

Noise can cause marine mammals to interrupt their feeding, alter their vocalizations, or leave important habitat, among other behavioural responses. The current North American paradigm for regulating activities that may result in behavioural responses identifies received sound levels (RL), at which individuals are predicted to display significant behavioural responses (often termed harassment). The recurrent conclusion about the need for considering context of exposure, in addition to RL, when assessing probability and severity of behavioural responses led us to conduct a systematic literature review (370 papers) and analysis (79 studies, 195 data cases). The review summarized the critical and complex role of context of exposure. The analysis emphasized that behavioural responses in cetaceans (measured via a linear severity scale) were best explained by the interaction between sound source (continuous, sonar or seismic/explosion) and functional hearing group (a proxy for hearing capabilities). Importantly, more severe behavioural responses were not consistently associated with higher RL, and *vice versa*. This indicates that monitoring and regulation of acoustic effects from activities on cetacean behaviour should not exclusively rely upon generic multi-species RL thresholds. We recommend replacing the behavioural response severity score with a response/no response dichotomous approach that can represent a measure of impact in terms of habitat loss and degradation.

Key words: cetaceans, whales, dolphins, noise, behaviour, received level, regulations.

Introduction

Sounds in the ocean generated by human activities such as hydrocarbon exploration, shipping, construction, and military-exercises are becoming more prevalent and are causing large-scale changes in the marine acoustic environment (Hildebrand 2009). Human-generated sounds (hereon referred to as noise) can have a broad variety of effects on marine mammals including loss of hearing sensitivity, deafness, behavioural change, displacement from important habitat, induced stress responses, and interference with an individuals' ability to detect, recognize and/or discriminate sounds used for foraging, conspecific communications, navigation, and predator/hazard avoidance (Richardson et al. 1995; Nowacek et al. 2007; Southall et al. 2007; Weilgart 2007; Wright and Soto 2007; Rolland et al. 2012; Ketten 2014; Erbe et al. 2015; NOAA 2015). In extreme cases, sounds from mid-frequency active sonar and multi-beam sonar might lead to mass stranding (Balcomb and Claridge 2001; Southall et al. 2013), blasts from underwater explosions have likely damaged ear structures (Todd et al. 1996), and underwater detonations have led to mortalities (Danil et al. 2011). While these extreme cases may lead to lethal impacts¹ on individual animals, disruption of behaviour (e.g., cessation of feeding) could affect numerous individuals and thus lead to sub-lethal impacts at the population level (see Pirotta et al. 2014).

The effects of noise exposure on marine mammals have typically been regulated using some threshold for the level of sound intensity (received level, RL) to which an individual (receiver) is predicted to display significant behavioural responses (often

¹ We use the term 'impact' throughout the manuscript to refer to the myriad direct or indirect consequences of changes in physical or physiological condition of the exposed animal, behaviour responses, stress responses, masking, or any other change in organismal or environmental condition resulting from that exposure. We do not intend it to refer to any specific legal definition in any nation.

termed harassment) (NOAA 2015; Scholik-Schlomer 2015). In general terms, and ignoring location-specific patterns of sound signal constructive and destructive interference, RL thresholds are applied using models of sound propagation where RL decreases with increasing distance from the sound source, and where the severity of the effect is expected to parallel this change in RL (Richardson et al. 1995). In practice, what this means is that sounds of higher intensity experienced by an individual close to the sound source could cause a loss of hearing sensitivity (Zone 1 in Figure 1), whereas the same sounds have reduced intensity farther from the source and could cause other effects, such as a change in behaviour (Zone 2 in Figure 1).

RL thresholds for injury (Zone 1 in Figure 1) are now considered to be specific to marine mammal's functional hearing group (a proxy for individual's hearing capabilities) and have been expressed in terms of Sound Pressure Level (SPL) and Sound Exposure Level (SEL) (Southall et al. 2007; NOAA 2015). Currently, there are no specific RL thresholds for explicitly assessing or regulating masking or stress responses (Zone 2 in Figure 1). However, there are tools available to quantify the potential loss of acoustic communication space, and thus, to potentially include this effect as part of noise impact assessments (Clark et al. 2009; Hatch et al. 2012; Moore et al. 2012; Erbe 2015; Erbe et al. 2015).

The generic RL thresholds for behavioural disturbance in current use in North America originated from a few field studies in the 1980's on baleen whales (e.g., Malme et al. 1983; 1984; Richardson et al. 1986; 1990). Thresholds for behavioural disturbance were

proposed for various types of sounds (impulsive or continuous) and applied to all marine mammals regardless of their functional hearing group (Figure 1, HESS 1999, see Scholik-Schlomer 2015). These generic RL thresholds, although not applied universally (e.g., Federal Register 2003; Wood et al. 2012), are still used and have been applied in several impact assessments (e.g., LGL 2010; DFO 2012; 2014; Statoil 2015; BP environment plan²). Efforts have been made to improve them using new data (e.g., Southall et al. 2007; Federal Register 2009), including the addition of unique RL thresholds for behavioural disturbance specific to species that were deemed particularly sensitive to noise: harbour porpoises (*Phocoena phocoena*) and beaked whales (family: Ziphiidae) (e.g., Federal Register 2015).

Current scientific knowledge recognizes that acoustic characteristics of the sound source, marine mammal's hearing sensitivity, and context of exposure must be considered in addition to RL and species sensitivity to predict the probability and severity of behavioural response of a marine mammal exposed to a sound source (Wartzok et al. 2004; Southall et al. 2007; Ellison et al. 2012). Context of exposure, beyond that of the sound source type and spectral characteristics, includes variables such as movement and depth of the sound source, the ratio of signal to background noise, sound level above hearing threshold, proximity of the source to the receiver, bathymetry in the exposure area, as well as the receivers' species, sex, age, reproductive state, behavioural state prior to sound exposure, prior experience, motivation, and behavioural conditioning (Richardson et al. 1995; Wartzok et al. 2004; Nowacek et al. 2007; Southall et al. 2007; Weilgart 2007; Ellison et al. 2012).

² <http://www.bpgabproject.com.au/go/doc/5771/2501234/Environment-Plan-FAQ-s-and-information->

Substantial efforts are under way to amend the generic single multi-species RL thresholds paradigm to fit the current scientific understanding that RL alone is inadequate for predicting behavioural disturbance (Southall et al. 2007; Ellison et al. 2012; Miller et al. 2012; Southall et al. 2012; Goldbogen et al. 2013; Alves et al. 2014; Antunes et al. 2014; Miller et al. 2014). In particular, NOAA is currently developing “national guidance for better predicting significant behavioural effects” (Scholik-Schlomer 2015).

A comprehensive framework has been proposed to start to account for the effect of context of exposure on individual response to noise (Southall et al. 2007; Ellison et al. 2012). This framework includes a scoring scale for severity of behavioural responses of marine mammals to noise (Table 4 in Southall et al. 2007). The intent of this response severity scale was to reflect increasing severity of impact by delineating behavioural responses that are relatively minor and/or brief (scores 0-3); those with higher potential to affect foraging, reproduction, or survival (scores 4-6); and those considered likely to affect foraging, reproduction, or survival (scores 7-9) (Southall et al. 2007). The definitions in this behavioural response severity scale approach can be very subjective; for example, “brief or minor cessation/modification of vocal behaviour” results in a score of two while “minor cessation or modification of vocal behaviour (duration < duration of source operation)” corresponds to a score of three (Table 4 in Southall et al. 2007). This type of vague definitions led to amendments of this severity scale in other studies to make it more detailed and less imprecise, though definitions were often arbitrary due to lack of information (Miller et al. 2012; Williams et al. 2014). The behavioural response

severity scoring approach has additional problems; for example, Tougaard et al. (2015a) discouraged its use as it does not take into account long-term changes in animal behaviour (Moore and Clarke 2002; Bejder et al. 2006a,b; Pirodda et al. 2015; Weaver 2015), and most importantly, response severity may be falsely interpreted as equating to severity of impacts on individuals or populations. This proportionality was not implied by Southall et al. (2007) or Ellison et al. (2012) but “it is tempting to make this inference and hence use the severity scores as a proxy for population impact” (Tougaard et al. 2015). Despite these recognized issues, this severity scoring is the *de facto* standard measure of behavioural responses in the recent proposed framework (Ellison et al. 2012) and many noise impact assessments (e.g., SMRU Consulting 2014; Heinis and de Jong 2015).

The generic behavioural disturbance RL thresholds are an example of the disparity between science and policy. On one hand, North American regulators have retained RL-based thresholds (Table 1 in Scholik-Schlomer 2015) despite the increasing evidence that RL alone is inadequate for predicting behavioural responses: the acoustic characteristics of the sound source, marine mammal’s hearing sensitivity and context of exposure must be considered (e.g., Nowacek et al. 2007; Southall et al. 2007; Weilgart 2007; Ellison et al. 2012). On the other hand, although attempts are under way to update these thresholds (see Scholik-Schlomer 2015), regulatory guidance in Canada and in the United States have not yet provided updated behavioural disturbance RL thresholds that encompass the most up-to-date scientific knowledge. The discrepancy between science and policy is further complicated by the fact that the effects of noise on

marine mammals are a very controversial topic in the scientific, regulatory, and public realm that has resulted, among other things, in several lawsuits in the United State and elsewhere. In some instances, conclusions of research have favoured the interests of sponsors that funded studies potentially used in the process of informing policy decisions (Wade et al. 2010).

In this context, a systematic literature review and analysis within the behavioural response severity-based framework proposed by Southall et al. (2007) and Ellison et al. (2012) was essential. We compiled and summarized information on species, sound sources, context of exposure, and marine mammal behavioural responses with the goal of evaluating which variable(s) best explained marine mammal behavioural responses to noise. Our initial goal was to advance the process of improving the current RL threshold(s) paradigm by incorporating the best available science to highlight the relative importance of RL and context of exposure in predicting probability and severity of behavioural responses, and propose more meaningful metrics to replace previous generic metrics where single RL thresholds are applied across species. This paper provides a synthesis of the results of the systematic literature review and analysis, which instead, lead us to call for a paradigm-shift in the way in which behavioural responses of marine mammals to noise are evaluated.

Materials and Methods

A systematic literature review was conducted to gather studies, available from 1971 up to June 2015, on the behavioural responses of wild marine mammals to noise. The

public database PubMed was searched using the specific terms [MARINE MAMMAL or WHALE or DOLPHIN or PORPOISE or PINNIPED or CETACEAN or SEAL or SEA LION] and [NOISE or SONAR or SEISMIC or VESSEL or PILE-DRIVING or CONSTRUCTION or ALARM] and [BEHAVIOUR or BEHAVIOR]). Review papers (Nowacek et al. 2007; Southall et al. 2007; Weilgart 2007) were used to manually extract additional references that were not identified while using the search terms. Studies were filtered so only those that provided information on the behavioural responses of marine mammals to noise were reviewed further (Figure 2, S1³).

Only studies in the wild were considered in this review because of the important differences in the life history and sound exposure contexts for captive and wild individuals (Parsons et al. 2009; see Tougaard et al. 2015), and because the objective of this study was to review behavioural responses of marine mammals in their natural environment. The participation of captive individuals in sound exposure experiments is often reinforced with food, which likely motivates them to tolerate high RL (Scheifele et al. 2005; Southall et al. 2007). Further, cetaceans in captivity cannot leave their tank, and therefore, cannot express their full repertoire of behavioural reactions (e.g., cannot move away several kilometres as in Tougaard et al. 2015). Response thresholds derived from older animals might also not be representative of younger animals in the wild (e.g., old individuals lose their ability to perceive certain frequencies) (see Parsons et al. 2009; Tougaard et al. 2015). Consequently, studies focused on measuring

³ Supplementary material 1: The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.

permanent and temporary thresholds shifts, natural sound sources, and studies in captivity, were discarded.

For the studies further considered, information on 1) behavioural responses of marine mammals to noise, 2) sound source, and 3) context of exposure, was compiled. Review papers were used to extract estimated RL values that were not provided in the original studies. All data cases gathered from the studies included in either the qualitative or quantitative synthesis (Figure 2) are available in S2⁴.

Marine mammal behavioural responses to noise

For each study, a summary of the behavioural reactions reported during sound exposure was provided in a narrative extracted or summarized from each paper. Some studies provided detailed information on the response of different individuals or populations; in those cases one study was represented by multiple data cases (see raw data in S2). Southall et al. (2007) developed a qualitative, 10-step index to rank the severity of behavioural responses on the basis of the observed physical magnitude of each response (e.g., minor change in orientation, change in respiration rate, fleeing the area), and its potential biological significance and thus impact (e.g., cessation of feeding, separation of mother and offspring). Using this index, a severity score was assigned to each behavioural reaction reported during noise exposure as per Table 4 in Southall et al. (2007). Note that the behavioural response severity scoring was based exclusively on the information readily available in the studies reviewed (i.e., additional

⁴ Supplementary material 2. Interactive supplementary material to download and display the publications and data cases included in the systematic review and meta-analysis.

unpublished information about context, methodology, or details of the experiments that might be available through contact with the authors or colleagues were not considered).

Studies in this review also reported very dissimilar levels of information regarding the behavioural response of the subject marine mammal. Information ranged from very detailed data that were compiled through focal follows and/or digital acoustic recording tags attached to study animals, to general and subjective descriptions of changes in speed in which an animal was traveling without any details about the magnitude or duration of the observed change. This variability in detail reporting did not permit a well-defined differentiation among severity scoring definitions (e.g., brief, minor, moderate response, etc., as per Table 1), making it unfeasible to have a clear distinction between numerical scores. Therefore, we further grouped the response variable (behavioural response severity score) into four wider categories: very high, high, moderate and low (Table 2). This partitioning represents a compromise between the loss of too much information from the more subjective studies against the loss of precision from better-presented information in the more detailed studies. For comparison, the Southall et al. (2007) 10-point scale has been simplified in other studies to a binary cut-off for “likely to affect” by Miller et al. (2012) and more detailed definitions by Williams et al. (2014).

Factors affecting marine mammal behavioural responses to noise

Data on sound source

Sounds can be characterized by their amplitude or source level (SL), by their pitch or frequency and by their duration. Hence, when available, the following information was recorded: peak or predominant source frequency (Hz), duration or duty cycle, SL and

RL (either measured or estimated). The sound sources most frequently reported were divided into four categories based on similarities in terms of their frequency range of emission, duty-cycle and relative SL: continuous (shipping including ice-breaking, construction, vibratory drilling), low-frequency active sonar (LFAS), mid-frequency active sonar (MFAS), and seismic/explosions.

Studies were often not explicit as to whether the frequency reported for a sound source corresponded to the peak, predominant, or range with the most energy. Furthermore, there was incongruence in the definitions provided by some studies when categorizing LFAS and MFAS based on frequency range of emission: while the United States Department of the Navy defines MFAS in the range between 1-10 kHz, several studies in Europe have used the term LFAS for sources that are within 1-2 kHz (e.g., Miller et al. 2012; Alves et al. 2014; Antunes et al. 2014). For consistency, sonar in the range between 1-10 kHz was classified as MFAS (Table 3).

The majority of studies that reported RL provided them as broadband sound pressure level (SPL) in dB re 1 μ Pa using a root-mean-square (rms) sound amplitude (or mean square pressure) within a given time window and frequency range. RL were also expressed as peak sound pressure level, measured in frequency bands (e.g., 1/3-octave band level), reported as a range of values (e.g., 150-200 dB re 1 μ Pa), or as a specific point value, such as an average or a sound dose received by the animal (Antunes et al. 2014). In the latter, experiments were designed based on dose escalation of levels of sounds to determine at what sound dose individuals changed their behaviour (e.g., Miller et al. 2012).

Data on species: accounting for hearing sensitivity

Five functional hearing groups were proposed by Southall et al. (2007), and are used by the National Oceanic and Atmospheric Administration (NOAA) and other groups conducting environmental impact assessments with the aim of representing similarities among marine mammals in known or expected hearing capabilities (Table 4, e.g., LGL 2010; JASCO 2015; NOAA 2015). We used a similar partitioning also as a proxy for known or expected hearing sensitivity, with the addition of manatees: 1) low-frequency hearing (LF baleen whales), 2) mid-frequency hearing (MF toothed cetaceans; toothed cetaceans other than those in the high-frequency hearing category), 3) high-frequency hearing (HF toothed cetaceans; e.g., harbour porpoises [*Phocoena phocoena*] and river dolphins), 4) phocid pinnipeds (seals), 5) otariid pinnipeds (eared seals) and 6) manatees.

Data on context of exposure

To examine the importance of factors other than RL in the severity and probability of behavioural response, information on the following contextual variables was compiled for each study, when possible: movement of the sound source (moving, moving/approaching, stationary), proximity of the source to the observed individual(s), depth of the sound source (m), number of individuals (sex, age, and individual identification number), naivety, behaviour prior to sound exposure, and location of study.

Naivety was categorized as “1” if the population or individual was believed to have been exposed to a given noise for the first time in the study reviewed, or “0” if the population or individuals were already familiar with the sound. Studies were recorded as being

conducted in either captivity or the wild, and for studies in the wild, a general geographical location was recorded (e.g., Ligurian Sea, Italy).

Assessment of sample size and data quality for analysis

The next step after standardizing variables gathered in the literature consisted in conducting an assessment of the quality and number of data cases suitable for analysis (Figure 3, more detailed information can be explored using the raw data available in S2). First we ensured that data cases included a representation of the type of sound source, RL and a proxy for marine mammal's hearing sensitivity (functional hearing group) (Panel A in Figure 3). The majority of data cases were for MF toothed cetaceans and LF baleen whales exposed to continuous, MFAS, and seismic/explosion sound sources ($N = 199$).

