

Contents lists available at [ScienceDirect](http://ScienceDirect)

# Marine Pollution Bulletin

journal homepage: [www.elsevier.com/locate/marpolbul](http://www.elsevier.com/locate/marpolbul)

## Cetacean noise criteria revisited in the light of proposed exposure limits for harbour porpoises

Jakob Tougaard<sup>a,\*</sup>, Andrew J. Wright<sup>a,b</sup>, Peter T. Madsen<sup>c,d</sup><sup>a</sup> Aarhus University, Department of Bioscience, Frederiksborgvej 399, DK-4000 Roskilde, Denmark<sup>b</sup> Department of Environmental Science and Policy, George Mason University, 4400 University Drive, Fairfax, VA 22030, USA<sup>c</sup> Aarhus University, Department of Bioscience, DK-8000 Aarhus C, Denmark<sup>d</sup> Murdoch University Cetacean Research Unit, Murdoch University, WA, Australia

### ARTICLE INFO

#### Article history:

Available online 20 November 2014

#### Keywords:

Underwater noise  
Temporary threshold shift  
Impact assessment  
Behavioural response  
Hearing  
Frequency weighting

### ABSTRACT

The impact of underwater noise on marine life calls for identification of exposure criteria to inform mitigation. Here we review recent experimental evidence with focus on the high-frequency cetaceans and discuss scientifically-based initial exposure criteria. A range of new TTS experiments suggest that harbour and finless porpoises are more sensitive to sound than expected from extrapolations based on results from bottlenose dolphins. Furthermore, the results from TTS experiments and field studies of behavioural reactions to noise, suggest that response thresholds and TTS critically depend on stimulus frequency. Sound exposure levels for pure tones that induce TTS are reasonably consistent at about 100 dB above the hearing threshold for pure tones and sound pressure thresholds for avoidance reactions are in the range of 40–50 dB above the hearing threshold. We propose that frequency weighting with a filter function approximating the inversed audiogram might be appropriate when assessing impact.

© 2014 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

### 1. Introduction

Hearing is the primary sense for much marine life for detecting signals from prey, predators, conspecifics, competitors and the environment. Noise introduced into the environment by human activities thus has the potential to interfere with auditory detection and thereby affect the animals directly as well as indirectly via prey and predators (e.g. Richardson et al., 1995; National Research Council, 2003, 2005; Nowacek et al., 2007; Weilgart, 2007). The magnitude of this problem has been realised slowly over the last four decades and as a consequence underwater noise has gradually moved up on the political agenda (see Simmonds et al., 2014). In an influential review of the field, Richardson et al. (1995) focused primarily on the descriptive; a collection of all available information on relevant noise sources and studies of their impact on marine mammals. A decade later Southall et al. (2007) made an updated review of the literature and offered the first published scientific guidance regarding noise exposure criteria.

Although scientifically based, this paper is heavily influenced by, and targeted to, policy in the United States because the suggested criteria are based on the legal definitions of injury and behavioural harassment under the U.S. Marine Mammal Protection

Act (MMPA). As a consequence of this, the scientific recommendations provided by Southall et al. (2007) may not be appropriate for direct application in other countries, or even under other domestic legislations in the United States, such as the Endangered Species Act. Several non-U.S. legal frameworks, such as the European Habitats Directive (European Commission, 1992) and the Marine Strategy Framework Directive (European Commission, 2008) focus more on sustaining populations, the habitats that support them and the ecosystems of which they are a part, rather than accounting for takes of individual animals as is the case in the MMPA.

Nevertheless, the criteria suggested by Southall et al. (2007) filled a large global policy vacuum. As a consequence of the pressing need for actual exposure criteria, the suggestions of Southall et al. (2007) have quickly acquired status as the *de facto* standard in many political processes around the world, despite the limitations, caveats and lack of information and imperfect understanding carefully emphasised by Southall et al. (2007). Further there has been surprisingly little constructive debate over the contents of the initial recommendations of this paper and no substantive alternatives or augmentations have yet been offered (but see Ellison et al., 2011). As a result, the recently proposed criteria for acoustic injury by the U.S. National Oceanic and Atmospheric Administration (NOAA, 2013) builds directly on methodology and exposure criteria presented by Southall et al. (2007), updated in light of new experimental data obtained after 2007 and considering additional discussions on weighting functions (Finneran and Jenkins, 2012).

\* Corresponding author.

E-mail address: [jat@bios.au.dk](mailto:jat@bios.au.dk) (J. Tougaard).

Southall et al. (2007) should be credited for suggesting actual exposure criteria, as such a move inevitably attracts criticism from those believing them to be too high, those asserting that they are too low, and those merely seeking to improve the science upon which they are based. One of the clearly stated shortcomings of Southall et al. (2007) is that exposure criteria are provided for all cetaceans, divided into three groups, low-frequency (LF), mid-frequency (MF) and high-frequency (HF) cetaceans even though they are all based on experiments on a few species of MF-cetaceans. The MF criteria were then extrapolated to the LF-cetaceans (large whales) and HF-cetaceans (narrow-band high-frequency odontocetes, including porpoises).

Since the review by Southall et al. (2007) a substantial number of studies on other species, particularly the harbour porpoise (*Phocoena phocoena*), have become available. In the light of these new results and in attempt to further the process of identifying meaningful approaches to mitigate noise effects on marine mammals we here undertake a critical, but constructive review of the guidance and recommendations presented by Southall et al. (2007). We then proceed to apply the consequences of that review in a discussion of exposure criteria for the harbour porpoise, a high frequency species where much recent data has been collated. While we use porpoises as a model species, we believe that the scope of the proposed impact assessment has a wider relevance beyond the consideration of any specific legislative standards or species.

## 2. Exposure criteria for injury

Southall et al. (2007) developed the first general exposure criteria regarding injury for marine mammals. In that process, a number of principal and necessarily simplifying assumptions were made. Of these, three are considered key assumptions by us: (1) the auditory system, being adapted for detection of very low sound levels, will be first organ system to suffer injury from sound exposure; (2) that risk of impact scales with loudness-weighted sound exposure level (signal energy); and (3) that the onset permanent threshold shift (PTS) is the basis for defining safe exposure limits. While the following discussion is specifically targeted the criteria proposed by Southall et al. (2007) it is noteworthy that NOAA's (2013) proposed acoustic injury criteria conform to these assumptions.

### 2.1. The auditory system as the most sensitive to injury

Marine mammals in general have acute underwater hearing, and for that reason, the auditory system is considered to be the first to suffer from injuries when exposed to increasingly powerful noise (Southall et al., 2007). Sound exposure, however, may also induce other, potentially injurious effects that are more subtle or hard to measure, and hence be overlooked, particularly in marine animals. One such parameter is physiological discomfort, which is very hard to quantify unless it is extensive enough to materialize in the forms of increased levels of stress hormones or reduced fitness over long periods of time. Several studies on human divers indicate that sound exposures can cause long term physiological effects with consequences for fitness that may not be reliably detected through elevated levels of stress hormones or reduced auditory capabilities in the form of temporary threshold shift (TTS).

Steevens et al. (1999) report two cases of what appears to be noise-induced neurological disturbances in two navy divers. The first diver was exposed to 160 dB re. 1  $\mu$ Pa (rms) for 15 min (190 dB SEL) at 240 Hz, causing no measurable TTS. However, at the end of the exposure he reported light-headedness, somnolence, blurred vision and a vibratory feeling in his extremities, and he was unaware that the sound source had been turned off. Half an hour after the exposure and after being decompressed from

3 atmospheres, the diver again reported to experience nausea and only responded to strong verbal stimuli. These symptoms abated within 30 min and none of physiological variables measured during or after the exposure could explain his response. Three weeks later he again suffered from an episode of light-headedness, memory loss and nausea and, despite an intense neurological examination, no cause could be established. After 9 months he subjectively assigned continued impairments in the forms of insomnia and memory loss to the sound exposure. Sixteen months after the exposure he was undergoing anti-seizure and anti-depressant therapy.