Subsequently, the feasibility of including contextual variables was examined. The studies reviewed were conducted in various geographic locations, years and seasons; studies used different sampling methods, observers, and reported various sample sizes. In addition, most studies reported limited information on contextual variables, meaning that little contextual data could be compiled. For instance, behaviour of marine mammals prior to sound exposure has been proposed as one of the key contextual features influencing marine mammal behavioural responses to noise. However, this information was only provided in a limited number of the studies reviewed, and was presented in various ways when provided (Table 5). After attempting to standardize this variable, it was evident that most contextual data cases (80%) were linked only to MF toothed cetaceans exposed to continuous and MFAS (Panel B in Figure 3). This is

further complicated by the inconsistencies among studies in the definition of LFAS, as presented in Table 3.

Information reported on sound source context (e.g., movement) was even less detailed and consistent (Panel B in Figure 3), as were details on proximity of individuals to the source ($N = 116$ data cases) and number of individuals ($N = 128$; ranging from detailed information [e.g., 1 female] to very vague information [e.g., several individuals]) (see data in S2). In the case of naivety, it was nearly impossible to determine with certainty whether individual animals or populations were naïve, with the exception of one study (Finley et al. 1990).

An analysis with this degree of heterogeneity, lack of standardization across definitions and small sample sizes, although mathematically possible (if all combinations are represented, which is not the case here), is inadvisable (see McKenzie et al. 2013). Thus, the inclusion of this type of data violates numerical and statistical standards in the pursuit of noise exposure criteria (Wright 2015). Consequently, a detailed qualitative analysis of the critical and complex role of context of exposure is presented in detail in the discussion and S3⁵ but the statistical analysis did not include most contextual explanatory variables. The full list of variables that could be considered in the analysis was: functional hearing group, type of sound source and RL.

⁵ Supplementary material 3. Analysis and data exploration of additional variables.

Meta-analysis

The main goal of this analysis was to determine what variable(s) best explained marine mammal behavioural responses to noise. For consistency only the maximum RL expressed as sound pressure level (RL_{SPL}) and the maximum behavioural response severity score assigned was considered (Figure 3). The analysis was conducted by 1) combining RL_{SPL} measured as rms, peak or peak-to-peak ($N = 195$ data cases), and 2) using only RL_{SPL} measured as rms ($N = 173$ data cases, Figure 3).

The response variable (behavioural response severity score) is a discrete variable with a clear order that ranges from low, moderate to high. Therefore, a regression model for ordinal dependent variables was performed (ordinal logistic regression, Agresti 2002). We considered as explanatory variables: 1) RL (hypothesized to parallel the change in the severity of the behavioural response), and controlled for functional hearing group (to account for known or expected cetacean hearing sensitivity) and sound source (to account for similarities in terms of their frequency range of emission, duty-cycle and relative SL). Potentially, the effects of either variable on the behavioural responses exhibited depend on the level of the remaining variables. We therefore included second order interactions. As the main goal of this meta-analysis was to determine what variable(s) best explained marine mammal behavioural responses to noise, all possible candidate models were built considering second order interactions, and their relative empirical support was weighted.

$$\text{Behavioural Response Severity Score} \sim \text{RL} + \text{Functional hearing group} + \text{Sound Source} + \text{Functional hearing group} * \text{Sound Source} + \text{Functional hearing group} * \text{RL} + \text{Sound Source} * \text{RL}$$

Model comparison and selection was based on the Akaike Information Criterion corrected for small sample sizes (AICc, Burnham and Anderson 2002) and the derived measure evidence ratio (E_i , Anderson 2008). The evidence ratio quantifies model selection uncertainty and the weight of evidence for each model (Anderson 2008). Models with a delta AICc less than 7 (Burnham et al. 2011) and those with $E_i < 10$ are presented. The relative variable importance $w^+(j)$ was calculated to quantify the evidence of the importance of each explanatory variable (j) in the set of candidate models (Burnham and Anderson 2002). Values for $w^+(j)$ range from 0 to 1; larger values indicate higher importance of a variable (j) relative to the other variables. Finally, to explore the fit of the model selected, we plotted the proportion of the observed and predicted relative proportion of data cases for each behavioural response severity score.

Results

Of the 792 publications identified, 370 publications were screened and 219 met the criteria for inclusion in the qualitative analysis (Figure 2). Data cases that reported RL and behavioural responses of cetaceans to noise ($N = 256$, Figure 3, S2) included information on 35 species of marine mammals across five functional hearing groups: HF toothed cetaceans ($N = 18$), LF baleen whales ($N = 75$), MF toothed cetaceans ($N = 145$), phocid pinnipeds ($N_{\text{in water}} = 8$, $N_{\text{in air}} = 7$), and manatees ($N = 2$) (one data case was for species aggregated within several groups).

The analysis was performed for those groups that contained the majority of data cases: MF toothed cetaceans and LF baleen whales exposed to continuous, MFAS, and

seismic/explosion sound sources (“very high” behavioural response was not included due to low sample size, see Figure 3). This corresponded to a total of 79 studies and 195 data cases (see Figure 3 and Table 6). Note that not all studies used in the analysis are referenced in this manuscript but a full list is presented in S3.

Sound levels received by the animals did not explain the severity of behavioural responses: more severe behavioural response severity scores were not consistently related to higher RL and less severe behavioural response severity scores were not consistently related to lower RL (Figure 4a). This pattern is consistent across all the data compiled (S2).

The lack of relationship between RL and the severity of behavioural responses was upheld for MF toothed cetaceans exposed specifically to both continuous (Figure 4b, $N = 49$) and MFAS (Figure 4c, $N = 81$). Similarly, for LF baleen whales exposed to seismic sounds, it was not possible to differentiate between the curves representing high and low response severity in relation to RL (Figure 4d, $N = 41$). However, in this case, behavioural responses of moderate severity were found to be more common at RLs that were lower than either high or low severity responses. In conclusion, even when comparing one cetacean functional hearing group (which might reasonably be considered a general proxy for species with similar known or expected hearing capabilities) with one type of sound (which provides a general proxy for frequency, duration and SL of the sound source) the RL still did not vary in relation with the severity of behavioural responses. Logically, the application of a standard (and group-specific) frequency (hearing sensitivity) weighting (e.g., like M-weightings or US Navy Type-2 weightings) to received SPLs would result in a similar lack of RL relationship. This is

because the frequency weighting effect on the SL of a particular sound source would be similar across the entire frequency hearing group.

The most parsimonious model to describe the severity of behavioural response included the interaction between sound source and functional hearing group, but did not include RL (Table 7). There are few data cases for baleen whales (LF) exposed to MFAS and the model generally fits better for toothed cetaceans than baleen whales. Clearly this is reflecting the effect of sample sizes, but there could be other type of effects (e.g., baleen whales and toothed cetaceans have a different repertoire of behavioural responses, see discussion). With regard to sound source, continuous sources were found to elicit less severe behavioural responses from MF toothed cetaceans (Figure 5). Additionally, seismic/explosion sound sources were found to elicit substantially more high-severity behavioural responses than either moderate- or low-severity responses from LF baleen whales. It is unadvisable to further interpret these patterns of severity due to the uneven and small sample sizes.

The most parsimonious model had approximately twice the empirical support of the second best model ($E_i = 2.42$), which included RL in addition to the interaction between sound source and functional hearing group (Table 7). The inclusion of RL into the model did not improve the fit (difference in log likelihoods: 0.20). This indicates that the small difference in the value of the delta Akaike between the first and second best models ($\Delta AIC = 1.77$, see Table 7) is due to the bias correction term not being large enough, thus the variable RL conforms to the definition of “pretending variable” (Anderson 2008). Pretending variables may enter a model set with a relatively small ΔAIC , thus misleadingly informing that it is one of the best models. A pretending variable is

identified when the fit of the model does not improve with the added parameter (difference in log likelihoods ~ 0) (Anderson 2008). Pretending variables can increase model selection uncertainty and may bias multi-model inference; thus, pretending variables should be removed from the model set (Anderson 2008).

The analysis of relative variable importance showed that the type of sound ($w_{(sound\ source)}=0.988$) and functional hearing groups ($w_{(hearing\ group)}=0.976$) were important variables for describing the severity of behavioural response of wild cetaceans, while RL was superfluous ($w_{(RL)}=0.344$).

The best model showed a reasonably good fit: there was little discrepancy between the predicted relative proportion of data cases among the behavioural response severity scores (black points with corresponding 95% confidence intervals) and the reported observations (bars) (Figure 5). As would be expected, prediction-observation discrepancy was often larger for combinations with smaller samples sizes, such as the severity of response of LF baleen whales exposed to MFAS and MF toothed cetaceans exposed to seismic-explosion sound sources.

The results presented above were for a combination of RL_{SPL} measured as rms, peak or peak-to-peak. However, the patterns in the results were unaltered when we repeated the analysis using only RL_{SPL} rms measurements (Section I - Analysis 1 in S3).

Similarly, the patterns in the results were unaltered when we replaced the behavioural response severity score with a dichotomous response (avoidance or no-avoidance; Section I - Analysis 2 in S3). The main patterns also remained the same when removing data cases that included replicates of the same individual (Section II in S3), and when

removing data cases from playbacks of recorded sounds of airguns and drilling/construction (Section III in S3). Accordingly, the main conclusions drawn from this exercise are unaffected by the different measurements of RL, by pseudo-replication, when considering only avoidance and no-avoidance behavioural responses, and when restricting the analysis to sounds from real operating sound sources only: more severe behavioural response severity scores were not consistently related with higher RL, and vice versa.

Discussion

The studies reviewed have persistently concluded that context of exposure plays a critical and complex role in modulating the severity of behavioural responses of marine mammals to noise (Wartzok et al. 2004; Southall et al. 2007; Ellison et al. 2012).

However, it was impossible to include many contextual factors due to lack of reporting, specificities, and lack of standardization across the many studies reviewed.

Consequently, based on the information available in the literature, the analysis supports three main conclusions:

1. More severe behavioural responses in cetaceans (measured via a linear severity scale) were not consistently associated with higher RL, and vice versa (Figure 4, S2);
2. Behavioural responses in cetaceans were best explained by the interaction between sound source type (continuous, MFAS or seismic/explosion) and functional hearing group (a proxy for hearing capabilities) (Table 7); and

3. The different cetacean functional hearing groups have, at least for certain sound sources with adequate sample sizes, dissimilar response profiles. Baleen whales (with low-frequency hearing) and toothed cetaceans (with mid-frequency hearing) exhibited dissimilar severity of behavioural responses depending on the type of sound source. For example, mid-frequency hearing toothed cetaceans had relatively less severe behavioural responses when exposed to continuous sounds (vessels), compared to baleen whales which showed more severe responses when exposed to continuous and seismic/explosion sounds (Figure 5).

Differential responses of marine mammals with different functional hearing capacity to sound sources of various types have been described extensively in the literature (for a review see Southall et al. 2007), and this study confirms quantitatively the interplay between these two factors and their importance in predicting severity of behavioural response. There are many examples of this in the literature. For example, feeding gray whales (*Eschrichtius robustus*) avoided airgun sounds (RL: 100 – 120 dB re 1 μ Pa), but not drillship sounds at higher RLs (Malme et al. 1986, RSL between 149 – 176 dB re 1 μ Pa). Similarly, blue whales (*Baleanoptera musculus*) were less likely to produce calls during exposure to MFAS (RL: < 160 dB re 1 μ Pa rms), but more likely to produce calls when a ship was nearby (Melcón et al. 2012).

An overlap in the frequency range of the sound source and hearing is obviously an important factor in determining whether or not an individual is likely to detect and respond to a sound source. However, the differences in the severity of behavioural response observed among functional hearing groups may not solely result from hearing

capabilities, but possibly also from other factors that are taxon-specific given that low-frequency hearing cetaceans are baleen whales and mid-frequency hearing cetaceans are toothed cetaceans. If a match in the frequency band of sound sources and hearing capabilities were solely responsible for the results, one would expect relatively more severe responses to mid-frequency sounds such as MFAS in MF toothed cetaceans compared with LF baleen whales. However, this does not always appear to be the case (e.g., Figure 5, Melcon et al. 2012). The interpretation of the finer details in the results of Figure 5, however, must be considered carefully as the majority of data cases yielding low behavioural scores for MF toothed cetaceans to continuous sources were drawn from a single study (Williams et al. 2014). Similarly, the majority of cases yielding high behavioural response scores in LF baleen whales exposed to seismic sources originated from one study (Ljungblad et al. 1988). Accordingly, the designs of these studies, including any associated sampling biases, may be influencing these conclusions substantially.

This review assessed the importance of a generic multi-species RL in the severity of the behavioural response elicited in cetaceans, and concluded that this variable, without any further contextual information, was trivial in determining the severity of behavioural response severity score, even in combination with sound source type and functional hearing group. To put it simply, and noting that we were not able to include many contextual variables into the model, more severe behavioural response severity scores were not consistently reported at higher RL, and weaker responses were not consistently reported at lower RL (Figure 4, S2). This pattern persisted when replacing the behavioural response severity scale with a dichotomous response (avoidance or no-

avoidance, S3). This suggested that the intrinsic characteristics of each sound source (e.g., frequency, duration, etc.) may be of greater importance in predicting the severity of behavioural responses than RL.

One of the key assumptions of behavioural response severity scales, such as the scale proposed by Southall et al. (2007), is that the severity score parallels the RL. In their review, however, Southall et al. (2007) proposed that for LF baleen whales exposed to non-pulsed sound sources (e.g., vessels, aircraft, drilling or wind turbines, and many active sonar systems), context of exposure had a role at least as important as exposure level in predicting severity of behavioural responses. There was also no clear relationship between RL and severity of response for MF toothed cetaceans exposed to non-pulsed sound sources (Southall et al. 2007). Subsequently, Ellison et al. (2012) proposed to separate behavioural response severity scoring into two classes: 1) lower-level responses (score 0 - 4), which are more likely described and assessed according to sound exposure and context, and 2) higher-level responses (scores 5 – 9) described and assessed according to a dose-response approach. In this conceptual framework, higher and lower RL values are expected to be associated with higher and lower behavioural response severity scores, respectively (Figure 2 in Ellison et al. 2012).

Both Southall et al. (2007) and Ellison et al. (2012) concluded that establishing some sort of dose-response relationship between RL and behavioural response severity was not possible without much additional information. This analysis strongly supports those conclusions. However, the review also demonstrates that there is currently not enough data in the published literature to include consideration of all the necessary contextual information in a generalised statistical model. Furthermore, the results of our review,

summarized in Figure 4a, indicate that while behavioural responses were observed starting at approximately 110 dB re 1 μ Pa, high-severity behavioural responses were equally likely as low- and moderate-level behavioural reactions at these RLs⁶. This suggests that a relatively low RL might result in biologically significant impacts (by affecting foraging, socializing, reproduction, or overall survival).

The available literature has demonstrated cases where the severity of the acoustic behavioural responses to noise indeed escalated in combination with increasing RLs. These species, source, and context- specific examples include: increasing call amplitude in northern right whales (*Eubalaena glacialis*) when background sound levels rose (Parks et al. 2011), reduction in the number of humpback whales (*Megaptera novaengliae*) singing with increasing seismic survey pulse RL (Cerchio et al. 2014), and increased humpback whale song length in response to higher-levels of LFAS playback (Fristrup et al. 2003). Similar results have been presented for several other species, including, but not limited to, striped (*Stenella coeruleoalba*), Atlantic spotted (*Stenella frontalis*), and short-beaked common dolphins (*Delphinus delphis*, Papale et al. 2015), bowhead whales (*Balaena mysticetus*, Ljungblad et al. 1988), California sea lions (*Zalophus californianus*, Houser et al. 2013a,b), Australian fur seals (*Arctocephalus pusillus doriferus*, Tripovich et al. 2012), blue whales (McKenna 2011), and killer whales (*Orcinus orca*, Holt et al. 2009; Miller et al. 2014). Following this pattern, but adding further complexity, a study examining behavioural responses of bowhead whales to increasing RL of seismic sounds showed an initial increase in calling rates at low RL

⁶ For reference, the mean broadband source level of several ship types combined is 173 (+/- 7) dB re 1 μ Pa at 1 m (s (20 – 40,000 Hz) (Table 2 in Veirs et al. 2016).

(~94 dB re $1 \mu \text{Pa}^2\text{-s}$), followed by a decrease in calling rate at higher RL (~127 dB re $1 \mu \text{Pa}^2\text{-s}$) to near silence at the highest RL (160 dB re $1 \mu \text{Pa}^2\text{-s}$, Blackwell et al. 2015).

The studies mentioned above were all for single species exposed to one sound source type, and many of those examples (e.g., Holt et al. 2009; McKenna 2011; Cerchio et al. 2014; Papale et al. 2015) were specifically for one type of acoustic behavioural response: the Lombard vocal behavioural response. The Lombard response is the tendency of a person or animal to modify their vocal amplitude or the acoustic frequency of the sounds they produce when challenged by a noisy environment (Patricelli and Blickley 2006; Hotchkin and Parks 2013). Accordingly, this behavioural response may be inherently more connected to RL than other possible behavioural responses (e.g., change in respiration rates, avoidance). In fact, if all behavioural responses were a highly non-linear function of RL (e.g., a step function such that there was some level at which the probability of a response changes very rapidly), this may have been detected in the analysis. This is not the case and in many instances reported in the literature, behavioural responses were explained by a combination of context variables, where RL was merely one of them. For example, killer whale behavioural responses to ship passages was best explained by a combination of year, month, age of the individual, number of ships, and RLs (Williams et al. 2014). These species/source specific results are valuable for predicting severity of behavioural responses in similar contexts of exposure. However, the results of our analysis indicate that these studies should not be used to extrapolate to other species, sound sources and/or context in a generalised framework.

Despite being confident in our conclusion that the types of sound and functional hearing groups were important variables for describing the severity of behavioural response of wild cetaceans, while RL was extraneous, we acknowledge the limitations of this study. We believe that there are at least three additional potential non-exclusive reasons for why the severity in the behavioural response was explained so poorly by RL, which we will explore in the following sections:

1. The response variable (behavioural response severity score and corresponding definitions) are not providing accurate measures of behavioural responses of marine mammals to noise; and/or
2. RL across studies were measured, estimated and/or reported inconsistently, and thus were not comparable across the majority of studies; and/or
3. The lack of data on additional, and potentially more important explanatory variables (e.g., context of exposure and natural variation), is overshadowing any small contribution from RL.