In the second case study of Steevens et al. (1999), a diver was exposed to 181 dB re. 1  $\mu$ Pa (rms) for 15 min (210 dB SEL) at 1000 Hz. This exposure caused along with a moderate TTS of 19 dB, a feeling of light-headedness, inability to concentrate, agitation, blurred vision and head vibrations. The following day the diver reported an increased sensitivity to noise. Two weeks after the exposure, he still experienced heightened sensitivity to noise, increased irritability and concentration problems. A year later he reported that he still felt that he had not recovered fully having concentration problems and mood swings. Similar conditions were also reported over the short-term by Montague and Strickland (1961), although in this case the 23 subject divers also all displayed at least 6–7 dB TTS five minutes after the exposures of unspecified durations to a 1500 Hz pure tone at maximum tolerable levels (50% felt this had been reached by 200 dB re. 1  $\mu$ Pa, rms)

While correlation does not equate causality, these studies nevertheless suggest that exposures with SELs around or even below those shown to cause TTS may, at least under some circumstances, lead to long term neurological disturbances. The consistent effects on the visual system suggest that the mechanism behind these neurological disturbances relate to strong stimulation of the vestibular system that has a strong oculomotor feedback (Parker et al., 1978). There are to our knowledge no dedicated studies of this in marine mammals, but during intense sound exposures some navy dolphins have displayed behaviours, such as biting the experimental setup or refusing to return to the bit plate that may perhaps be indicative of some level of annoyance or physiological discomfort arising from the exposures (Schlundt et al., 2000; Finneran et al., 2002). For example, the simple act of reorienting their heads for exposures may represent an effort to reduce their own exposures due to their highly directional hearing (Au and Moore, 1984). Also Gray and Waerebeek (2011) reported apparent akinesia and possible death of a dolphin incidentally exposed to sounds from a seismic air gun at close range. It is presently unknown if such responses are short term and merely reflect annoyance on the part of the animal or if they in fact are tell-tale signs of neurological disturbances and discomfort similar to that indicated for some human divers at similar or lower sound levels.

The methodology currently in use during intense sound exposure experiments in marine mammals is, in our opinion, unable to uncover such effects. It is therefore possible that marine mammals may, in at least some cases of exposure to high intensity, low frequency sound, suffer from noise-induced neurological disorders that go undetected, but which are potentially more problematic than TTS. This possibility shed in our view doubt on the idea that PTS is an appropriate general threshold for concern about physiological effects and should accordingly motivate specific investigations on this issue for marine mammals.

### 2.2. Scaling of impact with loudness-weighted energy

It is impossible to develop individual exposure criteria for every possible sound source and thus there is a fundamental need for a common metric, which can be used to assess several types of sounds including those for which no experimental data are

available. As pointed out by Southall et al. (2007) and others (e.g. NOAA, 2013; Finneran, 2008; Terhune, 2013), it is essential to take the hearing of the animal into account when evaluating the impact of a particular sound in the form of frequency weighting.

Conceptually, frequency weighting is straightforward, consisting of a band pass filter that de-emphasises those parts of the signal which fall outside the range of best hearing for a given group of animals with similar audiograms. In human audiology this is done through a series of weighting curves, tailored to different types of sound exposure. Most routinely used in human noise assessments is the A-weighting filter (e.g. Anonymous, 2001). C-weighting is a less known and little used alternative to A-weighting that is intended to better reflect human sensitivity to louder sounds. The differences between these weighting functions are found in the steepness of the curves towards low and high frequencies and the extent of the flat pass band in the centre. Two additional weighting curves, the B-weighting and the D-weighting are no longer commonly used and are accordingly not discussed here.

Southall et al. (2007) introduce and discuss an analogue to C-weighting for marine mammals, the M-weighting. Here we will only consider the fundamental assumption that an analogue to the less known and little used C-weighting from human audiometry is used as model, rather than the ubiquitous A-weighting. Both A- and C-weighting functions are based on equal loudness contours for humans, reflecting the fact that the loudness of a sound to a human is not directly related to the sound pressure, but a function of both frequency and sound pressure. Loudness is measured in Phon and the A-weighting curve is based on the 40 Phon iso-loudness curve, corresponding to the sound pressure level at different frequencies where a sound is perceived as being as loud as a pure tone at 1 kHz at 40 dB SPL. The 40 Phon iso-loudness curve roughly parallels the shape of the human audiogram, but is by definition offset with 40 dB. C-weighting on the other hand was introduced to quantify the loudness of more intense sounds and is modelled to roughly follow the 100 Phon iso-loudness curve, the argument being that this weighting would be better at capturing the perceived magnitude of loud noise and presumably also better reflect the perception of levels that will cause impact. This assumption, however plausible it may seem, is nevertheless essentially without experimental support and C-weighting, even of loud sounds, have been replaced by A-weighting in almost all human community noise regulations (e.g. European Commission, 2000, 2002). Even for short impulsive sounds, such as gunshots, where C-weighting is still recommended when recording peak levels, A-weighted levels correlate significantly better with annoyance, as rated by human observers, than does the C-weighted levels (Meloni and Rosenheck, 1995; Vos, 2003). There is thus a need for a re-evaluation of the justification for using C-weighting as basis for frequency weighting in marine mammals. NOAA seems to have embraced the need for such a re-evaluation in their recently proposed injury criteria (NOAA, 2013), based on new experimental results from two bottlenose dolphins, noting that one displayed substantial hearing loss (Finneran and Schlundt, 2013). However, audiograms differ among different species, especially with respect to the specific frequencies of greatest sensitivity. Thus, the weighting should ideally be species specific, or (as empirical data is extremely limited) at least by species grouped according to actual similarity of their audiograms across all frequencies, rather than by the more generalised functional hearing groups presented by Southall et al. (2007) and largely adopted by NOAA (2013).

A second argument put forward by Southall et al. (2007) for a C-weighting analogue for marine mammals is that this weighting curve is wider than the A-weighting curve, thus its use results in conservative criteria. The reasoning being that weighted sound levels are, if anything, overestimated, which produces stricter than needed protection of the animals. However, this reasoning has

some problems: there is a fundamental difference between frequency weighting used to establish exposure criteria from experimental data and frequency weighting used to compare a given exposure to an already established exposure criterion (Fig. 1). These two applications of weighting should each be undertaken in a way that provides conservative judgements to the benefit of the animals in situations where data is scarce. However, what may not be immediately evident is that a clearly conservative approach to weighting when establishing limits could be the opposite when applying an already established limit to assess impact and vice versa.

No matter how limited information is available, an exposure limit must be based on some experimental data. Thus, there are two sides to an exposure limit (Fig. 1): establishment of the limit from data (research side) and use of the limit in assessments (application side). Frequency weighting can (and should) be performed on both sides. Ideally, the two weighting functions are identical and match the actual weighting function of the exposed auditory system. This is by and large commonly agreed to be the case for A-weighting in humans as this function is based on overwhelmingly large experimental and clinical data sets. However, the two functions need not be identical. Realising how little we know about hearing and effects of noise in marine mammals, it makes sense to be conservative if there is doubt about the true form of the function, i.e. deliberately select a weighting function that is more likely to lead to tighter regulation of activities than actually needed (the precautionary principle). However, it turns out that the two weighting functions are not symmetrical when it comes to errors. A conservative weighting function applied on the research side turns out to be the opposite on the application side. To realise this, consider the following relationships:

$$L_{\text{unweighted}} \geq L_{\text{wide}} \geq L_{\text{actual}} \geq L_{\text{narrow}}$$

where  $L$  denotes sound pressure level of any sound weighted by different weighting functions, of which one corresponds to the actual weighting in the ear, one is too wide and one is too narrow.

The unweighted (broadband) sound pressure level will always be equal to or greater than any weighted sound pressure level. The more restrictive (or the narrower) a weighting function is, the more sound energy is filtered out and the lower the weighted sound pressure level will be. When going from data to exposure limit (the research side in Fig. 1) the precautionary principle dictates the use of a filter not wider than the actual weighting in the ear of the animal, i.e. to the narrow side. This will lead to an underestimation of the exposure limit, i.e. below the level where effects actually incur. In contrast, when assessing impact (the application side in Fig. 1), where a given sound is compared to the established exposure level, the precautionary principle dictates the use of a filter that is at least as wide as the actual filter. This will lead to an overestimation of the potential impact of the assessed

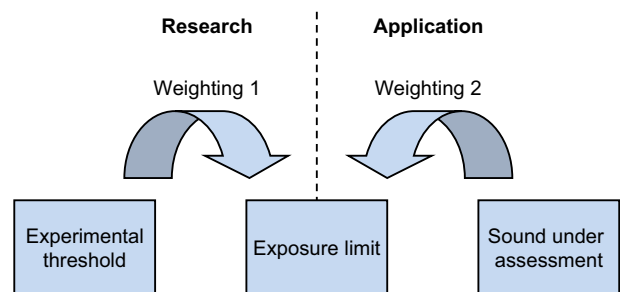


Fig. 1. Frequency weighting must be performed twice; first when establishing the exposure limit from data (research side), then when exposures are subsequently compared to the exposure limit (application side).

sound, i.e. the sound will be regulated to a level below the actual level at which effects incur.

Fig. 2 illustrates the consequences of selecting different combinations of weighting filters. The white spot in the centre illustrates the desirable point where frequency weighting actually conforms to the physiology of the ear. To the upper left is the point where a narrow filter (or weighting) is used in establishing the criterion and a wide filter (or weighting) used in assessments. This leads to certain over-regulation, which is not desirable, but it is at least safe to the animals. Opposite this point, in the lower right, is the least desired combination, leading to certain under-regulation. Along the lower left to upper right diagonal are the combinations with uncertain outcomes. In the best cases biases will cancel out, leading to the most appropriate level of regulation, but it is much more likely they will lead to errors to either side.

Now, returning to the suggested criteria of Southall et al. (2007), they discussed only the application side of Fig. 1, comparing sounds to the established exposure limits. When doing so, they deliberately selected the shape of the weighting function to be wider than they actually believed it to be, in correspondence with the precautionary principle, as described above. However, in that process they introduced an underestimation on the research side of Fig. 1. The result of this is that the proposed tolerance limits fall somewhere into the upper right part of Fig. 2, with an uncertain outcome, i.e. it is unclear whether the net result is underprotection of animals or overregulation of activities.

A few examples can illustrate this problem. First the simplest situation, where an exposure limit is established for a particular sound, for example an air gun pulse, and this limit is later used to evaluate exposure to the very same sound in a different context. In this case it does not matter if the weighting function is not completely correct, as the error committed on the research side cancels out completely with the error committed on the application side. A second example would be a sound outside the range of best hearing is used to establish an exposure limit which is then later used to assess potential impact of a second sound well within range of best hearing. The signal under assessment is compared to a criterion which is set too high, because a weighting function too wide was used. The weighted level of the sound used to establish the

limit will be too high ( $L_{\text{wide}} > L_{\text{actual}}$ ) because energy outside range of best hearing is included. In contrast, the weighted level of the sound being assessed against the exposure limit will be virtually identical to the unweighted level, as it falls within range of best hearing ( $L_{\text{wide}} = L_{\text{actual}}$ ). The implication is that the sound under assessment will only be considered to have an impact at levels that are substantially higher than is actually the case, i.e. underregulation. The third example is the reverse, an exposure limit based on exposure to a sound well within range of best hearing of the animal, but used to assess potential impact of a second sound which is outside range of best hearing. Here the signs switch, resulting in over-regulation. From these examples it follows that care should be taken when an exposure limit established with one type of sound is used to assess impact of another sound, in particular if the exposure limit is determined with sounds outside range of best hearing.

It should be noted that the proposed NOAA (2013) criteria have attempted to address these issues by adjusting the frequency weighting curves based on Finneran and Jenkins (2012). However, NOAA (2013) still uses the same weighting on both sides of the process, which means that the net result is not guaranteed to be conservative. Moreover, the use of this single-species audiogram-based weighting for a large number of different species with varied hearing capabilities is likely to produce errors of judgement. For example, any criteria based on low frequency sounds would be heavily down-weighted using the bottlenose audiogram, producing levels that would be far below those actually experienced by animals with better lower-frequency hearing, such as killer whales (Szymanski et al., 1999). The NOAA (2013) draft criteria violates the consistency further by extrapolating criteria values established for one hearing group under one weighting curve (based on bottlenose dolphin data) for use with other groups that then use different weighting systems.

### 2.3. PTS as basis for definition of limits to harmful exposures

Once a weighting standard is defined and accepted, exposure limits need to be established based on some threshold of impact, in this case for injury. What exactly constitutes an injury for a cetacean is ultimately more a legal than a scientific question. Nevertheless, direct injury involves sound exposures that are high enough to temporarily or permanently impair the function of some parts of the body of the exposed animals. In extreme cases, such as exposures to underwater explosions at close range, the injuries may be fatal. In the opposite end of the severity scale are subtle neurological effects and temporary and permanent increases in hearing thresholds (TTS and PTS, respectively).

While the permanent loss of hearing, PTS, definitely constitutes injury, the situation is less clear for TTS. TTS is by definition a temporary event, but recent experiments on terrestrial mammals suggest that the auditory system may not recover fully from severe TTS despite its definition to the contrary. TTS does not involve a destruction of hair cells, but seem instead to be caused at least in part by swelling of cochlear nerve terminals similar to that resulting from excitotoxicity (Kujawa and Liberman, 2009). It has furthermore been shown that levels causing reversible TTS can also lead to permanent degeneration of cochlear afferent nerves (Kujawa and Liberman, 2009). So while the hair cells may remain intact after a severe TTS, neurological degeneration may compromise more complex auditory processing and lead to a reduction of stimulus encoding under noisy conditions, tinnitus and hyperacusis. It is also possible that the damage from repeated severe TTS can ultimately accumulate to form a neurologically-based PTS (Kujawa and Liberman, 2009). This information was not available at the time when Southall et al. (2007) was made but its relevance

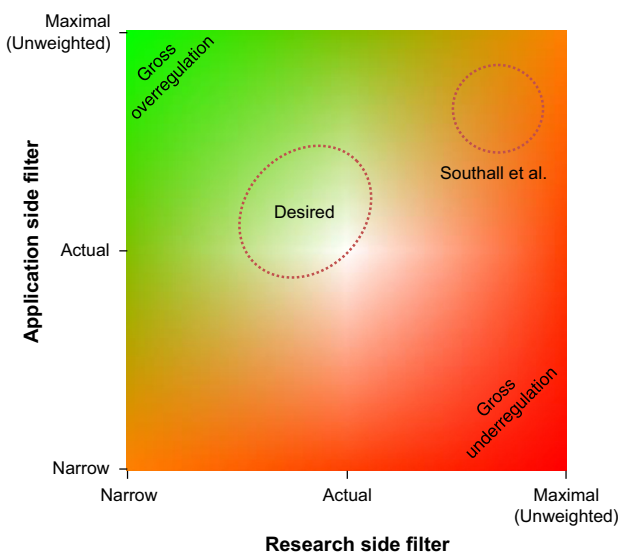


Fig. 2. Consequences of using different combinations of weighting functions in establishing criteria (horizontal axis) and assessing noise against criteria (vertical axis). Green indicates conservative areas, red indicates liberal areas. White spot in centre is the optimal point where weighting filter is exactly equal to the true weighting performed by the auditory system.

has been acknowledged later, for example in the NOAA (2013) draft guidelines.

Regardless of whether TTS or PTS thresholds are used to define injury thresholds from noise exposures, onset of TTS must be measured reliably and in a way that is representative for young, healthy animals with good hearing in the wild. This is because empirical data on PTS onset in marine mammals is extremely limited, for very good ethical reasons. Thus, the injury criteria as developed by Southall et al. (2007) are based on estimated thresholds for inducing PTS. Southall et al. (2007) assumed that PTS occurs at the same sound exposure levels that generate 40 dB of TTS, as is the case for terrestrial mammals (e.g. Ahroon et al., 1996). Accordingly, PTS onset values were extrapolated from observed TTS onset using TTS growth rates from chinchillas or humans, except in the one case (pinnipeds in air) where empirical data from marine mammals required use of a more precautionary value (see Southall et al., 2007).

Despite very careful and meticulous experimental approaches (e.g. Finneran et al., 2002; Nachtigall et al., 2003), the early TTS studies made on belugas and bottlenose dolphins are, potentially affected, at least to some degree, in two ways that may lead to an overestimation of levels required to generate TTS. Firstly, many earlier studies were conducted under noisy conditions in sea pens (Nachtigall et al., 2003, 2004) or an artificial masker was deliberately introduced to stabilize noise levels (Schlundt et al., 2000; Finneran et al., 2002), both approaches that lead to derivation of masked hearing thresholds set by the ambient noise. This means that the stimulus level required to induce onset for TTS could be overestimated because the true baseline hearing threshold was never established. The problem has in some cases been addressed by doing TTS experiments in quieter tanks (e.g., Finneran et al., 2002).