1. The behavioural response severity scoring scale and definitions used in this review are not appropriate

i. Scoring difficulties

This review intended to score every reported marine mammal behavioural response to noise exposure via a linear severity scale (see Southall et al. 2007; Ellison et al. 2012). This type of scoring has been successfully applied to some studies that provided very detailed information on behavioural responses (Miller et al. 2012), although they

sometime required more detailed albeit arbitrary definitions of the severity score (Williams et al. 2014). The process of assigning a score in this review was problematic, in part due to the simplistic, broad descriptions of changes in behaviour reported (e.g., “brief” versus “minor”, as per Table 1). Another issue, even with the simple descriptions, was that it was often nearly impossible to assign reported responses to specific numerical scorings (Table 1 and 2). For example, there was often not enough detail in the reviewed study to distinguish between the behavioural response severity score 4 (changes in locomotion speed) and 5 (extensive or prolonged changes in locomotion speed). This is the main reason for combining both numerical scores into a “moderate” response in this analysis.

ii. An absence of behavioural responses does not equate to an absence of impact: suitability of behaviour as a proxy for impact

The natural extension of this uncertainty in low-severity scoring is the consideration of reported absence of behavioural responses. In many cases, such reports may simply represent a mismatch between the study design and the actual responses of the animal, but there will also be occasions where these reports do indeed represent a true absence of behavioural response. However, it is here where the more fundamental limitations of a focus on behavioural responses become apparent. For example, in areas where noise is recurrent, sensitive animals might have already left the area prior to a study being conducted, leaving results driven only by the most noise-tolerant individuals of a population (Bejder et al. 2006a; 2009). There are occasions when individuals may be unwilling to display any avoidance response due to the value of a habitat for feeding or reproduction (Bejder et al. 2009). Remaining in a disturbed area of

higher sound exposure can impact foraging performance or reproductive rates via social disruption, leading to potential long-term consequences for energy budgets or fitness (Richardson et al. 1995; Williams et al. 2006; Lusseau et al. 2009; Pirodda et al. 2014; Tougaard et al. 2015).

In a similar manner, changes in behaviour may also impose an energetic cost on individuals (see Patricelli and Blickley 2006). For example, while the Lombard vocal response facilitates coping with noise, it carries an energetic cost and has physiological limits (Holt et al. 2015). Other vocal coping strategies (Table 8) may carry unquantified costs, including missed opportunities to detect sounds from predators, prey or conspecifics (Tyack 2009).

Furthermore, behavioural change does not account for potentially cumulative or synergistic effects. For instance, no behavioural changes were documented in feeding humpback whales exposed to underwater explosions in Newfoundland, Canada (Todd et al. 1996). However, a subsequent increase in entrapment rate in the area may have been related to the long-term effects of the exposure; dissections of the peripheral auditory systems of two stranded whales revealed damaged ear structure likely resulting from blast shock waves (Ketten et al. 1993; Ketten 1995; Todd et al. 1996).

In brief, the absence of a behavioural response to noise does not necessarily mean that there is an absence of a biological impact; and a strong behavioural response does not always indicate population-level effects as, for instance, non-responsive individuals might be habituated or may be reacting in ways that are not observed or measured (Lusseau et al. 2009; Wright and Kyhn 2014; Tougaard et al. 2015). Caution is thus

warranted when considering solely displacement as a behavioural response, particularly for populations that exhibit high site fidelity and thus are less likely to leave a disturbed area. Instead impact evaluation needs to consider both the potential short- and long-term adverse energetic (and other physiological) consequences of impacting critical life functions measured via behavioural changes (Castellote et al. 2012). This will more likely account for potential sub-lethal effects, such as the alteration of foraging performance, rather than relying solely on displacement or presence (McCarthy et al. 2011; Castellote et al. 2012; LaManna et al. 2013; Merchant et al. 2014; Pirota et al. 2014).

iii. **Incorrect scoring due to incomplete records of behavioural responses**

To assign a behavioural response severity score, one must first detect a behavioural response. However, very few studies would be capable of measuring every possible behavioural response parameter. For example, the majority of studies reported observable behavioural disruptions at the surface. However, behavioural responses may occur below the surface, when marine mammals are visually obscured from observers, and can persist beyond observational periods. Consequently, studies relying on bio-logging and telemetry may be less prone to potential biases and more accurate in documenting effects in these situations (e.g., Miller et al. 2012; Southall et al. 2012).

Similarly, studies lacking acoustic measurements may be producing behavioural response severity scores that are artificially very low. For example, Frankel and Clark (1998) found little if any change in humpback whale surface behaviour in response to LFAS, while later passive acoustic studies showed that humpback whales increased the

length of their songs in response to this sound source (Biassoni et al. 2000; Miller et al. 2000; Fristrup et al. 2003). Other similar examples include the responses of sperm whales (*Physeter microcephalus*) exposed to airgun arrays (Miller et al. 2009) and bottlenose dolphins (*Tursiops truncatus*) exposed to motorboats and trawlers (La Manna et al. 2013).

Another consideration is that the spatial and temporal scales over which marine mammals exhibit behavioural responses to sound exposures extend beyond the range of detectability and monitoring of many studies, potentially generating false negative results. For instance, sounds from seismic airguns during seismic surveys become very important components of non-impulsive, low-frequency sounds that can propagate over more than 4000 km from their source (McDonald et al. 1995; Niekirk et al. 2004, 2012). These sounds are likely detected by baleen whales, and may overlap with sounds they use for communicating over long distances (Clark 1990; Clark and Gagnon 2006). This is illustrated by a study where low-frequency sounds generated by an Ocean Acoustic Waveguide Remote Sensing experiment source, and recorded 200 km from the source, altered the acoustic behaviour of humpback whales (Risch et al. 2012 but see Gong et al. 2014). Effects can also extend beyond the period of monitoring, as illustrated in a study of effects from airgun low-frequency sounds and showing displacements that extended over a period longer than the period of exposure in fin whales (*Balaenoptera physalus*, Castellote et al. 2012). When considering these reports, it is very possible that the geographical and temporal scale at which marine mammals are being exposed to noise is significantly larger in comparison to the geographical and temporal scale monitored during the vast majority of behavioural studies available (see Clark and

Gagnon 2006). In the case where there is a mismatch between the scale of monitoring and exposure, the severity of response assigned to a data case could be biased downward. Similar scale-related biases may arise from studies of LFAS where, like seismic surveys, LFAS exposures can last several hours or days, and can occur several times within a year.

Matters are complicated further by potential coping mechanisms, such as animals actively reducing their noise exposure by 'hiding' in near-surface or near-source acoustic shadows (McCauley et al. 2000; Madsen et al. 2006), exploiting the complex, non-linear propagation patterns of underwater sounds (Madsen et al. 2006). The risk here is that 'severe' responses such as avoidance could be incorrectly interpreted as a much more subtle behavioural response. Accordingly, the Southall et al (2007) severity scale cannot always be applied with great confidence to what appears to be responses of low-severity.

iv. Difficulties when attempting to interpret biological significance of responses

To assign a score under the Southall et al. (2007) severity scale, it is necessary to identify the behaviour being altered and determine the extent to which it has changed. While the latter may require subjective decisions over what is "minor" and "moderate" the former can be more troublesome. For example, humpback whales alter their vocalisations in the presence of noise (Biassoni et al. 2000; Miller et al. 2000; Croll et al. 2001; Risch et al. 2012; Cerchio et al. 2014). While humpback whale vocalisations are known to be associated with reproduction, the extent to which an acoustic behavioural

response confers notable fitness consequences is unknown. Accordingly, cessation in the vocalizations could be considered to be a “minor” change in behaviour or a “major” change by disrupting reproduction. Other such examples include: longer blue whale calls in the presence of seismic (sparker) exploration (Di Iorio and Clark 2010); cessation of echolocation and shortened dive duration in beaked whales during exposures to vessel sound (Aguilar Soto et al. 2006); reduction or cessation of beaked whale foraging vocalizations in the presence of MFAS (Tyack et al. 2011); changes in the amplitude and rate of blue whale calls in the presence of ships (McKenna 2011); and a reduction in sperm whales buzz rates during exposure to airgun sounds (Miller et al. 2009).

In this analysis, we assigned a score of 4 and 5 to those types of acoustic behavioural responses (i.e., moderate and prolonged modification of vocalizations). However, such changes could have implications for critical activities such as foraging, socializing, and breeding display (Cerchio et al. 2014; Pirotta et al. 2014). As such, these changes might have instead been scored higher as moderate or prolonged cessation of breeding (severity 7 or 8) (Southall et al. 2007); or moderate and prolonged cessations of feeding/resting (severity 6 or 7) (per revisions by Ellison et al. 2012). Further difficulties arise in scoring changes to vocalisations that have not yet been ascribed an ecological significance, such as the majority of sounds produced by baleen whales. Strict adherence to Southall et al. (2007) would score changes in vocalisations as 4 to 6, while a score of 7 or 8 might be recorded by application of a more precautionary approach. Pilot whales proved particularly difficult to score since they imitated the sounds of, and moved towards, operating MFAS (Alves et al. 2014).

2. Inconsistent measurement, estimation, and reporting of RL

Although strong recommendations have been made repeatedly in the literature to appropriately and consistently measure, estimate, and report acoustic metrics (Richardson et al. 1995; Madsen 2005; Merchant et al. 2015) few studies have implemented this. This lack of reporting is reflected in the absence of information associated with RL units in some examples provided in this manuscript (see Discussion), such as if they arose from bandwidth, peak-to-peak or rms measurements. To allow for some level of comparison, we used the maximum RL_{SPL} in the analysis because it was the metric most often reported across studies, although it may not necessarily be the most appropriate response metric (Madsen et al. 2006; Miller et al. 2014; see also Section I in S3). For example, cumulative sound exposure level (CSEL, which integrates acoustic energy over time) could be a more relevant metric than SPL (Madsen 2005; Miller et al. 2012; Tougaard et al. 2015). In addition, it has been suggested that sound levels relative to a marine mammal's hearing threshold (sensation levels, MacGillivray et al. 2014) based on audiogram-specific hearing weightings may provide a more appropriate metric (Ellison et al. 2012). Additional acoustic metrics, such as the signal-to-noise ratio (i.e., the relative level of the sound signal in comparison with the background noise), should also be considered potentially relevant to responses (Dunlop et al. 2013). Unfortunately, such details are rarely presented in the current literature, if the information is available at all. Accordingly, many assessments have used the generalised auditory criteria that were tested in this review (e.g., LGL 2010; JASCO 2015; NOAA 2015; Statoil 2015). This functional hearing group approach, however, is not accepted nor applied universally in assessments (Tougaard et al. 2015)

and recent evidence suggests that the hearing mechanisms of marine mammals are not fully understood (Berns et al. 2015).

As a consequence, the lack of information and lack of measurement standardization may be underpinning our results that demonstrated that RL is an uninformative variable for predicting behavioural response. However, we do not believe this to be always the case as there are also several individual studies that have found no link between RL and behavioural responses. For example, Miller et al. (2000) found that the increase in humpback whale song length during exposure to LFAS transmissions (RL: 150 dB re 1 μ Pa rms) did not correlate with maximum RL. Similarly, behavioural responses of one gray whale (i.e., decreased blow interval, length of surfacing, and length of dive) exposed to drillship playback seemed to be more pronounced when exposed to lower, rather than higher, RLs (Malme et al. 1986).

3. Context of exposure and natural variation: it's complicated

Details about important contextual variables such as the behaviour of individuals prior to sound exposure, etc. were compiled during this review when available (see S2), but could not be included in the analysis. This was due to limited sample sizes that were partly a function of the high variability in the definitions employed in the reviewed studies to differentiate between a state, behaviour, or activity - if they were presented at all (Figure 3, Table 5). However, these studies did show that, for example, animals engaged in foraging may be less overtly responsive than those travelling, presumably due to the benefits of pursuing their activity in comparison to the costs of interrupting this behaviour through displacement (e.g., Miller et al. 2009). This tendency has been

reported in various species responding to different acoustic sources (e.g., dolphins, Richardson et al. 1995; manatees, Miksis-Old and Wagner 2012), but it has been perhaps most extensively documented in bowhead whales (Richardson et al. 1995; Koski et al. 2009; Robertson et al. 2013; Robertson 2014). This is reflected in the data cases gathered for baleen whales exposed to seismic; however this is not evident for other species and noise sources (see Data exploration 2 - Section IV in S3, and see plots in S2 when selecting foraging and traveling in the 'behaviour prior to sound exposure' filter). Killer whales and pilot whales (*Globicephala melas*), for example, ceased deep foraging dives when exposed to playbacks of active sonar in the range between 1-2 kHz (see Table 3) while individuals conducting non-feeding shallow dives (likely traveling or resting) did not alter their dive behaviour during sound exposure (Sivle et al. 2012). Similarly, blue whales that were either feeding in deep water, or not feeding, ceased their behaviour during controlled exposures to simulated MFAS, whereas individuals that were foraging closer to the surface typically did not change their behaviour (Goldbogen et al. 2013).

Complicating matters further, behavioural responses during playback experiments also differed from those responses observed when exposed to a real operating source at the same output level. For example, bowhead whales exposed to seismic activities (RL: 107-158 dB re 1 μ Pa) did not exhibit avoidance nor changes in their surface behaviour, while individuals exposed to experimentally-controlled airgun sounds (RL: 124-134 dB re 1 μ Pa) moved away from the area (Richardson et al. 1986). Similar results were reported for Cuvier's beaked whales (*Ziphius cavirostris*), which strongly avoided playbacks of MFAS at distances of 10 km, but reacted much less severely to naval

sonar operating 118 km away, despite similar RLs (DeRuiter et al. 2013a). In contrast, Blainville's beaked whales (*Mesoplodon densirostris*) ceased echolocating during deep foraging dives and moved away during both real and simulation playbacks of MFAS (Tyack et al. 2011). In contrast to earlier studies that in some cases used playbacks of recorded noise-generating activities (e.g., Richardson et al. 1990, 1995) recent behavioural response controlled studies use real operating sound sources (e.g., Miller et al. 2012). Many of these controlled experiments can last a few hours in a day while real operations can last for several days and can occur multiple times in a year. Behavioural response studies conducted during real operations in terms of time and sound source will ultimately provide the most realistic conditions of sound exposure.

It has been suggested that, at least on some occasions, the distance from the sound source might be important, regardless of the RL (DeRuiter et al. 2013a). Some support for this can be found in a study of sperm whales that found buzz production rates were linked more to the distance to a seismic source rather than RLs, which were higher with the source between 5 to 12.6 km than they were at 2 km (Madsen et al. 2006).

Therefore, distance from the sound source, in addition to RL, should be considered.

In addition to the variation in responses dependent on the behaviour of individuals prior to exposure and distance from the sound source, reactions may also vary by gender, age, sexual condition, individuals, and other physical factors such as habitat use. For example, groups of humpback whale females with calves consistently avoided a single operating airgun (RL: up to 143 dB re 1 μ Pa rms), while in contrast male humpbacks were attracted to it at higher RLs (RL: 179 dB re 1 μ Pa rms) (McCauley et al. 2000).

Similarly, sighting rates of female bottlenose dolphins decreased over eight years of bridge construction while sighting rates of males remained unchanged (Weaver 2015). Female killer whales seemed to be more likely than males to respond to the passage of a ship (Williams et al. 2014). Irrespective of gender, there are also inter-individual differences in the amount of cumulative time that dolphins spend in the presence of boats (Pirotta et al. 2015).

Animals' previous experience with particular sounds is another factor. For example, some of the most severe behavioural responses (scores 8 and 9) found in this review were for naïve, wild belugas (*Delphinapterus leucas*) in the Canadian high Arctic, responding to some of the lowest RLs reported: 90-120 dB re 1 μ Pa rms (LGL and Greeneridge 1986 in Southall et al. 2007; Finley et al. 1990). In contrast, beluga in the St. Lawrence Estuary generally exhibit much less severe behavioural responses to passing vessels (Lesage et al. 1999), suggesting that they may be more tolerant to vessels (Wartzok et al. 2004). Belugas in the St. Lawrence are not naïve to shipping noise as they live in an area with a high proportion of vessels with a broadband SPL rms exceeding 120 dB re 1 μ Pa for about 8-32% of the time, depending on the site (McQuinn et al. 2011). However, despite the differences in the magnitude of responses at each study area, behaviour of belugas at both sites was changed significantly compared with baseline conditions. Similarly, blue whales in feeding areas off California have altered their diving and acoustic behaviour during simulated MFAS playback, in some cases at very low RLs: 130 dB re 1 μ Pa (Melcon et al. 2012; Goldbogen et al. 2013). This area is subject to a high level of naval activity so these individuals were likely not naïve to MFAS sounds (Goldbogen et al. 2013).

Other contextual factors are likely environmental, which can make them harder to ascertain. For example, travelling bowhead whales avoided drilling operations when they were not confined by ice (RL: 120 dB re 1 μ Pa); but exhibited tolerance of higher RLs (RL: up to 135 dB re 1 μ Pa) when there was no alternative route through the ice (Richardson et al. 1991a in Richardson et al. 1995). However, in most cases in this review, it is unclear what contextual metrics, or combinations thereof, triggered observed behavioural responses. For example, migrating humpback whales in Australia showed avoidance reactions to alarm sounds in one study (2.0-2.1 kHz, source level: 148-153 dB re 1 μ Pa, range: 0.3-2 km, Dunlop et al. 2013), but not in another study with similar contextual circumstances (3 kHz, source level: 135 dB re 1 μ Pa, range: up to 500m, Harcourt et al. 2014). Similarly, variability in responses of foraging sperm whales exposed to seismic surveys has been seen in Norway (Madsen et al. 2002; Miller et al. 2009) and the Gulf of Mexico (Mate et al. 1994). Other examples can be seen in the between-individual response variability in gray whales and humpback whales responding to airgun sound exposures (Malme et al. 1985; Malme et al. 1986; McCauley et al. 2000).