The use of quiet tanks does not, however, preclude a potential second problem: Both dolphins used in the Finneran et al. (2002) TTS experiment were older animals that had elevated hearing thresholds in the relevant frequency band; about 10 dB above the lowest thresholds measured for younger animals of the same species (Johnson, 1968b). It thus may be reasonable to question if the TTS thresholds derived for bottlenose dolphins and belugas based on data from older animals are also representative for young, healthy animals. There is limited and inconclusive information available, even from terrestrial mammals, on the possible effect of age on susceptibility to TTS (summarized in Finneran and Schlundt, 2010). Several recent studies on porpoises (Lucke et al., 2009; Popov et al., 2011; Kastelein et al., 2012, 2013) have demonstrated TTS at SELs considerably lower than reported for bottlenose dolphins and belugas. Future studies will hopefully elucidate if such differences are indeed species specific or whether they at least in part are related to different methodologies, masking levels or age and health of the used animals. Irrespectively, it is clear that the reliability of TTS extrapolations from a limited number of studies on a few individuals critically hinge on any problems with these studies.

### 3. Exposure criteria for behavioural responses

Exposure criteria for behavioural responses to underwater sound were discussed at length by Southall et al. (2007), but in the end they concluded that it was premature to establish general limits for safe exposure. This conclusion is reached in prudent realisation that the experimental results upon which such criteria should rest are too scattered across taxonomic groups, types of sounds/noises and contexts under which observations or experiments were conducted. NOAA's (2013) proposed guidelines also acknowledge this by limiting their bounds to injury. Behavioural

criteria are, however, also under production by NOAA. Additionally, Finneran and Jenkins (2012) presented behavioural response functions specifically relevant to military activities, such as sonar.

Instead of setting firm criteria, Southall et al. (2007) presented a framework for further progress towards establishing criteria for behavioural disturbance, as defined by the MMPA. The core of this framework was a list of behavioural responses that were scaled by their perceived severity. However, because this framework is tailored specifically to U.S. legislation (specifically the MMPA definition of impact: a "take"), it is difficult to generalise these recommendations (and those of Finneran and Jenkins, 2012) to other legislative frameworks, especially those pertinent to European countries. Any such attempts must, in our view, consider two important points. Firstly, a response severity scale for behavioural reactions, while useful in many ways, cannot be used directly or exclusively to infer the severity of actual impact. This is primarily due to the fact that such responses may only provide a relative indication of the immediate unpleasantness of the noise to the animal that is not necessarily suitable for assessing the long term fitness or population consequences of the noise exposure. Secondly, the use of such a scale implicitly presumes that an absence of any detectable reaction can be equalled with an absence of impact, which may not necessarily be the case.

#### 3.1. Scaling of impact with severity of response

The use of behavioural reactions to infer impact from noise exposure is inherently attractive to legislators and managers (as well as researchers) as the indirect and long-term consequences are, in most cases, beyond our present capability to study in wild animals, while behavioural responses are easier to observe. Southall et al. (2007) follow this approach in their review of the diverse literature by classifying the various experimental results in relation to a response severity scale. This response severity scale ranks behavioural reactions in nine steps from mild reactions (e.g. orientation response to acoustic stimulus), through medium reactions (e.g. short interruption of feeding behaviour) to strong reactions (e.g. startle and outright panic). As already discussed by Ellison et al. (2011), there are, however, a number of inherent problems with such a severity scaling.

Most importantly, Ellison et al. (2011) discussed the need for behavioural responses to be considered in terms of their audibility to the receiving animals. Furthermore, they noted that several contextual factors can be highly influential in determining behavioural responses, especially at lower severities. Such contextual factors include prior experience of the receiving animal, its current state (i.e., health and motivation) and activity, and its spatial relation to the sound source. In any case, even if these factors are considered, the focus of the scaling is fundamentally based upon the immediate perception by the animal, without any consideration for longer-term changes in behaviour or the exposure history of the animal. As a result, any estimation of the ultimate population-level impact based primarily on the severity of immediate responses, risk to establish a false proportionality between scores on the severity response scale and population impact. While such proportionality was not implied by Southall et al. (2007), it is tempting to make this inference and hence use the severity scores as a proxy for population impact.

In fact, a general correlation between impact in terms of fitness consequences with the severity of the immediate behavioural reaction is not supported. For example, some of the behaviours scoring very high on the severity scale are particularly immediate and short-term reactions to acoustic exposure, for example alarm responses and aggression and/or defensive responses. Such reactions can be very dramatic, but their short-term nature means that

many are unlikely to have long-term fitness consequences (e.g. Finneran and Jenkins, 2012, although see Götz and Janik, 2011).

Accordingly, it is reasonable to ask if fitness consequences can in any way be assessed through short-term studies of behaviour (Beale and Monaghan, 2004; Bejder et al., 2006; Lusseau and Bejder, 2007; Bateson, 2007; Beale, 2007; Johnson et al., 2009). While there is general agreement about the overall mechanisms leading from changes in behaviour to population level effects, most of which are captured in the PCAD model (National Research Council, 2005), there is an increasing realisation that in most cases it is impossible to establish these connections for real world systems from studies of behaviour alone. This realisation has led to new initiatives, centred on agent based models, such as the more general PCoD (Population Consequences of Disturbance) model (Lusseau et al., 2012; Nabe-Nielsen et al., 2013). Such models may incorporate some of the contextual elements required for better interpretation of behavioural responses, such as health and motivational state. However, they typically require large amounts of data, are very sensitive to all the various uncertainties of the input data, and are generally in their infancy.

### 3.2. Absence of detected reaction equals absence of impact

One problem of presuming proportionality between severity index scoring of behavioural responses and population impact arises when the logic is reversed through the concept that a lack of reaction implies that there has been no impact. Such a conclusion is clearly unsupported, but rarely, if ever, articulated directly (Madsen et al., 2006). However, this problem must always be kept in mind when it comes to actual implementation of legislation. For example, the failure to detect an impact may simply be because the behavioural reactions were too subtle, to be quantified with available parameters or occurred out of view as is often the case with diving animals and thus overlooked (Miller et al., 2009). Similarly, experimental design may be at fault, with behavioural reactions of one type undetected and unreported in a study focusing on responses of another kind. Furthermore, even if the study is collecting the right data on the right response, it is entirely possible that small sample sizes will produce a null result due to a lack of statistical power. Behavioural responses may also go unnoticed if they are simply changes in behavioural state that might otherwise be considered normal (e.g., a change from foraging to travelling or resting; Williams et al., 2006).

Another alternative that is unrelated to methodology is the possibility that there is in fact no obvious external reaction at all. However, it remains entirely possible that one or more physiological and psychological responses can take place, without any external symptoms. Such reactions include acute, prolonged or repeated stress responses (Fair and Becker, 2000; Wright et al., 2007a,b), diversion of attention that may lead to increased susceptibility to predators or other human interactions; (Todd et al., 1996; Wright et al., 2013; see also Nielsen et al., 2012) or failure to take advantage of foraging opportunities. These all have the potential to lead to long-term fitness consequences under certain situations or conditions.

The last and potentially most widespread and important way for noise to generate fitness effects without behavioural responses is through masking of other acoustic stimuli of importance to the animals (Møhl, 1980, 1981; Clark et al., 2009). The physiology of masking can and has been studied in captivity (e.g. Johnson, 1968a; Erbe and Farmer, 1998). However, while a calculation of communication range reduction is a useful proxy for impact (e.g., Møhl, 1980, 1981; Clark et al., 2009; Jensen et al., 2009; Hermanssen et al., 2014), direct assessment and quantification of masking effects in wild animals has so far proven impossible. The main reason for this is that masking, by definition, represents an

elevation of detection thresholds. Resulting consequences are thus typically the absence of behaviours, which are not outwardly detectable, such as failure to detect a conspecific, a potential prey or a potential predator, one way or the other leading to a decrease in fitness. The methodological difficulties involved in quantifying masking effects on fitness means that this effect has largely been ignored in legislation and it is clear that further progress in understanding of how masking affects real animals in the real world is required.

Accordingly, there are many good reasons to support the development of a fitness-based index for impact, but it is also clear that this may not be realistic at present. Therefore, due to the abovementioned difficulties in applying a response severity scale, we suggest a more simplified analysis in the interim. We propose that the onset of negative phonotaxis is a suitable starting point for discussion of criteria for fitness impacts of behavioural reactions to noise. Sustained negative phonotaxis manifests itself in a decreased abundance of animals in a smaller or larger area around the sound source and the impact of the sound can be quantified by the deterrence distance.