Given that this review represents a thorough assessment, it is concerning that we were unable to incorporate any statistical assessment of context due to the specificities and lack of standardization across studies (S2, Figure 3). Addressing this issue would require a comprehensive global research programme to assess acoustic effects in a comparative way across numerous geographic locations, species, ages, sexes, reproductive states and environmental conditions to generate the data required for a generalised model. Accordingly, it may remain impossible to incorporate context into a

response severity scale in anything more than the most general manner (e.g., this is a migration route, or a feeding area), at least for the foreseeable future until more well-designed and standardized studies are undertaken (see Erbe et al. 2016 for a summary of the protocols and standards necessary for research on noise impacts on marine life).

Summary, final remarks, and recommendations

This review provides a synthesis of the many behavioural responses of marine mammals to noise. The results of the analysis showed that the severity of the behavioural responses is different for baleen whales and MF toothed cetaceans, depending on the type of sound source to which they are exposed. However, incorporating the intensity of the sound received by the individuals or RL, did not improve model predictions of the severity of the behavioural responses of whales and dolphins to noise. Context of exposure plays a critical and complex role in modulating severity of behavioural response, however, it was impossible to include contextual variables in the analysis due to the very small sample sizes, the specificities, and lack of standardization across the many studies reviewed.

It is clear from this review that a **generalised** predictive model of acoustic behavioural impact, mediated via behavioural responses (i.e., via a linear severity scale), based upon RL and context, is challenging at best, and currently unfeasible with the datasets available in the published literature. This is because the studies we reviewed, which represent the best available science on the effects of noise on the behaviour of marine mammals to date, while aggregated and analysed in detail, did not provide complete,

clear and consistent information on pressure/response relationships and related contextual information.

While the findings of this systematic review should not preclude a species-by-species approach to management when sufficient data is available (Williams et al. 2014; Tougaard et al. 2015), they do suggest that an alternative approach is required for making more generalised or multi-species management decisions (Dekeling et al. 2014). As a starting point to developing an alternative framework, we recommend substituting the behavioural response severity score with an approach focused on dichotomous measures of effects, conveyed through observed behavioural responses:

- i. physical avoidance (e.g., were individuals displaced, if yes, for how long and how far from the noise source);
- ii. acoustic behavioural change (e.g., did individuals change their vocal behaviour (Table 8), if yes, for how long and how far from the noise source), which will include Lombard effect responses to masking.

Both of these can represent measures of impact in terms of habitat loss and degradation of either physical habitat or acoustic habitat, and are relatively easy to measure (e.g., based on direct field measurements). For example, during a seismic survey, a proportion of the population of harbour porpoises was displaced by 5 to 10 km, although the level of displacement declined through the 10 day survey; harbour porpoises remaining in the impact area reduced their vocal behaviour by 15% for 10 days (Thompson et al. 2013; Pirodda et al. 2014). Assessing vocal behaviour changes provides insight into potential sub-lethal effects on the individuals that remain within the

disturbed area (e.g., alteration of foraging performance, Thompson et al. 2013, see Williams et al. 2014). Importantly, when including all data cases (with and without RL reported) for studies that measured acoustic behavioural changes, 90% reported a change in acoustic behaviour while 10% did not report acoustic behavioural change (see Data exploration 1 - Section IV in S3). This strongly supports our contention that it is critical to include acoustic behavioural change in studies that aim at evaluating behavioural responses of marine mammals to noise.

The literature compiled in this review can be used to continue the process of summarizing information on behavioural responses of marine mammals to noise. Future studies could follow an approach similar to those used by Thompson et al. (2013) and Pirotta et al. (2014) to develop specific look up tables that provide the best available or precautionary “distance of effect” for a given geographic location, species, sound source and context of exposure. As in those studies, we recommend that behavioural responses of marine mammals to noise should be investigated under real conditions of operation.

In addition to using behavioural responses as indicators of habitat loss and degradation, a quantification of the potential loss of acoustic communication space can also be incorporated simply into this framework (Møhl, 1980, 1981; Clark et al. 2009; Hatch et al. 2012; Moore et al. 2012; Erbe 2015) as a proxy for additional sub-lethal impacts such as masking and/or stress responses (zone 2 in Figure 1). Methods are currently being refined to incorporate various noise effects into population consequences of disturbance (PCoD) models (e.g., King et al. 2015) and new drone technology is starting

to allow scientist to monitor not simply changes in behaviour due to anthropogenic stressors but changes in body condition and health (Smith et al. 2016).

Monitoring and regulating the effect of noise on the behaviour of marine mammal populations should be based on variables that are associated with potential effects on foraging, socializing, reproduction, or survival. The North American current paradigm used to evaluate and regulate the effects of noise on the behaviour of marine mammals relies on establishing allowable generic multi-species RL thresholds at which individuals are predicted to display significant behavioural responses (often termed harassment). The results presented in this review summarize what many individual studies have also concluded: the monitoring and regulation of sub-lethal impacts of noise on cetaceans should not rely entirely and solely upon generic RL thresholds. As a result, regulators continuing to base assessments of behavioural and other sub-lethal impacts of noise exposure on marine mammals solely on generic RL are failing to properly evaluate these impacts, and thus also fall short of properly managing and mitigating these effects.

Acknowledgements

C. Gomez was supported by a post-doctoral Visiting Fellowship managed via the Natural Sciences and Engineering Research Council (NSERC) of Canada, funded by the Strategic Program for Ecosystem-Based Research and Advice (SPERA) at Fisheries and Oceans Canada. This review benefitted from discussions during two international workshops conducted as part of this project (also funded by SPERA); in particular we thank Mark Tasker, Hilary Moors-Murphy, Gary Stenson, Karin Forney,

John Hildebrand, Christopher Clark, Christine Erbe, Jason Wood, Rob Williams, and Joe Crocker. We thank Christopher Daley for his detailed and helpful comments on the manuscript and we thank Paul Regular for helping in the development of the interactive supplementary material. We also thank the anonymous reviewers for their comments and suggestions to improve the quality of this manuscript.

References

- Agresti, A. 2002. *Categorical Data Analysis*. John Wiley and Sons, New York, NY.
- Aguilar Soto, N., Johnson, M., Madsen, P., Tyack, P., Boconcelli, A., and Borsani, F. 2006. Does Intense Ship Noise Disrupt Foraging in Deep-Diving Cuvier's Beaked Whales (*Ziphius cavirostris*)? *Mar. Mamm. Sci.* **22** (3): 690–99.
- Alves, A., Antunes, R., Bird, A. Tyack, P. Miller, P. Lam, F., and Kvadsheim, P. 2014. Vocal Matching of Naval Sonar Signals by Long-Finned Pilot Whales (*Globicephala melas*). *Mar. Mamm. Sci.* **30** (3): 1248–57. doi:10.1111/mms.12099.
- Anderson, D. 2008. *Model Based Inference in the Life Sciences: A Primer on Evidence*. Springer Science+Business Media, LLC, New York, NY.
- Antunes, R., Kvadsheim, P., Lam, F., Tyack, P., Thomas, L., Wensveen, P., and Miller, P. 2014. High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (*Globicephala melas*). *Mar. Poll. Bull.* **83** (1): 165–80.
- Azzara, A.J., Wyndylyn, M., von Zharen, M., and Newcomb, J. 2013. Mixed-Methods analytic approach for determining potential impacts of vessel noise on sperm whale click behavior. *J. Acoust. Soc. Am.* **134** (6): 4566–74.

<http://link.aip.org/link/JASMAN/v134/i6/p4566/s1&Agg=doi>.

- Balcomb, K., and Claridge, D. 2001. A mass stranding of cetaceans caused by Naval sonar in the Bahamas. *Bahamas J. Sci.* **01** (05): 2–12.
- Bejder, L., Samuels, A., Whitehead, H., Finn, H., and Allen, S. 2009. Impact Assessment Research: Use and Misuse of Habituation, Sensitisation and Tolerance in Describing Wildlife Responses to Anthropogenic Stimuli. *Mar. Ecol. Prog. Ser.* **395**: 177–85.
- Bejder, L., Samuels, A., Whitehead, H., Gales, N., Mann, J. Connor, R., Heithaus, M., Watson-Capps, J., Flaherty, C., and Krützen, M. 2006a. Decline in relative abundance of bottlenose dolphins exposed to long-term disturbance. *Conserv. Biol.* **20** (6): 1791–98.
- Bejder, L., Samuels, A., Whitehead, H., and Gales, N.. 2006b. Interpreting short-term behavioural responses within a longitudinal perspective. *Anim. Behav.* **72** (5): 1149-1158.
- Berns, G., Cook, P., Foxley, S., Jbabdi, S., Miller, K., and Marino, L. 2015. Diffusion tensor imaging of dolphin brains reveals direct auditory pathway to temporal lobe. *Proc. R. Soc. Lond. B. Biol. Sci.* **282** (1811): 20151203.
doi:<http://dx.doi.org/10.1098/rspb.2015.1203>.
- Biassoni, N., Miller, P., and Tyack, P. 2000. Preliminary results of the effects of SURTASS-LFA sonar on singing humpback whales. *WHOI Technical Report*. Vol. 6. Available from <http://nova.wh.who.edu/palit/Biassoni, Miller>,

Tyack 2000 Unknown Preliminary results of the effects of SURTASS-LFA Sonar on singing humpback whales [accessed January 20 2016].

Blackwell, S., Nations, C., McDonald, T., Thode, A., Mathias, D. Kim, K., Greene, C., and Macrander, M. 2015. Effects of airgun sounds on bowhead whale calling rates: evidence for two behavioral thresholds. PLoS ONE **10** (6): 1-29.

doi:10.1371/journal.pone.0125720

Buckstaff, K. 2004. Effects of watercraft noise on the acoustic behaviour of bottlenose dolphins in Sarasota Bay, Florida. Mar. Mamm. Sci. **20** (4): 709–25.

doi:10.1111/j.1748-7692.2004.tb01189.x.

Burnham, K., and Anderson, D. 2002. Model selection and multimodel inference: A practical information-theoretic approach. Springer-Verlag, New York, NY.

Burnham, K., Anderson, D., and Huyvaert, K. 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons, Behavioral Ecology and Sociobiology **65**: 23–35.

Castellote, M., Clark, C., and Lammers, M. 2012. Acoustic and behavioural changes by fin whales (*Balaenoptera physalus*) in response to shipping and airgun noise. Biol. Conserv. **147** (1): 115–22.

Cerchio, S, Strindberg, S., Collins, T., Bennett, C., and Rosenbaum, H. 2014. Seismic surveys negatively affect humpback whale singing activity off Northern Angola. PLoS ONE **9** (3): e86464. doi:10.1371/journal.pone.0086464

Clark, C., and Gagnon, G. 2006. Considering the temporal and spatial scales of noise

- exposures from seismic surveys on baleen whales. (IWC/SC/58/E9). Submitted to Scientific Committee, International Whaling Commission. 9pp. Available from http://www.bluewhalealliance.com/Clark_and_Gagnon.pdf [accessed January 20 2016]
- Clark, C., William E., Southall, B. Hatch, L. Van Parijs, S., Frankel, A., and Ponirakis, D. 2009. Acoustic masking in marine ecosystems: intuitions, analysis, and implication. *Mar. Ecol. Prog. Ser.* **395**: 201–22.
- Clark, C. 1990. Acoustic behavior of mysticete whales. *In* *Sensory Systems of Aquatic Mammals*. Edited by J. Thomas and R. Kastelein, Plenum Press. pp. 571– 583.
- Croll, D., Clark, C., Calambokidis, J., Ellison, W., and Tershy, B. 2001. Effect of anthropogenic low-frequency noise on the foraging ecology of *Balaenoptera* whales. *Anim. Conserv.* **4** (01): 13-27. doi: <http://dx.doi.org/10.1017/S1367943001001020>
- Dahlheim, M. 1987. Bioacoustics of the Gray whale (*Eschrichtius robustus*). Ph.D. Thesis, University of British Columbia. 266 pp.
- Danil, K., and St. Leger, J.A.. 2011. Seabird and dolphin mortality associated with underwater detonations exercises. *Mar. Technol. Soc. J.* **45** (6): 89 -95.
- Dekeling, R.P.A., Tasker, M., Van der Graaf, A.J., Ainslie, M.A, Andersson, M.H. et al., 2014. Monitoring guidance for underwater noise in European Seas, Part I: Executive Summary, Luxembourg. 166 pp. Available from <http://www.nrcresearchpress.com/page/cjz/authors#9d> [accessed January 15

2016]

- DeRuiter, S., Southall, B., Calambokidis, J., Zimmer, W., Sadykova, D., Falcone, Friedlaender, A. Joseph, J., Moretti, D., Schorr, G., Thomas, L., Tyack P. 2013a. First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. *Biol. Lett.* **9** (4): 20130223.
- DeRuiter, S., Boyd, I., Claridge, D., Clark, C., Gagnon, C., Southall, B., and Tyack, P. 2013b. Delphinid whistle production and call matching during playback of simulated military sonar. *Mar. Mamm. Sci.* **29** (2): E46–E59.
- DFO. 2012. Science review of Baffinland's Mary River Project final environmental impact statement. DFO Can. Sci. Advis. Sec. Sci. Rep. No. 2012/016
- DFO. 2014. Impacts of rerouting marine traffic in the St. Lawrence estuary on beluga (*Delphinapterus leucas*): Science in support of risk management. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. No. 2014/004.
- Di Iorio, L., and Clark, C. 2010. Exposure to seismic survey alters blue whale acoustic communication. *Biol. Lett.* **6** (1): 51–54.
- Dooling, R., and Therrien, S. 2007. Hearing in birds: what changes from air to water. *In* the effects of noise on aquatic life. *Edited by* A. Popper and A. Hawkins. Springer, New York, NY. pp. 77–82.
- Doyle, L., McCowan, B., Hanser, S., Chyba, C. Bucci, T., and Blue, J. 2008. Applicability of information theory to the quantification of responses to anthropogenic noise by Southeast Alaskan humpback whales. *Entropy* **10** (2): 33–

46.

Dunlop, R., Noad, M., Cato, D., Kniest, E. Miller, P., Smith, J., and Stokes, M. 2013.

Multivariate analysis of behavioural response experiments in humpback whales

(*Megaptera novaeangliae*). *J. Exp. Biol.* **216** (5): 759–70. doi:10.1242/jeb.071498.

Ellison, W., Southall, B., Clark, C., and Frankel, A. 2012. A New context-based

approach to assess marine mammal behavioral responses to anthropogenic

sounds. *Conserv. Biol.* **26** (1): 21–28.

Erbe, C. 2015. The Maskogram: A tool to illustrate zones of masking. *Aquat. Mamm.* **41**

(4): 434–43. doi:10.1578/AM.41.4.2015.434.

Erbe, C., Ainslie, M., de Jong, C., Racca, R., and Stocker, M. 2016. Summary report

panel 1: The need for protocols and standards in research on underwater noise

impacts on marine life. *In* The Effects of Noise on Aquatic Life. *Edited by* A. Popper

and A. Hawkins. Springer, New York, NY. pp. 1265-1271.

Erbe, C., Reichmuth, C., Cunningham, K. Lucke, K., and Dooling, R. 2015.

Communication masking in marine mammals: a review and research strategy. *Mar.*

Poll. Bull., December. **103** (1-2): 15-38. doi:10.1016/j.marpolbul.2015.12.007. pp

1265-1271.

Finley, K., Miller, G., Davis, R., and Greene, C. 1990. Reactions of Belugas

Delphinapterus leucas and Narwhals *Monodon monoceros* to ice-breaking ships in

the Canadian High Arctic. *Can. Bull. Fish. Aquat. Sci.* **224**: 97–117.

Federal Register. 2003. Small takes of marine mammals incidental to specified

activities; Oceanographic surveys in the Hess Deep, Eastern Equatorial Pacific Ocean. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service. July 11, 2003. 68 FR 133, pp. 41314-41321. Available from <http://www.gpo.gov/fdsys/pkg/FR-2003-07-11/pdf/03-17622.pdf> [accessed March 27 2016]

Federal Register. 2009. Taking and Importing Marine Mammals; U.S. Navy Training in the Hawaii Range Complex. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service. January 12, 2009. 74 FR 7 pp. 1456-1491. Available from <http://www.gpo.gov/fdsys/pkg/FR-2009-01-12/pdf/E9-37.pdf> [accessed March 27 2016]

Federal Register. 2015. Taking and Importing Marine Mammals; U.S. Navy Training in the Hawaii Range Complex. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service. June 3, 2015. 74 FR 7 pp. 31738-31817. Available from <http://www.regulations.gov/#!documentDetail;D=NOAA-NMFS-2015-0031-0001> [accessed March 27 2016]

Foot, A., Osborne, R., and Hoelzel, R. 2004. Environment: whale-call response to masking boat noise. *Nature*, **428** (6986): 910.

Frankel, A., and Clark, C. 1998. Results of low-frequency playback of M-sequence noise to humpback whales, *Megaptera novaeangliae*, in Hawaii. *Can. J. Zool.* **76**: 521–535.

Fristrup, K., Hatch, L., and Clark, C. 2003. Variation in humpback whale (*Megaptera*

- novaeangliae*) song length in relation to low-frequency sound broadcasts. J. Acoust. Soc. Am. **113** (6): 3411–24.
- Goldbogen, J., Southall, B., DeRuiter, S., Calambokidis, J., Friedlaender, A., Hazen, E., Falcone, E., Schorr, G., Douglas, A., Moretti, D., Kyburg, C., McKenna, M., and Tyack, P. 2013. Blue whales respond to simulated mid-frequency military sonar. Lond. B Biol. Sci. **280** (1765): 20130657. doi: <http://dx.doi.org/10.1098/rspb.2013.0657>
- Gong, Z., Jain, A., Tran, D., Hoon Yi, D., Wu, F., Zorn, A., Ratilal, P., and Makris, N. 2014. Ecosystem scale acoustic sensing reveals humpback whale behavior synchronous with herring spawning processes and re-evaluation finds no effect of sonar on humpback song occurrence in the Gulf of Maine in fall 2006. PLoS ONE **9** (10): e104733. doi:10.1371/journal.pone.0104733
- Harcourt, R., Pirota, V., Heller, G., Peddemors, V., and Slip, D. 2014. A whale alarm fails to deter migrating humpback whales: an empirical test. Endanger. Species Res. **25** (1): 35–42. doi:10.3354/esr00614.
- Hatch, L., Clark, C., Van Parijs, S., Frankel, A., and Ponirakis, D. 2012. Quantifying loss of acoustic communication space for right whales in and around a U.S. National Marine Sanctuary. Conserv. Biol. **26** (6): 983–94.
- Heinis, F., de Jong, C., and Rijkswaterstaat Underwater Sound Working Group 2015. Framework for assessing ecological and cumulative effects of offshore wind farms: cumulative effects of impulsive underwater sound on marine mammals. TNO 2015 R10335-A. Available from <https://tethys.pnnl.gov/publications/framework->

[assessing-ecological-and-cumulative-effects-offshore-wind-farms-cumulative.](#)

[accessed on January 15 2016]

High Energy Seismic Survey (HESS). 1999. High energy seismic survey review process and interim operational guidelines for marine surveys offshore Southern California. Prepared for The California State Lands Commission and the Minerals Management Service Pacific Outer Continental Shelf Region. Camarillo, California: High Energy Seismic Survey Team.

Hildebrand, J. 2009. Anthropogenic and natural sources of ambient noise in the ocean. *Mar. Ecol. Prog. Ser.* **395**: 5–20. doi:10.3354/meps08353.

Holt, M., Noren, D., Veirs, V., Emmons, C., and Veirs, S. 2009. Speaking up: killer whales (*Orcinus orca*) increase their call amplitude in response to vessel noise. *J. Acoust. Soc. Am.* **125** (1): EL27–L32.

Holt, M., Noren, D., Dunkin, D., and Williams, T. 2015. Vocal performance affects metabolic rate in dolphins: implications for animals communicating in noisy environments. *J. Exp. Biol.* **218**: 1647–1654. doi: 10.1242/jeb.122424

Holt, M., Noren, D., and Emmons, C. 2011. Effects of noise levels and call types on the source levels of killer whale calls. *J. Acoust. Soc. Am.* **130** (5): 3100.

Hotchkin, C, and Parks, S. 2013. The Lombard effect and other noise-induced vocal modifications: insight from mammalian communication systems. *Camb. Philos. Soc.* **88** (4): 809–24.

Houser, D., Martin, S., and Finneran, J. 2013a. Behavioral responses of California sea

lions to mid-frequency (3250-3450 Hz) sonar signals. *Mar. Environ. Res.* **92**: 268–78. <http://www.ncbi.nlm.nih.gov/pubmed/24183102>.

Houser, D., Martin, S., and Finneran, J. 2013b. Exposure amplitude and repetition affect bottlenose dolphin behavioral responses to simulated mid-frequency sonar signals. *J. Exp. Mar. Biol. Ecol.* **443**: 123–33.

JASCO 2015. JASCO Applied Sciences Underwater Acoustic Assessment of Marine Terminal Construction Rio Tinto Alcan Terminal A Extension Project, Kitimat BC. 61pp. Available from https://a100.gov.bc.ca/appsdata/epic/documents/p414/1439404447930_RK7nVLR_NZl1FFXLTZTc2hB5tCHZGpSX9F7gpB46ztgPLp1C0j5ZJ!-1966804633!1439404333004.pdf [accessed on March 26 2016]

Ketten, D., Lien, J., and Todd, S. 1993. Blast injury in humpback whale ears: evidence and implications. *J. Acoust. Soc. Am.* **94**: 1849–50.

Ketten, D. 2014. Sonars and strandings: are beaked whales the aquatic acoustic canary? *Acoustics Today* **10**: 46–56.

Ketten, D. 1995. Estimates of blast injury and acoustic trauma zones for marine mammals from underwater explosions. *In* *Sensory Systems of Aquatic Mammals*. Edited by R. Kastelein, J. Thomas, and P. Nachtigal. Woerdwen, The Netherlands: De Spil. pp. 391–407.

King, S., Schick, R., Donovan, C., Booth, C., Burgman, M., Thomas, L., and Harwood, J. 2015. An interim framework for assessing the population consequences of

disturbance. *Methods. Ecol. Evol.* **6**: 1150–1158. doi: 10.1111/2041-210X.12411

Koski, W., Funk, W., Ireland, D., Lyons, C., Macrander, A., and Blackwell, S. 2009. An update on feeding by bowhead whales near an offshore seismic surveys in the Central Beaufort Sea. SC/61/BRG3. Available from [http://www.north-slope.org/assets/images/uploads/feeding_bowhead whales near Offshore Seismic in Beaufort Sea.pdf](http://www.north-slope.org/assets/images/uploads/feeding_bowhead_whales_near_Offshore_Seismic_in_Beaufort_Sea.pdf) [accessed on January 15 2016]

La Manna, G., Manghi, M., Pavan, G., Lo Mascolo, F., and Sarà, G. 2013. Behavioural strategy of common bottlenose dolphins (*Tursiops truncatus*) in response to different kinds of boats in the waters of Lampedusa Island (Italy). *Aquat. Conserv.* **23** (5): 745–57. doi:10.1002/aqc.2355.

Lesage, V., Barrette, C., Kingsley, M., and Sjare, B. 1999. The effect of vessel noise on the vocal behavior of belugas in the St. Lawrence river estuary, Canada. *Mar. Mamm. Sci.* **15** (1): 65–84. doi:10.1111/j.1748-7692.1999.tb00782.x.

LGL 2010. A Conceptual Framework for Tiered Risk Assessment to Evaluate the Effects Of Sound from E&P Operations on Marine Mammals. 120 pp. Available from <http://www.soundandmarinelife.org/research-categories/behavioural-reactions-and-biological-significant-effects/application-of-risk-assessment.aspx> [accessed January 15 2016]

Ljungblad, D., Wursig, B., Swartz, S., and Keene, J. 1988. Observations on the behavioral responses of bowhead whales (*Balaena mysticetus*) to active geophysical vessels in the Alaskan Beaufort Sea. *Arctic* **41** (3): 183-194. <http://dx.doi.org/10.14430/arctic1717>

- Lusseau, D., Bain, D., Williams, R., and Smith, J. 2009. Vessel traffic disrupts the foraging behavior of southern resident killer whales *Orcinus orca*. *Endanger. Species Res.* **6** (3): 211–21.
- MacGillivray, A., Racca, R., and Li, Z. 2014. Marine mammal audibility of selected shallow-water survey sources. *J. Acoust. Soc. Am.* **135** (1): EL35–40.
doi:10.1121/1.4838296.
- Madsen, P. 2005. Marine mammals and noise: problems with root mean square sound pressure levels for transients. *J. Acoust. Soc. Am.* **117** (6): 3952–57.
- Madsen, P., Johnson, M., Miller, P., Aguilar Soto, N., Lynch, L., and Tyack, P. 2006. Quantitative measures of air-gun pulses recorded on sperm whales (*Physeter macrocephalus*) using acoustic tags during controlled exposure experiments. *J. Acoust. Soc. Am.* **120** (4): 2366–79.
- Madsen, P., Møhl, B., Nielsen, B., and Wahlberg, W. 2002. Male sperm whale behaviour during exposures to distant seismic survey pulses. *Aquat. Mamm.* **28** (3): 231–40.
- Malme, C., Miles, P., Clark, C., Tyack, P., and Bird, J. 1983. Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behaviour. BBN Report 5366, Bolt Beranek & Newman Inc., Cambridge, MA, for U.S. Minerals Management Service, Anchorage, AK. NTIS PB86-174174. Available from <http://www.boem.gov/BOEM-Newsroom/Library/Publications/1983/rpt5366.aspx>. [accessed January 15 2016]

- Malme, C., Miles, P., Clark, C., Tyack, P., and Bird, J. 1984. Investigations of the potential effects of underwater noise from petroleum industry activities on migrating Gray whale behaviour. Phase II. BBN Report 5586, Bolt Beranek & Newman Inc., Cambridge, MA, for U.S. Minerals Management Service, Anchorage, AK. NTIS PB86-218377. 358 pp. Available from <http://www.boem.gov/BOEM-Newsroom/Library/Publications/1983/rpt5586.aspx>. [accessed January 15 2016]
- Malme, C., Miles, P., Tyack, Clark, C., and Bird, J. 1985. Investigation of the potential effects of underwater noise from petroleum industry activities on feeding humpback whale behaviour. BBN Report 5586, Bolt Beranek & Newman Inc., Cambridge, MA, for U.S. Minerals Management Service, Anchorage, AK. Report to U-S. Department of the Interior. Anchorage, AK. Available from http://www.boem.gov/BOEM-Newsroom/Library/Publications/1985/85_0019.aspx. [accessed January 15 2016]
- Malme, C., Würsig, B., Bird, J., and Tyack, P. 1986. Behavioral responses of gray whales to industrial noise: feeding observations and predictive modeling. Outer Continental Shelf Environmental Assessment Program, Final Report of Principal Investigators. 207 pp. Available from <http://catalog.northslope.org/catalogs/4136-behavioral-responses-of-gray-whales-to-industri>. [accessed January 15 2016]
- Mate, B., Stafford, K., and Ljungblad, D. 1994. A Change in sperm whale (*Physeter macrocephalus*) distribution correlated to seismic surveys in the Gulf of Mexico. J. Acoust. Soc. Am. **96** (5):3268. Acoustical Society of America: 3268. doi: <http://dx.doi.org/10.1121/1.410971>
- McCarthy, E., Moretti, D., Thomas, L., DiMarzio, N., Morrissey, R., Jarvis, S., Ward, J.,

- Izzi, A., and Dilley, A. 2011. Changes in spatial and temporal distribution and vocal behavior of Blainville's beaked whales (*Mesoplodon densirostris*) during multiship exercises with mid-frequency sonar. *Mar. Mamm. Sci.* **27** (3): E206-226.
- McCauley, R., Fewtrell, D., Duncan, A., Jenner, C., Jenner, M., Penrose, J., Prince, R., Adhitya, A., Murdoch, J., and McCabe, K. 2000. Marine seismic surveys - a study of environmental implications. *APPEA J.*, **40**: 692–708.
- McDonald, M., Hildebrand, J., and Webb, S. 1995. Blue and fin whales observed on a seafloor array in the Northeast Pacific. *J. Acoust. Soc. Am.* **98** (2 Pt 1): 712-721.
- McKenna, M. 2011. Blue whale response to underwater noise from commercial ships. PhD thesis, University of California. 242 pp. Available from <http://gradworks.umi.com/34/57/3457564.html>. [accessed January 15 2016]
- McQuinn, I., Lesage, V., Carrier, D., Larrivée, G., Samson, Y., Chartrand, S., Michaud, R., and Theriault, J. 2011. A threatened beluga (*Delphinapterus leucas*) population in the traffic lane: vessel-generated noise characteristics of the Saguenay-St. Lawrence Marine Park, Canada. *J. Acoust. Soc. Am.* **130** (6): 3661-3673.
- Melcón, M., Cummins, A., Kerosky, S., Roche, L., Wiggins, S., and Hildebrand, J. 2012. Blue whales respond to anthropogenic noise. *PLoS ONE* **7** (2): e32681. doi:10.1371/journal.pone.0032681.
- Merchant, N., Pirotta, E., Barton, T., and Thompson, P. 2014. Monitoring Ship noise to assess the impact of coastal developments on marine mammals. *Mar. Poll. Bull.* **78** (1-2): 85–95.

- Miksis-Olds, J., and Tyack, P. 2009. Manatee (*Trichechus manatus*) vocalization usage in relation to environmental noise levels. *J. Acoust. Soc. Am.* **125** (3): 1806–15. doi:10.1121/1.3068455.
- Miller, P., Biassoni, N., Samuels, A., and Tyack, P. 2000. Whale songs lengthen in response to sonar. *Nature* **405** (6789): 903.
- Miller, P., Johnson, P., Madsen, P., Biassoni, N., Quero, N., and Tyack, P. 2009. Using at-sea experiments to study the effects of airguns on the foraging behavior of sperm whales in the Gulf of Mexico. *Deep-Sea. Res. Part 1 Oceanogr. Res. Pap.* **56** (7): 1168–81.
- Miller, P., Antunes, R., Alves, C., Wensveen, P., Kvadsheim, P., Kleivane, L., Nordlund, N., Lam, F., van Ijsselmuide, S., Visser, F., and Tyack, P. 2011. The 3S experiments: studying the behavioural effects of naval sonar on killer whales (*Orcinus orca*), sperm whales (*Physeter macrocephalus*) and long-finned pilot whales (*Globicephala melas*) in Norwegian water. SOI technical report, Scotland. 290 pp.
- Miller, P., Kvadsheim, P., Lam, F., Wensveen, P., Antunes, R., Alves, C., Visser, F., Kleivane, L., Tyack, P., and Sivle, L. 2012. The severity of behavioral changes observed during experimental exposures of killer (*Orcinus orca*), long-finned pilot (*Globicephala melas*), and sperm (*Physeter macrocephalus*) whales to naval sonar. *Aquat. Mamm.* **38** (4): 362-401. Available from <http://soi.st-andrews.ac.uk/documents/424.pdf> [accessed January 15 2016]
- Miller, P., Antunes, R., Wensveen, P., Samarra, F., Alves, A., Tyack, P., Kvadsheim, P.,

- Kleivane, L., Lam, P., Ainslie, M., and Thomas, L. 2014. Dose-response relationships for the onset of avoidance of sonar by free-ranging killer whales. *J. Acoust. Soc. Am.* **135** (2): 975–93.
- Møhl, B. 1980. Marine mammals and noise. *Arctic Seas Bull.* **2**, 1–2.
- Møhl, B. 1981. Masking effects of noise; their distribution in time and space. The question of sound from icebreaker operations. *In Proceedings of a Workshop: The Question of Sound from Icebreaker Operations*, Toronto, 23–24 February 1981. Arctic Pilot Project, Toronto, pp. 259–268.
- Moore, S., and Clarke, J. 2002. Potential impact of offshore human activities on gray whales (*Eschrichtius robustus*). *Fisheries Sci.* **4** (1): 19–25.
- Moore, S., Reeves, R., Southall, B., Ragen, T., Suydam, R., and Clark, C. 2012. A new framework for assessing the effects of anthropogenic sound on marine mammals in a rapidly changing Arctic. *BioScience* **62** (3): 289–295.
doi:10.1525/bio.2012.62.3.10.
- Nieukirk, S., Stafford, K., Mellinger, D., Dziak, R., and Fox, C. 2004. Low-frequency whale and seismic airgun sounds recorded in the mid-Atlantic ocean. *J. Acoust. Soc. Am.* **115** (4): 1832–1843.
- Nieukirk, S., Mellinger, D., Moore, S., Klinck, K., Dziak, R., and Goslin, J. 2012. Sounds from airguns and fin whales recorded in the mid-Atlantic ocean, 1999–2009. *J. Acoust. Soc. Am.* **131** (2): 1102–1112. doi: 10.1121/1.3672648.
- NOAA. 2015. National Oceanic and Atmospheric Administration DRAFT Guidance for

assessing the effects of anthropogenic sound on marine mammal hearing - underwater acoustic threshold levels for onset of permanent and temporary threshold shifts. Available from [http://www.nmfs.noaa.gov/pr/acoustics/draft-acoustic-guidance July 2015.pdf](http://www.nmfs.noaa.gov/pr/acoustics/draft-acoustic-guidance-July-2015.pdf). [[accessed January 15 2016](#)]

Nowacek, D., Thorne, L., Johnston, D., and Tyack, P. 2007. Responses of cetaceans to anthropogenic noise. *Mammal Rev.* **37** (2): 81-115. doi: 10.1111/j.1365-2907.2007.00104.x

Papale, E., Gamba, M., Perez-Gil, M., Martel Martin, V., and Giacoma, C. 2015. Dolphins adjust species-specific frequency parameters to compensate for increasing background noise. *PloS One* **10** (4): e0121711. doi:10.1371/journal.pone.0121711.

Parks, S., Clark, C., and Tyack, P. 2007. Short- and long-term changes in right whale calling behavior: the potential effects of noise on acoustic communication. *J. Acoust. Soc. Am.* **122** (6): 3725–31. doi:10.1121/1.2799904.

Parks, S., Johnson, E., Nowacek, D., and Tyack, P. 2011. Individual right whales call louder in increased environmental noise. *Biol. Lett.* **7** (1): 33–35.

Parsons, E., Dolman, S., Jasny, M., Rose, N., Simmonds, M., and Wright, A. 2009. A critique of the UK's JNCC seismic survey guidelines for minimizing acoustic disturbance to marine mammals: best practice? *Marine Poll. Bull.* **58** (5): 643–51.

Patricelli, G., and Blickley, J. 2006. Overview: Avian Communication in urban noise: the causes and consequences of vocal adjustment. *Auk* **123**: 639–49.