Such avoidance reactions remain relatively easy to observe and constitute a temporary loss of habitat. If occurring during times of foraging or reproduction, such loss of habitat can carry some level of fitness cost (even if only realised through a slight reduction in mating opportunities or calf survival) and, depending on the circumstances, could potentially affect the long term population size in an area. In other situations, the fitness cost of a displacement could be minimal, in particular in situations where the population is not limited by food but other factors such as bycatch or harvest. Similarly, avoidance on migration routes may carry a much lower cost, if any, and thus could be ignored for these reasons. Conversely, there are situations where animals may be unwilling or unable to avoid the source, but may still suffer fitness-related consequences of acoustic exposure (for example masking). Thus, we are not suggesting that avoidance be used directly as criterion to establish onset of fitness impact, as there need not be a direct correlation between the two. Instead, we proffer that the use of avoidance onset as a proxy from which the onset of fitness impact can be calculated offers a simple, and therefore importantly, an operational criterion, as the effect can be detected readily by standard survey methods and passive acoustic monitoring and without the need for long-term tracking of individuals.

## 4. Case study – the harbour porpoise

To illustrate many of the concepts presented above, we discuss an example of how they might be applied to an animal of particular concern in Europe: the harbour porpoise (*P. phocoena*). The harbour porpoise is the most common cetacean in European shelf waters (Hammond et al., 2002, 2013), and it is generally believed to be one of the most sensitive species of marine mammals when it comes to acoustic disturbance, making it a key species in discussions of impact of increasing anthropogenic noise in the oceans. Despite this, there are no commonly accepted limits for noise exposure for this species. Empirical data available to Southall et al. (2007) was insufficient to establish usable exposure limits. Accordingly, while Southall et al. (2007) extrapolated their criteria for HF cetaceans (such as porpoises) from MF criteria, they did so clearly stressing the need for additional experimental data and hinted that the real limits were likely to be lower for the HF-cetaceans than the MF-cetaceans. Fortunately, since 2007 several studies have produced new information on the susceptibility of porpoises to injury and behavioural disturbance. Much of this has been considered in the NOAA (2013) draft criteria, as well as other government reports and criteria (e.g., the below-mentioned

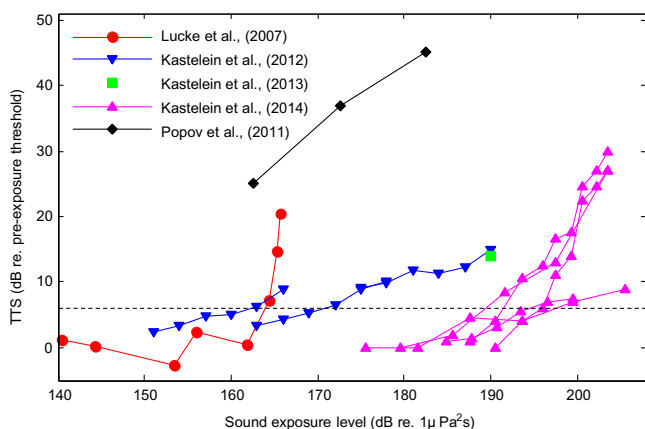
Lucke et al., 2009, formed the core of the German standard for harbour porpoise exposure to pile driving noise). Here we critically review this new information in an attempt to discuss the first suggestions for generalised exposure limits for this species. This represents a methodological example only and it is not intended that these criteria be directly extrapolated to other species.

#### 4.1. Noise induced injury (TTS) in porpoises

Five studies are now available on TTS in porpoises. Lucke et al. (2009) measured TTS at 4 kHz, induced in a single animal by exposure to short, broadband air-gun pulses and found that 6 dB of TTS could be induced by a single pulse with a received un-weighted sound exposure level (SEL) of 164 dB re. 1  $\mu\text{Pa}^2\text{s}$  (Fig. 3). Peak frequency of the pulse was below 500 Hz but with substantial energy at higher frequencies up to and possibly above 20 kHz.

Kastelein et al. (2012) exposed a different animal to octave band noise centred at 4 kHz for longer periods of time (between 7.5 min and 4 h). Three different sound pressure levels were used and TTS were induced by all three levels at different exposure times, indicating a trade-off between intensity and duration in line with what has been observed for bottlenose dolphins (Nachtigall et al., 2003). TTS as a function of SEL is shown in Fig. 3. Six dB of TTS could be induced by an SEL of 163 dB re. 1  $\mu\text{Pa}^2\text{s}$  for the low-intensity noise and 172 dB re. 1  $\mu\text{Pa}^2\text{s}$  for the medium-intensity noise. The high-intensity noise induced 9 dB of TTS at the lowest SEL of 175 dB re. 1  $\mu\text{Pa}^2\text{s}$ , but with the curve as a near-perfect continuation of the mid-intensity curve. The disparity of these two curves on one hand and the low-intensity curve on the other, implying higher sensitivity to low-intensity noise, is unexplained and counter-intuitive, but may have to do with the uncertainties involved in estimating the received sound exposure of a free-swimming animal in a net pen over very long durations.

In a third study, Kastelein et al. (2013) induced 14 dB of TTS in a harbour porpoise by exposing it to a long, continuous 1.5 kHz tone with a total SEL of 190 dB re. 1  $\mu\text{Pa}^2\text{s}$ . As only one exposure level was used no threshold for eliciting TTS onset could be determined. Extrapolating to an onset threshold from a single TTS point is questionable due to the extreme non-linear nature of the growth of TTS with SEL (cf. the strong disparity between the steepness of the curves in Lucke et al., 2009 and Kastelein et al., 2012). Threshold for eliciting 6 dB of TTS could thus have been anywhere between 172 dB re. 1  $\mu\text{Pa}^2\text{s}$  (using the shallow slope from Kastelein et al., 2012) and 188 dB re. 1  $\mu\text{Pa}^2\text{s}$  (using the steep slope of Lucke et al., 2009).



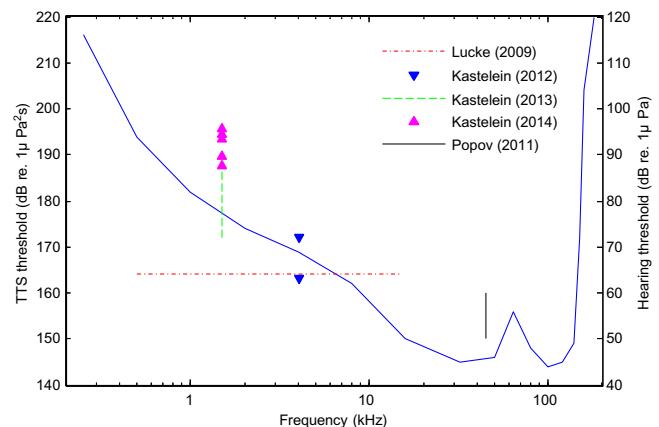
**Fig. 3.** Temporary threshold shifts (TTS) as function of sound exposure level (cumulated energy) of the fatiguing noise. Results from four different studies, of which one (Popov et al., 2011) was on finless porpoise and the three others on harbour porpoises.

In a very similar study Kastelein et al., 2014 induced up to about 30 dB of TTS by exposure to frequency modulated sweeps between 1 and 2 kHz in various combinations of duration, intensity and duty cycle, with combined SEL between 175 and 205 dB re. 1  $\mu\text{Pa}^2\text{s}$ . Here thresholds for TTS were found between 188 and 196 dB re. 1  $\mu\text{Pa}^2\text{s}$ .

A fifth study by Popov et al. (2011) was conducted on a different, but closely related species, the Yangtze finless porpoise (*Neophocaena asiaeorientalis*). Very little work has been done on the auditory function of this species, but direct comparison of their hearing to harbour porpoises has been made with respect to auditory filters (Popov et al., 2006) and no differences were seen between the two species suggesting no major differences in their overall hearing abilities. The study of Popov et al. (2011) was on one animal exposed to high levels of 3 min pulses of half-octave band noise, centred at 45 kHz. The lowest exposure level of 163 dB re. 1  $\mu\text{Pa}^2\text{s}$ , induced 25 dB of TTS (Fig. 3) that in turn precluded determination of a threshold for eliciting TTS. Clearly the threshold must have been below 163 dB re. 1  $\mu\text{Pa}^2\text{s}$  and a simple downward extension of the curve suggest that it could be as low as 145 dB re. 1  $\mu\text{Pa}^2\text{s}$ . On the other hand, as the growth of TTS with increasing exposure is highly non-linear (Lucke et al., 2009) the threshold could equally well be only slightly lower than 163 dB re. 1  $\mu\text{Pa}^2\text{s}$ . It thus seems reasonable only to conclude that the TTS threshold of the finless porpoise was likely somewhere in the range from 145 to 160 dB re. 1  $\mu\text{Pa}^2\text{s}$  at 45 kHz.

#### 4.2. Proposed exposure limit for TTS

Although there are only a few data points and these are derived by two different methods (operant conditioning and evoked potentials), the thresholds and estimated threshold ranges for eliciting TTS in the five studies mentioned above show a clear trend of lower thresholds at higher stimulus frequencies (Fig. 4). This tendency is more or less in parallel with the shape of the audiogram, strongly suggesting that summed SEL alone is not a good predictor of TTS thresholds, and that some form of frequency weighting would be appropriate. This finding is in line with more recent data on TTS in bottlenose dolphins (Finneran and Schlundt, 2013), but in contrast to the (precautionary) M-weighting curve suggested by Southall et al. (2007). It is clearly important to further test this notion by conducting further TTS experiments with signals of different (higher) frequencies and with different bandwidths. That will enable derivation of the most appropriate weighting function



**Fig. 4.** Sound exposure levels (SEL) required to elicit 6 dB of TTS taken from the four studies (broken line in Fig. 3). Estimated threshold range only is indicated for the data from Popov et al. (2011) and Kastelein et al. (2013). Stimulus in Lucke et al. (2009) was a broadband airgun pulse, hence frequency is also indicated only as a range.

and determine whether weighting with a curve resembling the inversed audiogram is appropriate or rather that curves based on iso-loudness contours should be used, such as those recently presented for porpoises (Wensveen et al., 2014).

Based on the limited information available at this point a preliminary exposure limit with respect to eliciting TTS in porpoise suggests itself from Fig. 4 as being an SEL of 100–110 dB above the porpoise pure tone hearing threshold at the relevant frequency. It is reasonable to ask whether it can at all be justified to compare thresholds expressed as SEL (a unit of energy) to thresholds expressed as SPL (a unit of pressure). Such a comparison can be made if SEL is viewed as a normalisation to a stimulus duration of 1 s ( $SEL = SPL + 10 \log(\text{dur})$ ) and recalling that the porpoise audiogram is valid for stimulus durations from 0.5 s to at least 5 s (Kastelein et al., 2010).

#### 4.3. Non-injurious, fitness-based impacts of noise in porpoises

Auditory scene analysis and detection of predators will be impacted, to some extent, by masking from anthropogenic noise, with some level of fitness impact. However, the consequences of such a masking for foraging and reproduction in porpoises are entirely unknown and thus presently impossible to assess. Due to the high frequencies involved (Møhl and Andersen, 1973; Kastelein et al., 2002; Hansen et al., 2008) in combination with propagation at these frequencies (DeRuiter et al., 2010) any effects are likely to be at short range (Hermannsen et al., 2014). In lieu of a more direct measure of fitness impact, we defer at this time to avoidance-based criteria, that would need to be applied with due precaution, as justified previously in the general discussion. A substantial number of experiments have been conducted on the impact of noise on the behaviour of harbour porpoises, both in captivity and in the wild. We focus on the field studies because a substantial number is available and because of the difficulties involved in translating the results of captive studies to wild populations (Wright et al., 2009). These difficulties include animals habituated to noise exposure or reinforced to tolerate noise, as well as limited ability for captive animals to express their full repertoire of behavioural reactions.

#### 4.4. Temporal weighting of sound pressures

A critical element in any comparison of different experiments involving exposure to sound is to ensure that measures of sound exposures of the individual studies are comparable. Especially for short duration signals there is the trivial, yet often overlooked issue of comparing peak (or peak to peak) sound pressures with root-mean-squared (rms) sound pressures and the critical choice of averaging time for rms-averages (see Madsen, 2005). Whereas there is good evidence that sound exposure level (cumulated energy) is the primary (albeit not the only) parameter correlating with induced TTS (Nachtigall et al., 2003; Kastelein et al., 2012, although see Popov et al., 2014 and Kastelein et al., 2014) it is much less clear which parameter is appropriate when it comes to behavioural reactions. There is some consensus on rms-average as the appropriate measure (NMFS, 2003), but the critical issue of averaging time has rarely been addressed (Madsen, 2005). However, as it is not unreasonable to suggest some link between audibility or loudness of the sound and the capability of the sound to affect behaviour, it makes sense to base the choice of averaging time on these factors as well. It is well known that audibility of short duration signals increase with duration, known as temporal integration (Green et al., 1957; Johnson, 1968b; Kastelein et al., 2010). Detection thresholds for short signals improve with about 3 dB per doubling of duration until a critical duration, termed the integration time is reached, above which a further increase in dura-

tion does not affect the threshold. Experimental data are in most cases successfully explained by a simple integrator model with exponential decay, the leaky integrator (Plomp and Bouman, 1959). This leaky integrator is the basis of the integration performed in sound level meters used in monitoring of noise exposure to humans, with a time constant (integration time)  $\tau$  of 125 ms recommended for short duration signals and referred to as the “fast average” (e.g. Anonymous, 1983). The choice of 125 ms as the time constant is based on human integration time data (e.g. Plomp and Bouman, 1959) and allows for a direct comparison of thresholds for signals of varying duration. The “rms fast average” ( $L_{\text{eq-fast}}$  in the following) can thus be calculated from the rms-average over the duration of the pulse ( $L_{\text{eq}}$ ) as:

$$L_{\text{eq-fast}} = L_{\text{eq}} + 10 \log(1 - e^{-d/\tau}), \quad (1)$$

where  $d$  is the duration in seconds and  $\tau$  is the time constant (0.125 s). By definition,  $L_{\text{eq-fast}}$  is a time average and is referenced to 1  $\mu\text{Pa}$  rms. The level can be interpreted as the rms sound pressure level ( $L_{\text{eq}}$ ) of a 125 ms signal of constant amplitude having the same energy as a 125 ms window of the signal.

Integration time has been determined also for a bottlenose dolphin (Johnson, 1968b), a beluga (Johnson, 1991), a harbour porpoise (Kastelein et al., 2010, shown here in Fig. 5) and a harbour seal (Kastelein et al., 2009). Although there is clearly an effect of frequency on integration time, these marine mammal data show reasonably good consistency with human data (see Kastelein et al., 2010). To test the leaky integrator model, we converted the thresholds of Kastelein et al. (2010) to  $L_{\text{eq-fast}}$  by Eq. (1) and normalised them with the audiogram to remove frequency dependent differences in thresholds. The normalised thresholds are plotted in Fig. 5b and show that thresholds, which on an absolute scale (Fig. 5a) span a range of nearly 80 dB, can be normalised to within  $\pm 3$  dB for the mean values (Fig. 5b). Stimulus frequency and duration can thus explain almost all variation in the threshold data. Based on this we recommend  $L_{\text{eq-fast}}$  as a standard measure also for marine mammals for comparing thresholds for signals of different duration.

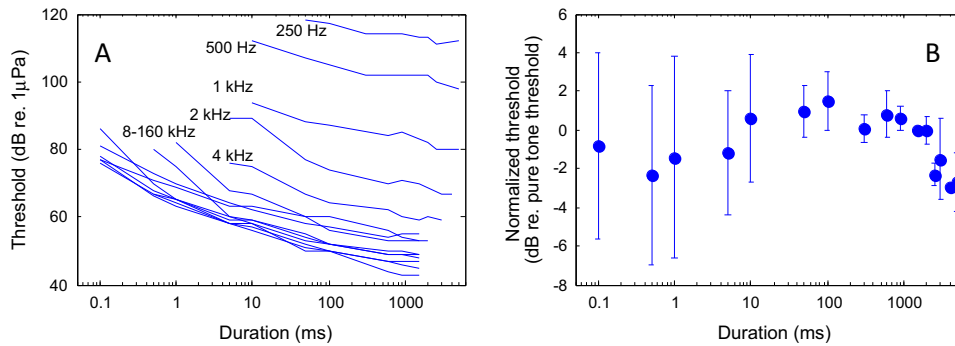
#### 4.5. Reactions to pingers

A number of studies have focused on reactions of porpoises to acoustic deterrent devices (pingers), two of which provide sufficient information to derive estimates of reaction thresholds. Both studies involved visually tracking porpoises around inactive vs. active pingers.