- Pirotta, E., Brookes, K., Graham, M., and Thompson, P. 2014. Variation in Harbour porpoise activity in response to seismic survey noise. *Biol. Lett.* **10**: 5. DOI: 10.1098/rsbl.2013.1090
- Pirotta, E., Thompson, P., Cheney, B., Donovan, C., and Lusseau, D. 2015. Estimating spatial, temporal and individual variability in dolphin cumulative exposure to boat traffic using spatially explicit capture–recapture methods. *Anim. Conserv.* **18** (1):20-31.
- Potter, J., Thillet, M., Douglas, C., Chitre, M., Doborzynski, Z., and Seekings, P. 2007. Visual and passive acoustic marine mammal observations and high-frequency seismic source characteristics recorded during a seismic survey. *IEEE J. Oceanic Eng.* **32** (2): 469–83.
- Rendell, L., and Gordon, J. 1999. Vocal response of long-finned pilot whales (*Globicephala melas*) to military sonar in the Ligurian Sea. *Mar. Mamm. Sci.* **15** (1): 198–204. doi:10.1111/j.1748-7692.1999.tb00790.x.
- Richardson, W., Greene, C., Malme, C., and Thomson, D. 1995. Marine mammals and noise. Academic Press Limited, San Diego, CA.
- Richardson, W., Greene, C., Hanna, J., Koski, W., Miller, M., Patenaude, M., and Smultea, M. 1995. Acoustic effects of oil production activities on Bowhead and white whales visible during spring migration New Pt. Barrow, Alaska - 1991 and 1994 phases: sound propagation and whale responses to playbacks of icebreaker noise. King City, Ontario. Available from <http://www.arlis.org/docs/vol1/A/34313549.pdf>. [accessed on January 15 2016]

- Richardson, W., Fraker, J., Würsig, B., and Wells, R. 1985. Behaviour of Bowhead whales *Balaena mysticetus* summering in the Beaufort Sea: Reactions to industrial activities. *Biol. Conserv.* **32** (3): 195–230.
- Richardson, W., Würsig, B., and Greene, B. 1986. Reactions of Bowhead whales, *Balaena mysticetus*, to seismic exploration in the Canadian Beaufort Sea. *J. Acoust. Soc. Am.* **80** (4): 1117–28.
- Richardson, W., Würsig, J., and Greene, C. 1990. Reactions of bowhead whales, *Balaena mysticetus*, to drilling and dredging noise in the Canadian Beaufort Sea. *Mar. Environ. Res.* **29** (2): 135–60. doi:10.1016/0141-1136(90)90032-J.
- Risch, D., Corkeron, P., Ellison, W., and van Parijs, S. 2012. Changes in humpback whale song occurrence in response to an acoustic source 200 Km away. *PLoS ONE* **7** (1): e29741
- Robertson, F. 2014. Effects of seismic operations on bowhead whale behaviour: implications for distribution and abundance assessments. Univeristy of British Columbia. PhD thesis. 231 pp. Available from http://www.marinemammal.org/wp-content/pdfs/Robertson_2014.pdf [accessed on January 2016]
- Robertson, F., Koski, W., Thomas, T., Richardson, W., Würsig, B., and Trites, A. 2013. Seismic operations have variable effects on dive-cycle behavior of bowhead whales in the Beaufort Sea. *Endanger. Species Res.* **21** (2): 143–60.
- Rolland, R., Parks, S., Hunt, K., Castellote, M., Corkeron, P., Nowacek, D., Kraus S., Wasser, S., and Kraus, S. 2012. Evidence that ship noise increases stress in Right

whales. *Proc. R. Soc. Lond. B. Biol. Sci.* **279**: 2363–68.

doi:10.1098/rspb.2011.2429.

Scheifele, P., Andrew, M., Cooper, R., Darre, M., Musiek, F., and Max, L. 2005.

Indication of a Lombard vocal response in the St. Lawrence River Beluga. *J.*

Acoust. Soc. Am. **117** (3 Pt 1): 1486–92.

Scholik-Schlomer, A. 2015. Where the decibels hit the water: perspectives on the

application of science to real-world underwater noise and marine protected species

issues. *Acoustics Today* **11** (3): 36–44.

Sivle, L. D., P. H. Kvadsheim, A. Fahlman, F. P A Lam, P. L. Tyack, and P. J O Miller.

2012. Changes in dive behavior during Naval sonar exposure in killer whales, long-

finned pilot whales, and sperm whales. *Front. Physiol.* **3** (400): 1–11. doi:

10.3389/fphys.2012.00400

Smith, C., Sykora-Bodie, S., Bloodworth, B., Pack, S., Spradlin, T., and LeBoeuf, N.

2016. Assessment of known impacts of unmanned aerial systems (UAS) on marine mammals: data gaps and recommendations for researchers in the United States.

Journal of Unmanned Vehicle Systems **4**(1): 31–44. doi: 10.1139/juvs-2015-0017

SMRU 2014. Determination of behavioural effect noise thresholds for Southern

Resident Killer Whales, Proposed Roberts Bank Terminal 2, Technical data report.

SMRU Canada Ltd. (Prepared for Hemmera) 268 pp. Available from

<http://www.robertsbankterminal2.com/wp-content/uploads/RBT2-Determination-of-Behavioral-Effect-Thresholds-for-Southern-Resident-Killer-Whales-TDR.pdf>

accessed on January 15 2016]

- Southall, B., Bowles, A., Ellison, W., Finneran, J., Gentry, R., Greene, C., Kastak, D., Ketten, D., Miller, J., Nachtigall, P., Richardson, J., Thomas, J., and Tyack, P. 2007. Marine mammal Noise exposure criteria: initial scientific recommendations. *Aquat. Mamm.* **33** (4): 411–522.
- Southall, B., Moretti, D., Abraham, B., Calambokidis, J., DeRuiter, S., and Tyack, P. 2012. Marine mammal behavioral response studies in Southern California: advances in technology and experimental methods. *Mar. Technol. Soc. J.* **46** (4): 48–59.
- Southall, B., Rowles, T., Gulland, F., Baird, R., and Jepson, P. 2013. Final report of the independent scientific review panel investigating potential contributing factors to a 2008 mass stranding of melon-headed whales (*Peponocephala electra*) in Antsohihy, Madagascar. Washington D.C. 5 pp. Available from http://www.cascadiaresearch.org/Hawaii/Madagascar_ISR_P_Final_report.pdf [accessed on January 15 2016]
- Statoil 2015. Marine noise inputs. Technical note on underwater noise. Statoil ASA. 28 Pp. Available from <http://www.statoil.com/no/EnvironmentSociety/Environment/impactassessments/NewEnergy/IntWind/Downloads/supporting%20studies/Underwater%20noise%20technical%20assessment.pdf> [accessed on January 15 2016]
- Stimpert, A., DeRuiter, S., Southall, B., Moretti, D., Falcone, E., Goldbogen, J., Friedlaender, A., Schorr, G., and Calambokidis J. 2014. Acoustic and foraging behavior of a Baird's beaked whale, *Berardius bairdii*, exposed to simulated sonar.

- Sci. Rep. **4** (7031) 1–8. doi:10.1038/srep07031.
- Terhune, J., Stewart, R., and Ronald, K. 1979. Influence of vessel noises on underwater vocal activity of harp seals. *Can. J. Zool.* **57** (6): 1337–38.
- Thompson, P., Brookes, K., Graham, I., Barton, T., Needham, K., Bradbury, G., and Merchant, N. 2013. Short-term disturbance by a commercial two-dimensional seismic survey does not lead to long-term displacement of harbour porpoises. *Proc. R. Soc. Lond. B. Biol. Sci.* **280** (1771): 20132001.
- Todd, S., Stevick, P., Lien, J., Marques, F., and Ketten, D. 1996. Behavioural effects of exposure to underwater explosions in humpback whales (*Megaptera novaeangliae*). *Can. J. Zool.* **74**: 1661–72.
- Tougaard, J., Wright, A., and Madsen, P. 2015a. Cetacean noise criteria revisited in the light of proposed exposure limits for harbour porpoises. *Mar. Poll. Bull.* **90** (1-2): 196–208. doi: <http://dx.doi.org/10.1016/j.marpolbul.2014.10.051>.
- Tripovich, J., Hall-Aspland, S., Charrier, I., and Arnould, J. 2012. The behavioural response of Australian fur seals to motor boat noise. *PLoS ONE* **7** (5): e37228.
- Tyack, P. 1983. Differential response of humpback whales, *Megaptera novaeangliae*, to playback of song or social sounds. *Behav. Ecol. Sociobiol.* **13** (1): 49–55.
- Tyack, P. 2009. Acoustic playback experiments to study behavioral responses of free-ranging marine mammals to anthropogenic sound. *Mar. Ecol. Prog. Ser.* **395**: 187–200.
- Tyack, P., Zimmer, W., Moretti, D., Southall, B., Claridge, D., Durban, J., Clark, C.,

- D'Amico, A., DiMario, N., Jarvis, S., McCarthy, E., Morrissely, R., Ward, J., and Boyd, I. 2011. Beaked whales respond to simulated and actual Navy sonar. *PLoS ONE* **6** (3): e17009.
- Van Parijs, S M, and P Corkeron. 2001. Boat traffic affects the acoustic behaviour of pacific humpback dolphins, *Sousa chinensis*. *J. Mar. Biol. Assoc. U. K.* **81** (3): 533–38.
- Veirs S, Veirs V., and Wood J. 2016. Ship noise extends to frequencies used for echolocation by endangered killer whales. *PeerJ* **4**: e1657; DOI 10.7717/peerj.1657
- Wade, L., Whitehead H., and Weilgart, L. 2010. Conflict of interest in research on anthropogenic noise and marine mammals: does funding bias conclusions? *Mar. Policy* **34** (2): 320-327.
- Wartzok, D., Popper, A., Gordon, J., and Merrill, J. 2004. Factors affecting the responses of marine mammals to acoustic disturbance. *Mar. Tech. Soc. J.* **37** (4):1-15.
- Watkins, W., Moore, K., and Tyack, P. 1985. Investigations of sperm whale acoustic behaviors in the Southeast Caribbean. *Cetology* **49**: 1–15.
- Watwood, S., Miller, P., Johnson, M., Madsen, P., and Tyack, P. 2006. Deep-diving foraging behaviour of sperm whales (*Physeter macrocephalus*). *J. Anim. Ecol.* **75** (3): 814–25.
- Weaver, A. 2015. Sex difference in bottlenose dolphin sightings during a long-term bridge construction project. *Animal Behavior and Cognition* **2** (1): 1–13.

doi:10.12966/abc.02.01.2015.

Weilgart, L. 2007. The impacts of anthropogenic ocean noise on cetaceans and implications for management. *Can. J. Zool.* **85**: 1091-1116. doi:10.1139/Z07-101

Wensveen, P. 2012. The effects of sound propagation and avoidance behaviour on naval sonar levels received by cetaceans. PhD. University of St. Andrews, Scotland. 116pp.

Williams, R., Erbe, C., Ashe, E., Beerman, A., and Smith, J. 2014. Severity of killer whale behavioral responses to ship noise: A dose-response study. *Mar. Poll. Bull.* **80** (1-2): 254–60.

Williams, R., Lusseau, D., and Hammond, P. 2006. Estimating relative energetic costs of human disturbance to killer whales (*Orcinus orca*). *Biol. Conserv.* **133** (3): 301–11.

Wood, J., Southall, B., and Tollit, D. 2012. PG&E Offshore 3-D Seismic Survey Project EIR – Marine Mammal Technical Report. Available from <http://www.coastal.ca.gov/energy/seismic/mm-technical-report-EIR.pdf>. [accessed January 15 2016]

Wright, A., and de Soto, N. 2007. Do marine mammals experience stress related to anthropogenic noise? *Int. J. Comp. Psychol.* **20** (2): 274–316.

Wright, A., and Kyhn, L. 2014. Practical management of cumulative anthropogenic impacts with working marine examples. *Conserv. Biol.* **29** (2): 1 – 8.
doi:10.1111/cobi.12425.

Wright, A. 2015 .Sound Science: Maintaining Numerical and Statistical Standards in the Pursuit of Noise Exposure Criteria for Marine Mammals. *Front. Mar. Sci.* **2**:99. doi: 10.3389/fmars.2015.00099

Draft

Tables

Table 1. Examples of definitions provided by Southall et al. (2007) to assign a behavioural response severity score. Subjective definitions are highlighted in bold.

Score	Corresponding Behaviour Types (e.g.,)
0	No observable response
1	Brief orientation response
2	Moderate or multiple orienting behaviours
3	Prolonged orientation behaviour
4	Changes in locomotion speed, direction, or diving
5	Extensive or prolonged changes in locomotion speed, direction, or diving
6	Minor or moderate individual or group avoidance of sound source
7	Severe and/or sustained avoidance of sound source
8	Long-term avoidance of area (>source operation)
9	Outright panic, flight, stampede, attack of conspecifics, or stranding events

Table 2. Proposed cut-offs of the 10-index score proposed by Southall et al. (2007), Miller et al. (2012), Williams et al. (2014) and this review.

Score	Southall et al. (2007)	Miller et al. (2012)	Williams et al. (2014)	This review
0	Minor/ brief responses	Not likely to affect vital rates	No change	Low/Nil
1			Minor	
2			Moderate	
3	Higher potential to affect foraging, reproduction, or survival	Likely to affect vital rates	Extensive	Moderate
4			Not applicable	High
5				Very high
6	Likely to affect foraging, reproduction or survival			
7				
8				
9				

Draft

Table 3. Number of data cases for studies that reported RL and behavioural responses of cetaceans to Low Frequency Active Sonar (LFAS) in the wild. Note that data cases can represent either one or multiple studies.

Studies	Frequency reported	Name reported	Name assigned	# cases (baleen whales)	# data cases (toothed cetaceans)
Clark et al. 1999; Buck and Tyack 2000 in Ellison et al. 2012	160 - 330 Hz	LFAS	LFAS	2	
Croll et al. 2001	130 - 320 Hz	US Navy SURTASS LFAS	LFAS	2	
Frankel and Clark 1998	centred at 75 Hz with a 30-Hz bandwidth	75-Hz M-sequence sound	LFAS	1	
Frankel and Clark 2000	75 Hz, with a noticeable third harmonic at 225 Hz	Acoustic Thermometry of Ocean Climate (ATOC) sounds	LFAS	1	
Fristrup et al. 2003	150 - 320 Hz	US Navy SURTASS LFAs	LFAS	1	
Miller et al 2011; 2012; 2014	1 - 2 kHz	LFAS	MFAS		9
Miller et al 2011; 2012; Antunes et al. 2014; Alves et al. 2014	1 - 2 kHz	LFAS	MFAS		15
Miller et al 2011, 2012; Sivle et al. 2012; Wensveen 2012	1 - 2 kHz	LFAS	MFAS		22
Miller et al. 2000; Biassoni et al. 2000	130 - 160 Hz (low component) and 260 – 320 Hz (high component)	US Navy SURTASS LFAS	LFAS	1	
Risch et al. 2012 (with rebuttal by Gong et al. 2014)	bandwidth of roughly 50 Hz, centre frequencies approx. 415, 735, 950 Hz	Ocean Acoustic Waveguide Remote Sensing (OAWRS) sounds	LFAS	1	
Number of data cases				9	46

Table 4. Functional hearing groups representing the frequency band of hearing for all species within the group, although individual species' hearing ranges are typically not as broad (Southall et al. 2007, NOAA 2015). The hearing range for the low-frequency group was estimated based on behavioural studies, recorded vocalizations, and inner ear morphology measurements (NOAA 2015).

Functional hearing group Species grouping or species	Hearing Range
High-frequency cetaceans (HF toothed cetaceans) Harbour porpoise (<i>Phocoena phocoena</i>), <i>Kogia</i> spp., river dolphins, Hector's dolphin (<i>Cephalorhynchus hectori</i>), Hourglass dolphins (<i>Lagenorhynchus cruciger</i>), Peale's dolphins (<i>Lagenorhynchus australis</i>)	200 Hz to 180 kHz
Mid-frequency cetaceans (MF toothed cetaceans) Beaked whales [Baird's beaked whale (<i>Berardius</i> spp.), Blainville's beaked whale (<i>Mesoplodon densirostris</i>), Cuvier's beaked whale (<i>Ziphius cavirostris</i>)] Beluga whale (<i>Delphinapterus leucas</i>) Dolphins [Bottlenose dolphin (<i>Tursiops truncatus</i>), common dolphin (<i>Delphinus delphis</i>), Indo-pacific bottlenose dolphin (<i>Tursiops aduncus</i>), rough-toothed dolphin (<i>Steno bredanensis</i>), white-beaked dolphin (<i>Lagenorhynchus albirostris</i>), white-sided dolphin (<i>Lagenorhynchus acutus</i>)] Killer whale (<i>Orcinus orca</i>) Sperm whale (<i>Physeter macrocephalus</i>)	150 Hz to 160 kHz
Low-frequency cetaceans (LF baleen whales) Bowhead whale (<i>Balaena mysticetus</i>), right whale (<i>Eubalaena</i> spp.), blue whale (<i>Baleanoptera musculus</i>), fin whale (<i>Balaenoptera physalus</i>), gray whale (<i>Eschrichtius robustus</i>), humpback whale (<i>Megaptera novaengliae</i>), minke whale (<i>Balaenoptera acutorostrata</i>)	7 Hz to 30 kHz
Phocid pinnipeds Harbor seal (<i>Phoca vitulina</i>), grey seal (<i>Halichoerus grypus</i>), ringed seal (<i>Phoca hispida</i>), bearded seal (<i>Erignathus barbatus</i>)	75 Hz to 100 kHz
Otariid pinnipeds Sea lions and fur seals, California sea lion (<i>Zalophus californianus</i>)	100 Hz to 40 kHz

Table 5. Behaviour category assigned for each study when information on the behaviour, state, or activities were reported on wild individuals or populations prior to being exposed to noise.

Behaviour prior to exposure	Examples of behaviour types, state and/or activities reported in the literature reviewed
Foraging	Feeding, foraging, vocalizations as proxy for foraging, deep feeding/foraging
Resting	Resting, hauling-out during non-breeding season
Socializing	Socializing, breeding, calving, mating, vocalizing (non-foraging calls such as reproductive calls)
Traveling	Traveling, migrating, directional swimming

Draft

Table 6. Number of data cases compiled for studies of wild individuals/populations that provided RL and behavioural responses for several types of sound sources ($N = 256$). The number of data cases for the behavioural score responses 'very high' are presented in parenthesis and the data cases highlighted in bold were used in the analysis. Each data case can represent one study or can represent multiple studies.

Sounds	HF toothed cetaceans	MF toothed cetaceans	LF baleen whales	Manatees	Phocid and otariid pinnipeds	Various
Acoustic alarm	14	3+(1)	1		2	1
Artificial sounds			2		2	
High frequency			1			
LFAS			9		1	
MFAS		81+(1)	4			
Missile launch					3	
Pile/pipe driving	2				2	
Sonic booms					1	
Seismic-Explosion	1	6	42		2	
Aircraft		2	2		1	
Continuous	1	49+(2)	14	2	(1)	
Number of data cases	18	145	75	2	15	1

Table 7. Model selection statistics. $\Delta AICc$: delta Akaike information criterion; E_i : evidence ratio; k : number of estimated parameters (Burnham and Anderson 2002; Anderson 2008). Data used here correspond to the majority of data cases compiled (highlighted in bold in Table 6) ($N = 195$). Only results with $E_i < 10$ are presented.