Culik et al. (2001) studied reactions to a PICE pinger, which produced an FM-sweeping, multiharmonic signal in the frequency range 20–160 kHz. Observations showed avoidance out to a distance of at least 200 m. The transducer of the PICE pinger is highly resonant around 55 kHz, which is thus the most powerful part of the signal spectrum and likely also the most audible part of the signal to porpoises. A measured (rms) source level of 145 dB re. 1  $\mu\text{Pa}$ , this equates to a reaction threshold at an  $L_{\text{eq-fast}}$  of 93 dB re. 1  $\mu\text{Pa}$  (rms), assuming spherical spreading and an absorption coefficient at 55 kHz of 17.8 dB/km. The duration of these signals was 0.3 s, which results in a correction to  $L_{\text{eq-fast}}$  from Eq. (1) of 0.6 dB.

Carlström et al. (2009; with additional detail in Berggren et al., 2002) studied reactions to a Dukane Netmark1000 pinger and found reaction distances over the range of 125–375 m. The fundamental frequency of the signal was 11 kHz, but the 4th harmonic at 43 kHz was likely the most audible component to porpoises. Source level of the 4th harmonic was estimated to be 129 dB re. 1  $\mu\text{Pa}$  (rms) and absorption at 43 kHz set to 12.5 dB/km. As above, the duration of signals were 0.3 s, yielding a correction to  $L_{\text{eq-fast}}$





**Fig. 5.** Relation between sound duration and detection thresholds of a porpoise. (A) Data from Kastelein et al. (2010), showing higher sensitivity to longer signals and higher frequencies. The vertical offset between curves reflects the audiogram with increasing sensitivity to higher frequencies. (B) Plot of the same data, converted to  $L_{eq-fast}$  (rms weighting with exponential decay,  $\tau = .125$  s) and normalised by the audiogram (threshold to 1.5 s tones). See text for further explanation.

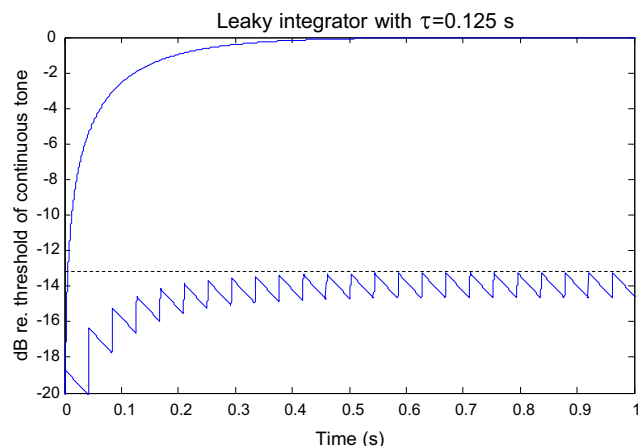
from Eq. (1) of 0.6 dB. Received  $L_{eq-fast}$  at a mean reaction distance of 250 m is thus estimated to have been 78 dB re. 1  $\mu$ Pa (rms).

#### 4.6. Reactions to seal scarers

Four studies are to our knowledge available on reactions to acoustic harassment devices (seal scarers), three of which are based on theodolite trackings from land, and the last (Brandt et al., 2012) is based on passive acoustic monitoring and aerial surveys.

Olesiuk et al. (2002) studied the reactions to a seal scarer (Airmar AHD, 10 kHz), and found a dramatic decline in observations when the seal scarer was on out to the maximum sighting distance of 3.5 km. Source level was stated to be 196 dB re. 1  $\mu$ Pa peak to peak, which can be converted to  $L_{eq}$  by subtraction of 9 dB ( $20 \log(2\sqrt{2})$ ), ratio between peak-peak amplitude and rms average for a sine wave), assuming that the signal was close to a sine wave pulse. Individual pulses were very short, 1.8 ms, but repeated in longer bursts with 25 pulses per second.  $L_{eq-fast}$  was estimated by modelling the sound pressure changes over subsequent pulses by a leaky integrator model (Fig. 6). The difference between the asymptotic level reached after about 10 pulses was 13 dB lower than the  $L_{eq}$  of a continuous signal. With an absorption coefficient at 10 kHz of 1 dB/km  $L_{eq-fast}$  was estimated to have been 97 dB re. 1  $\mu$ Pa (rms) at a range of 3.5 km.

In an almost identical study on a comparable seal scarer (Airmar dB II Plus), Johnston (2002) obtained similar results. Porpoises



**Fig. 6.** Output of leaky integrator with  $t = 0.125$  s when input is signal of Airmar ADH (1.8 ms pulses, 40 ms inter-pulse intervals), compared to output generated by a continuous signal beginning at  $t = 0$ . The difference between the two curves is the factor used to convert single pulse  $L_{eq}$  to  $L_{eq-fast}$ .

reacted to the 10 kHz signals out to at least 1 km (the limit of his observation field). Source level was given as 181 dB re. 1  $\mu$ Pa and a correction for short repeating pulses was made as above. Hence, the estimated  $L_{eq-fast}$  is estimated to have been 107 dB re. 1  $\mu$ Pa (rms) at a 1 km range.

The two most recent studies (Brandt et al., 2012, 2013) used a different type of seal scarer (Lofitech, 14 kHz, source levels 189 dB re. 1  $\mu$ Pa, pulse duration 0.5 s). Brandt et al. (2012) used passive acoustic monitoring (C-PODs) around the seal scarer and detected statistically significant reductions in porpoise acoustic activity out to distances of 7.5 km. These observations were supported by an aerial survey that showed that the porpoises had also vacated the area at least out to 7.5 km. The second study (Brandt et al., 2013) used the same seal scarer but observations were from a cliff by means of a theodolite. Although very few observations were made, reactions were seen at distances of at least 2.4 km from the seal scarer. The  $L_{eq-fast}$  at these ranges, and with an absorption rate of 1 dB/km, can be calculated to be 98 dB re. 1  $\mu$ Pa (rms) and 120 dB re. 1  $\mu$ Pa (rms) respectively.

#### 4.7. Reactions to pile driving

A number of studies have looked at reactions of porpoises to impact pile driving operations in connection with construction of large offshore wind farms. All have all been conducted by means of passive acoustic monitoring at various distances from the pile driving site. Tougaard et al. (2009) reported reactions to pile driving at Horns Reef 1 offshore wind farm at distances beyond 21 km. The peak energy in the pulses was at 500 Hz, although substantial energy was present at higher frequencies up to 100 kHz. Pulse duration was about 0.2 s, which corresponds to a correction of 1 dB to produce a  $L_{eq-fast}$  of 130 dB re. 1  $\mu$ Pa (rms) at 21 km from the piling site, using a source level of 237 dB re. 1  $\mu$ Pa peak to peak. Conversion from peak to peak measurements to  $L_{eq}$  for transient pulses is not trivial, but the data presented by Bailey et al. (2010) suggests that a conversion factor of 20 dB is appropriate for pile driving pulses.

Similar reaction distances (18 km) were observed by Brandt et al. (2011) at Horns Reef 2 offshore wind farm. Noise levels (single pulse SEL) were measured at 2.3 km, and extrapolated to 18 km, where porpoise presence was assessed by passive acoustic monitoring. The single pulse SEL can be converted to  $L_{eq}$  by adding 3 dB, taking into consideration that pulse duration was 0.5 s (as  $SEL = L_{eq} + 10 \log(d)$ ), which yields an estimated  $L_{eq-fast}$  of 149 dB re. 1  $\mu$ Pa (rms). This is most likely an overestimate of the actual level at 18 km distance, because of excess attenuation from a shallow sand barrier (3–5 m water depth) between the passive acoustic monitor and the construction site. Peak energy was around 200 Hz, but also with considerable energy at higher frequencies.

**Table 1**

Studies on harbour porpoise reactions to noise sources from which robust estimates of thresholds could be derived and included in Fig. 6.