Model	k	$\Delta AICc$	E_i
Functional hearing group + Sound source + Functional hearing group * Source	5	0.000	1.000
Functional hearing group + Sound source + RL + Functional hearing group * Sound source	6	1.77	2.42
Functional hearing group + Sound source	3	2.86	4.17
Functional hearing group + Sound source + RL	4	4.25	8.38
Functional hearing group + RL	6	5.28	14.02

Table 8. Studies that report changes in the vocal behaviour of wild marine mammals due to noise exposure.

Vocal behaviour change	Species	Reference
Vocalisation duration increase	killer whale, humpback whale, bowhead whale	Biassoni et al. 2000; Miller et al. 2000; Fristrup et al. 2003; Foote et al. 2004; Blackwell et al. 2015
Vocalisation duration decrease	fin whale, manatee	Castellote et al. 2012; Miksis-Olds and Tyack 2009
Vocalisation cessation	sperm whale, Blainville's beaked whale, beluga, blue whale, humpback whale, fin whale, bowhead whale, Baird's beaked whale, killer whale, Cuvier's beaked whale	Watkins et al. 1985; Lesage et al. 1999; Clark and Gagnon 2006; Aguilar Soto et al. 2006; McDonald et al. 1995; Tyack et al. 2011; Richardson et al. 1985; Cerchio et al. 2014; Stimpert et al. 2014; Miller et al. 2012; Risch et al. 2012; Blackwell et al. 2015; DeRuiter et al. 2013a
Vocalization rate decrease	right whale, Blainville's beaked whale, blue whales, sperm whale, manatees, pilot whale, bowhead whale, melon-headed whale, harp seals	Potter et al. 2007; Parks et al. 2007; Miksis-Olds and Tyack 2009; McCarthy et al. 2011; Melcón et al. 2012; Miller et al. 2012; Blackwell et al. 2015; DeRuiter et al. 2013b; Terhune et al. 1979; Azzara et al. 2013
Vocalization rate increase	bottlenose dolphin, beluga, Pacific humpback dolphin, sperm whale, false killer whale, humpback whale, killer whale, blue whale, pilot whale, harp seals	Lesage et al. 1999; Buckstaff 2004; Doyle et al. 2008; Tyack 1983; Van Parijs and Corkeron 2001; DeRuiter et al. 2013a; Tyack et al. 2011; Miller et al. 2012; Di Iorio and Clark 2010; Rendell and Gordon 1999
Vocalization frequency shift	gray whale, fin whale, right whale, beluga, bottlenose dolphin, short-beaked common, Atlantic spotted, and striped-dolphin, false killer whale, killer whale, pilot whale	Dahlheim 1987; Lesage et al. 1999; Parks et al. 2007; Castellote et al. 2012; La Manna et al. 2013; Papale et al. 2015; DeRuiter et al. 2013b; Tyack et al. 2011; Miller et al. 2012

Vocal behaviour change	Species	Reference
Vocalization intensity increase	killer whale, right whale, bottlenose dolphin, beluga whale, blue whale	Scheifele et al. 2005; Finley et al. 1990; Holt et al. 2009; Holt et al. 2009, 2011; La Manna et al. 2013; McKenna 2011; Parks et al. 2011
Vocalization redundancy increase	Beluga, pilot whale	Lesage et al. 1999; Rendell and Gordon 1999
Vocalization matching	pilot whale, false killer whale	Miller et al. 2012; Antunes et al. 2014; Alves et al. 2014; DeRuiter et al. 2013 <i>b</i>

Draft

List of figures

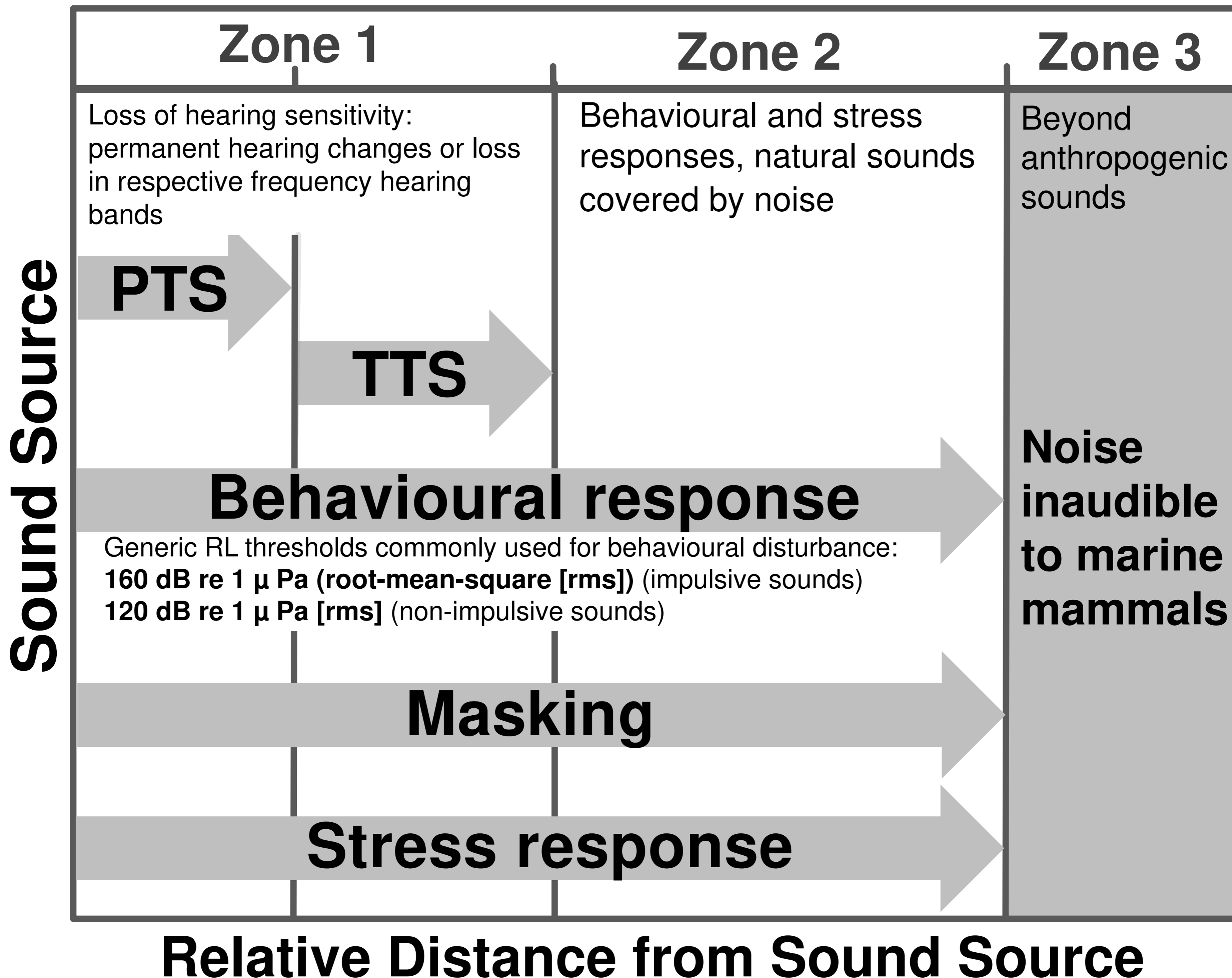
Figure 1. Examples of the effects of noise on marine mammals with respect to distance from the sound source. Zone 1 represents areas closer to the sound source, with RL of relatively high intensity that could cause a loss of hearing sensitivity: Permanent Threshold Shift (PTS) or Temporary Threshold Shift (TTS). Zone 2 represents areas farther from the sound source where individuals might show behavioural responses (e.g., cessation of feeding), stress responses, and/or experience acoustic masking (interference with individual's ability to detect, recognize and/or discriminate sounds). Zone 1 can include all types of impact. Zone 2 only includes long-range impacts. Zone 3 represents areas beyond those where impacts are expected. Adapted from Fig. 1 of Dooling and Therrin (2007).

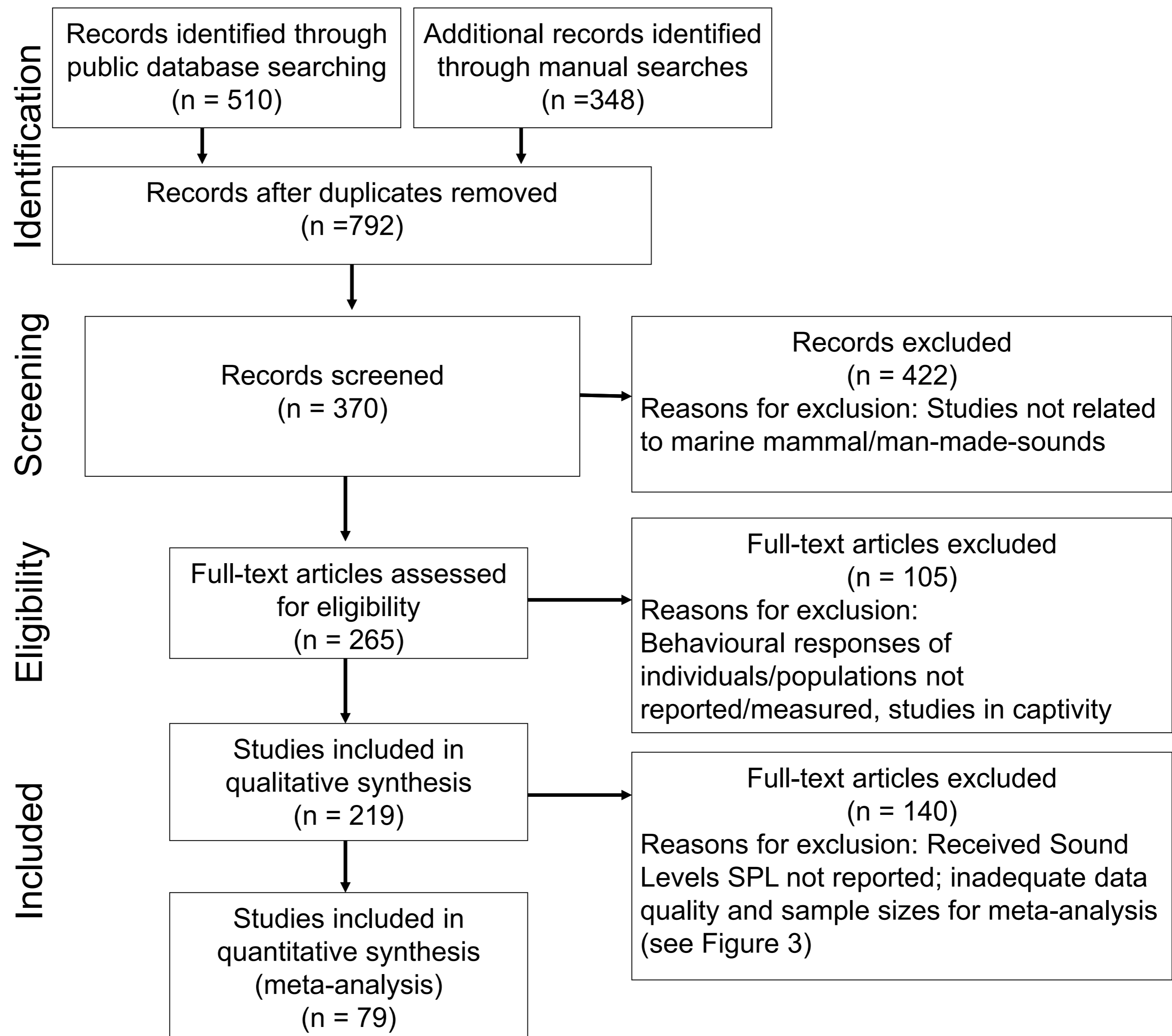
Figure 2. PRISMA (preferred reporting items for systematic reviews and meta-analyses) diagram illustrating the process for selection of publications included in the systematic review.

Figure 3. Process for selection of variables and data cases included in the analysis. Panel A illustrates the process for selecting data cases for the analysis. Panel B illustrates the reduction in the number of data cases when including two additional contextual variables that had the largest sample sizes: data cases equal to or larger than 10.

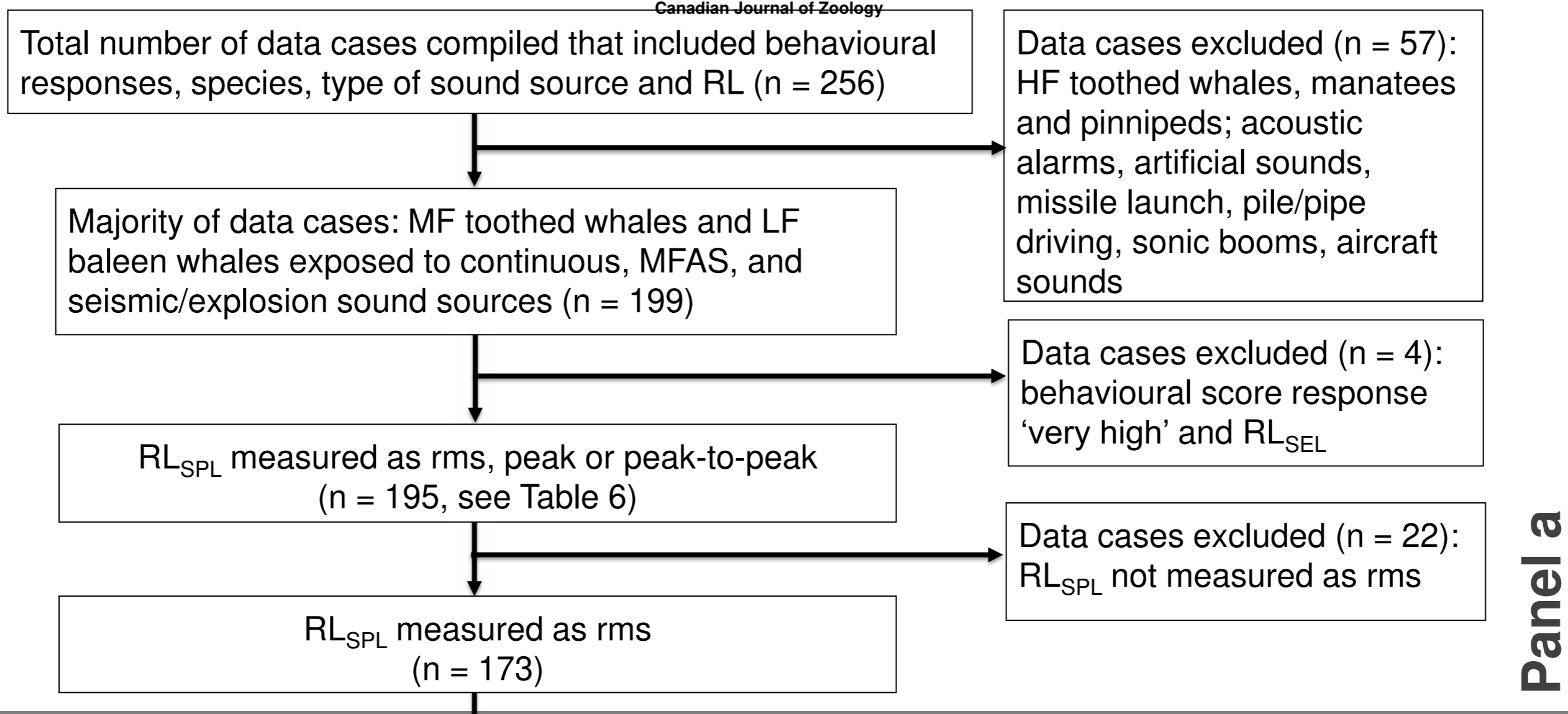
Figure 4. Probability density function (via Kernel density estimation) of the behavioural response severity score (low, moderate, high) of LF baleen whales and MF toothed cetaceans in relation to RL of continuous, MFAS, and seismic/explosion sound sources. The values provided for behavioural score responses and RL_{SPL} are the maximum reported for each study reviewed. RL SPL, measured as rms, peak or peak-to-peak, ranged between 82 to 180, see raw data in S2. N = number of data cases.

Figure 5. Relative proportion of behavioural response severity scores (high, moderate, low) for each sound source (sample size indicated on top of each bar) and functional hearing group (LF baleen whales and MF toothed cetaceans). Predicted model probabilities are presented as black dots with corresponding 95% confidence intervals. Model probabilities and relative proportion of data cases for each behavioural response severity score showed little discrepancy, suggesting overall good model fit. Due to small sample sizes, caution is warranted when interpreting the specific patterns of severity.



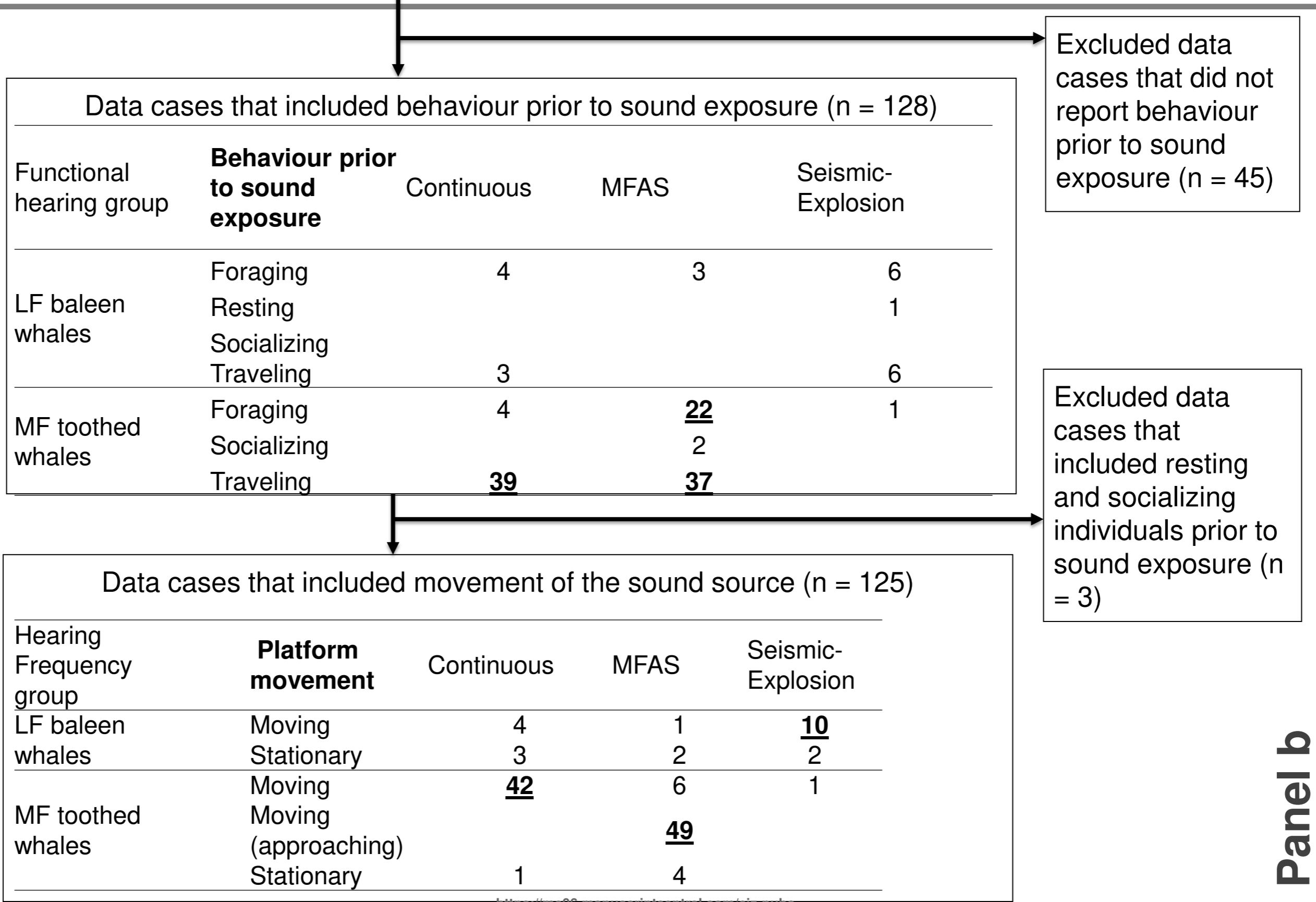


Included in meta-analysis

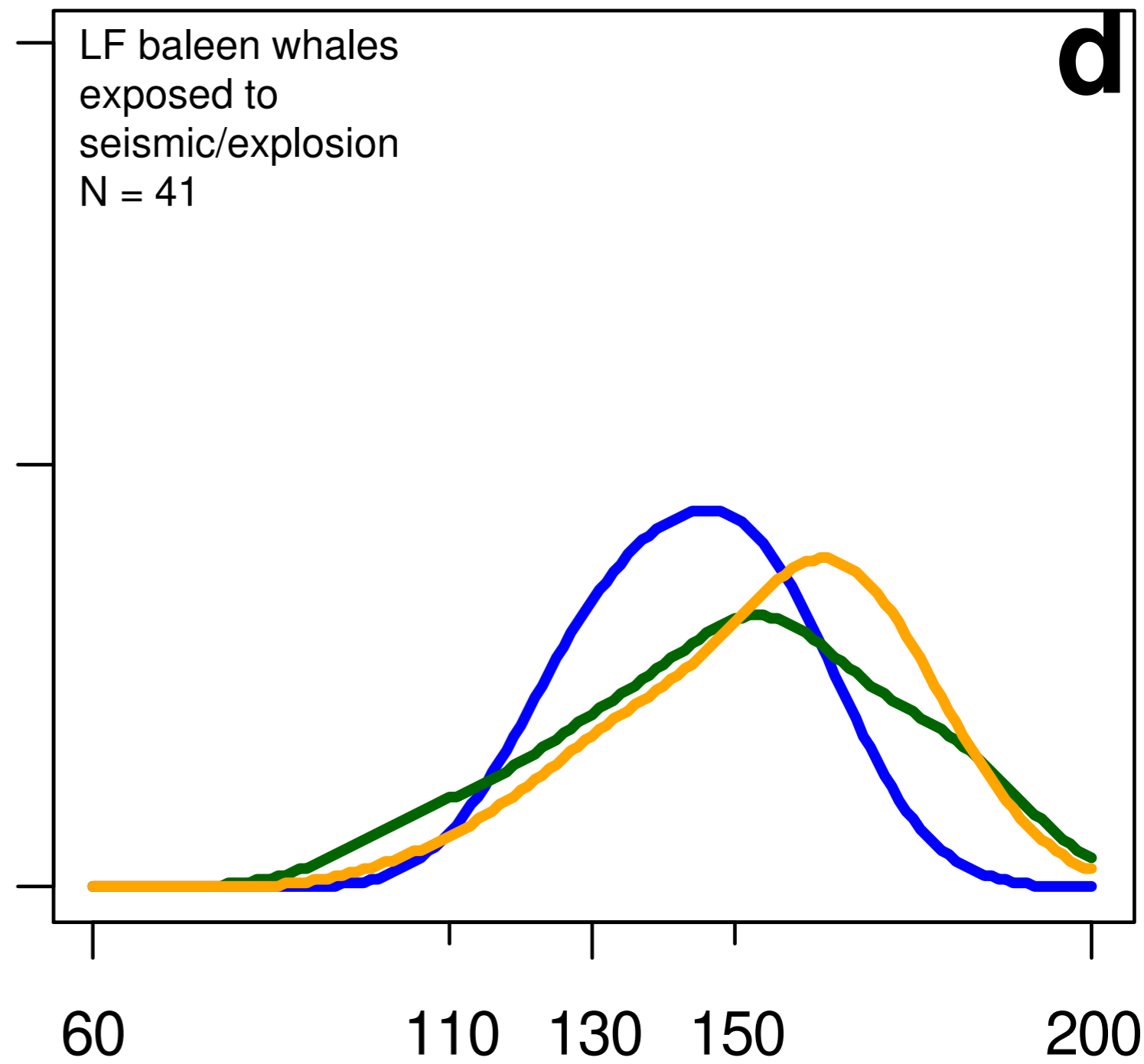
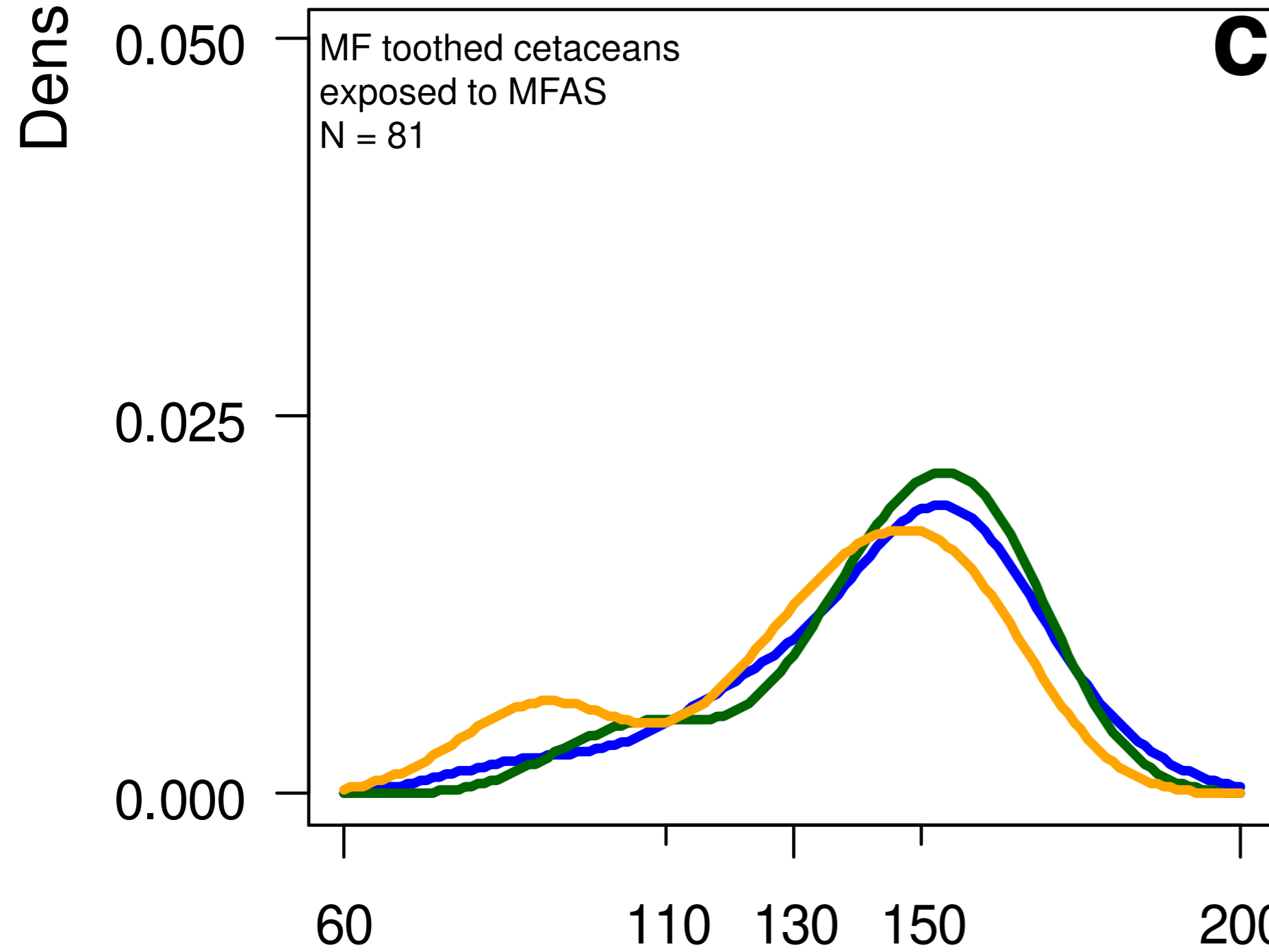
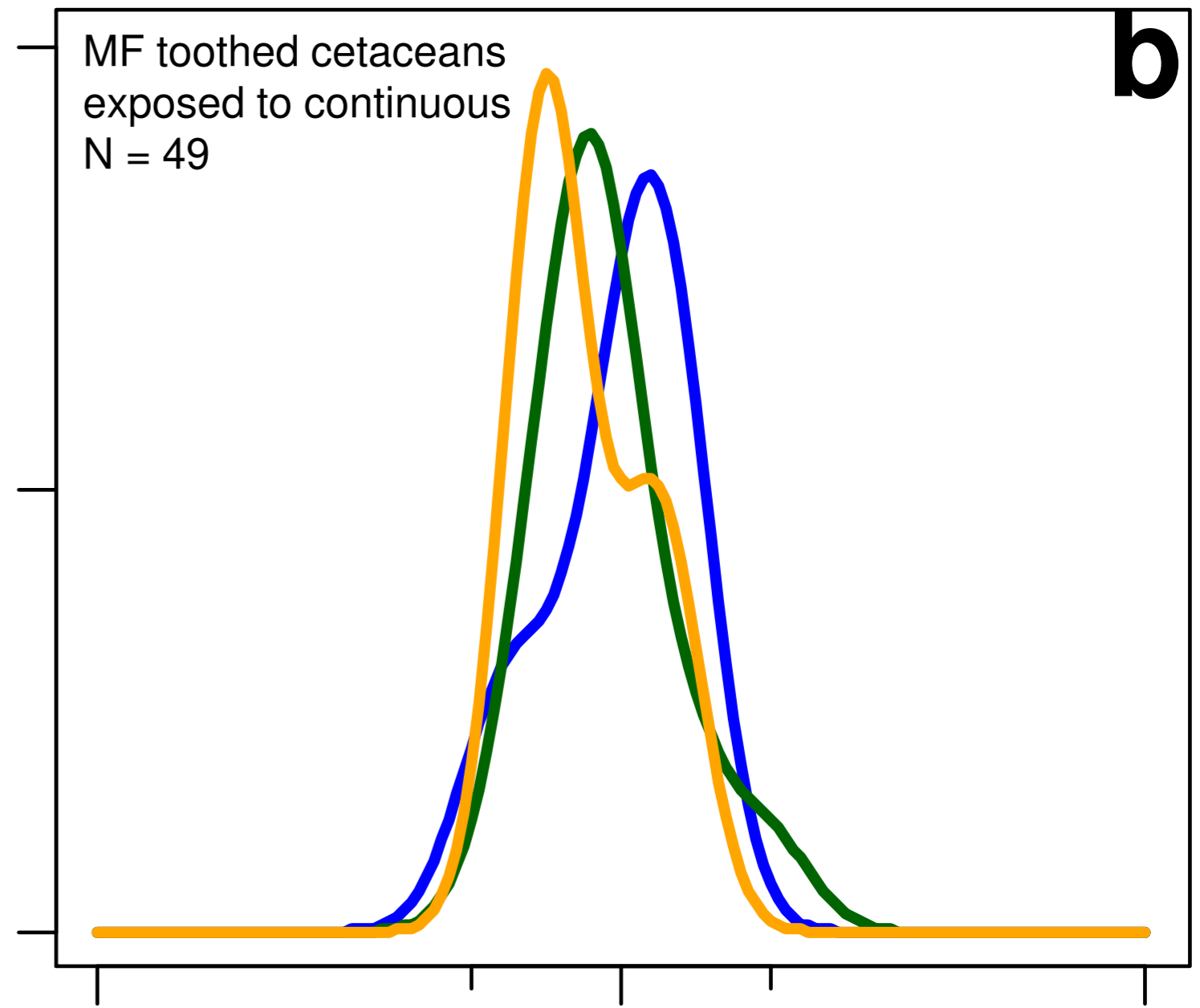
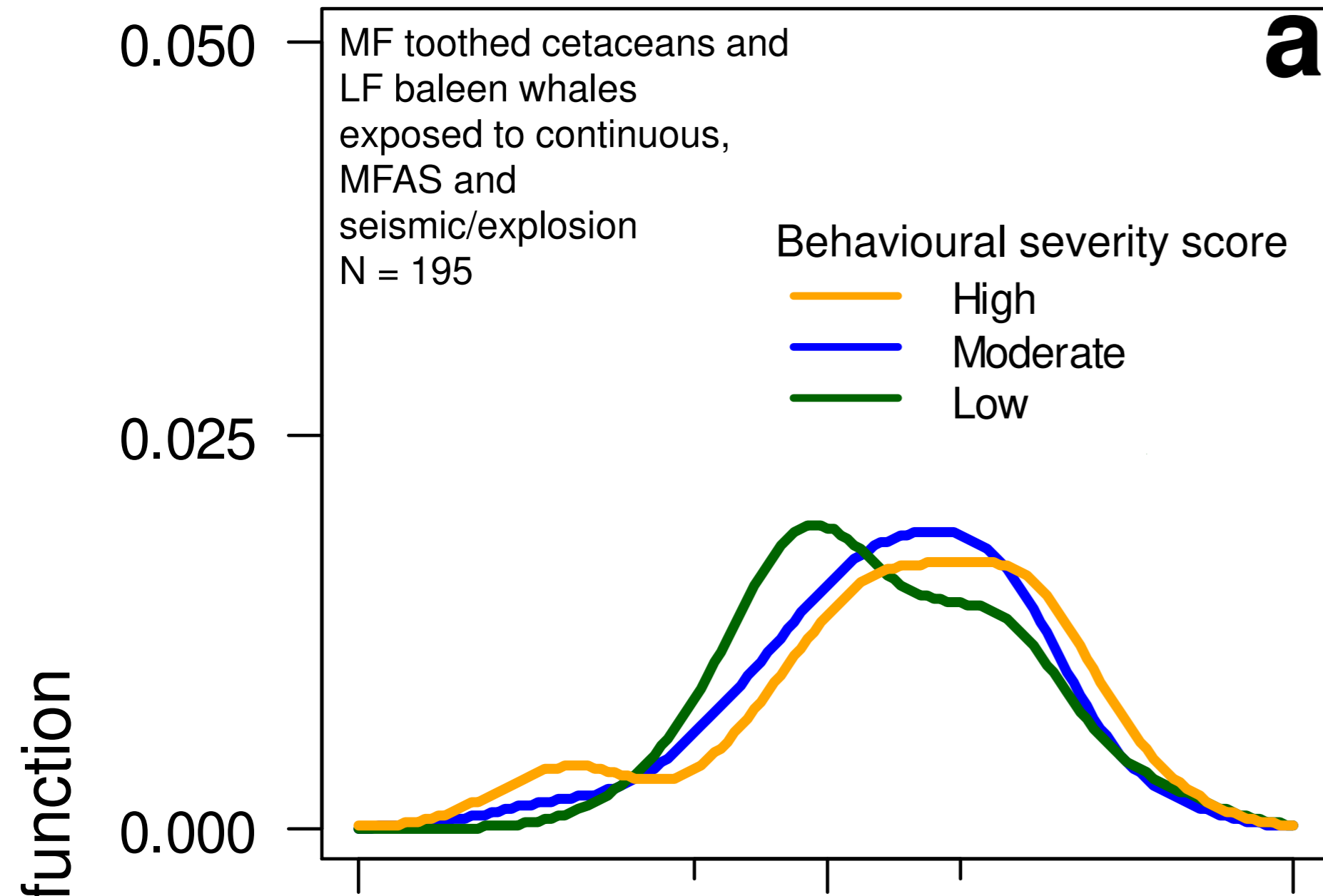


Panel a

Included in qualitative analysis



Panel b



Received Sound Levels (dB re 1 μ Pa)