Reference	Sound characteristics						Avoidance			
	Sound source	Peak frequency	Bandwidth	Signal type	Duration (ms)	Corr. to $L_{eq-fast}$	Source level	Distance	$L_{eq-fast}$	Comments
Culik et al. (2001)	Pice ADD	55 kHz	20–160 kHz	Sweep	0.3 s	–0.6 dB	145 dB re. 1 $\mu$ Pa rms	200 m <sup>a</sup>	93 dB re. 1 $\mu$ Pa	Assumed to react to 55 kHz part of sweep
Carlström et al. (2009)	Dukane Netmark1000 ADD	11 kHz	Narrow, with harmonics	Tonal pulse	0.3 s	–0.6 dB	138 dB re. 1 $\mu$ Pa rms <sup>b</sup>	125–375 m	78 dB re. 1 $\mu$ Pa	Reaction assumed to be to 4th harmonic <sup>c</sup>
Olesiuk et al. (2002)	Airmar AHD	10 kHz	Narrow, with harmonics	Tonal pulses	1.8 ms, 25 pulses/s	–13 dB	196 dB re. 1 $\mu$ Pa pp <sup>d</sup>	>3.5 km	97 dB re. 1 $\mu$ Pa	Minimum estimate of reaction distance
Johnston (2002)	Airmar dB II Plus AHD	10 kHz	Narrow, with harmonics	Tonal pulses	2.5 ms, 25 pulses/s	–13 dB	181 dB re. 1 $\mu$ Pa rms	>990 km	107 dB re. 1 $\mu$ Pa	Mean minimum approach distance
Brandt et al. (2012)	Lofitek AHD	14 kHz	Narrow, with harmonics	Tonal pulse	0.5 s	0 dB	189 dB re. 1 $\mu$ Pa rms	7.5 km	98 dB re. 1 $\mu$ Pa	Passive acoustic monitoring
Brandt et al. (2013)	Lofitek AHD	14 kHz	Narrow, with harmonics	Tonal pulse	0.6 s	0 dB	189 dB re. 1 $\mu$ Pa rms	2.4 km	120 dB re. 1 $\mu$ Pa	Very few observations
Tougaard et al. (2009)	Pile driving Horns Reef 1 OWF	1 kHz	Broadband, with strong LF emphasis	Impact pulse	0.2 s	–1 dB	237 dB re. 1 $\mu$ Pa pp <sup>e</sup>	>21 km	130 dB re. 1 $\mu$ Pa	Possible reaction to higher frequencies
Thompson et al. (2010) <sup>f</sup>	Pile driving Beatrice OWF	800 Hz	Broadband, with strong LF emphasis	Impact pulse	0.2 s	–1 dB	250 dB re. 1 $\mu$ Pa pp	<40 km	120 dB re. 1 $\mu$ Pa	No reaction observed
Brandt et al. (2011)	Pile driving Horns Reef 2 OWF	300 Hz	Broadband, with strong LF emphasis	Impact pulse	0.5 s	0 dB	252 dB re. 1 $\mu$ Pa pp	18 km	167 dB re. 1 $\mu$ Pa <sup>g</sup>	Possible reaction to higher frequencies
Dähne et al. (2013) <sup>h</sup>	Pile driving Alpha Ventus OWF	500 Hz	Broadband, with strong LF emphasis	Impact pulse	0.2 s	–1 dB	245 dB re. 1 $\mu$ Pa pp	25 km	150 dB re. 1 $\mu$ Pa	Possible reaction to higher frequencies
Tougaard et al. (2012)	Pile driving	1 kHz	Broadband, with LF emphasis	Impact pulse	0.2 s	–1 dB	186 dB re. 1 $\mu$ Pa pp	200 m	130 dB re. 1 $\mu$ Pa	Playback at reduced levels

<sup>a</sup> Estimated from their Fig. 2.<sup>b</sup> From Barlow and Cameron (2003), median value.<sup>c</sup> See spectrum in Berggren et al. (2002). 4th harmonic estimated to be 9 dB below fundamental.<sup>d</sup> Peak–peak values converted to rms by subtraction of 9 dB.<sup>e</sup> Conversion from peak–peak to rms by subtraction of 20 dB, cf. Bailey et al. (2010).<sup>f</sup> Acoustic measurements from Bailey et al. (2010).<sup>g</sup>  $L_{eq}$  at threshold overestimated, due to excess attenuation by sandbar.<sup>h</sup> Acoustic measurements from Betke and Matuschek (2010).

A third study at Alpha Ventus offshore wind farm (Dähne et al., 2013) gave comparable results with reaction distances around 25 km. Importantly, the associated aerial surveys supported the conjecture that porpoises leave the impact area rather than diminishing their vocal behaviour. Noise measurements 17 km from the construction site were used to calculate a received level at 25 km of around 155 dB re. 1  $\mu$ Pa peak (Betke and Matuschek, 2010). This can be converted to  $L_{eq-fast}$  by subtraction of 1 dB for pulse duration correction and 14 dB for conversion from peak to rms (+6 dB to convert from peak pressure to peak-to-peak pressure and –20 dB conversion factor from peak-to-peak to rms, from Bailey et al., 2010). This gives an estimated  $L_{eq}$  of 140 dB re. 1  $\mu$ Pa (rms). Peak energy was around 500 Hz, but again there was considerable energy at higher frequencies, making it unclear which part of the spectrum the animals reacted most strongly to.

In one additional study, Thompson et al. (2010) investigated reactions of porpoises to pile driving at a single station located 40 km from the piling site at the Beatrice offshore wind farm and could not detect changes in porpoise abundance. Sound pressure levels were measured out to distances beyond 40 km by Bailey et al. (2010), which allowed estimation of  $L_{eq-fast}$  of about 120 dB re. 1  $\mu$ Pa (rms).

#### 4.8. Scaled-down controlled exposure experiments

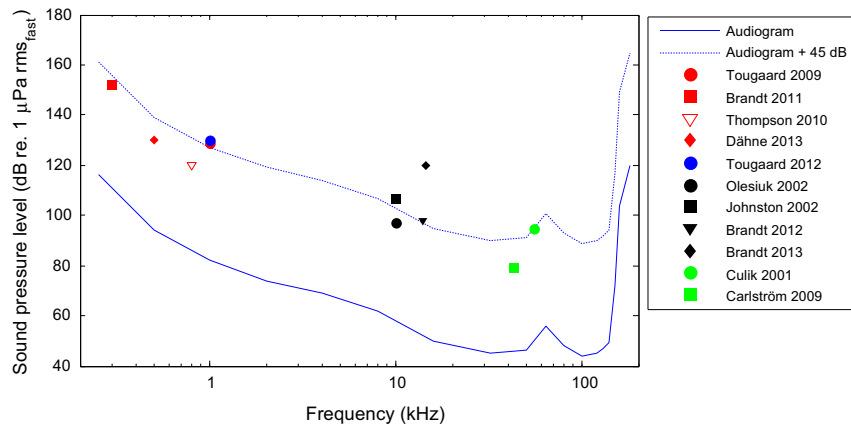
To overcome the difficulties of studying the large impact area in a real pile driving, Tougaard et al. (2012) conducted a play back experiment with pile driving noise but at reduced source level.

By this approach the animals could be tracked by theodolite and a deterrence distance of about 200 m could be determined corresponding to a received sound pressure level  $L_{eq}$  of about 130 dB re. 1  $\mu$ Pa (rms). Furthermore, it could be verified that animals did in fact leave the exposed area, something which could not be determined with certainty in the majority of the studies of real pile driving events as they relied exclusively on passive acoustic monitoring.

#### 4.9. Proposed exposure limit for fitness-based impacts

In summary, avoidance of mostly ‘low-frequency’ pile driving events appears to be on the order of 20 km, while reactions to ‘mid-frequency’ seal scarers are at ranges on the order of magnitude below – between 1 and 7.5 km, and response distances to ‘high-frequency’ pingers another order of magnitude lower still – around 200 m. The actual exposures in these three groups differ in several ways: source level, peak frequency, pulse duration and pulse repetition rate. In particular, the signals in two of the seal scarer studies (Olesiuk et al., 2002; Johnston, 2002) differ from those in the rest by being very short, but rapidly repeated pulses. Adjustment of sound pressure levels to  $L_{eq-fast}$ , however, results in thresholds very close to the threshold found by Brandt et al. (2012). The threshold found by Brandt et al. (2013) is 10–20 dB higher, but very few animals were observed in that study making it quite likely that the response threshold was overestimated.

The duration of sounds in the other studies were all 0.2 s or more with low repetition rates of one per second or less, resulting



**Fig. 7.** Estimated thresholds for behavioural reactions (negative phonotaxis) from a range of field studies involving pile driving (red symbols), scaled-down pile driving noise (blue), seal scarers (black) and pingers (green). For pingers and seal scarers the  $x$ -axis indicates the frequency which most likely was the most audible to porpoises. For pile driving noise the  $x$ -axis indicates the peak frequency of the noise. Porpoises most likely reacted to energy at higher frequencies than the peak frequency, which means that these points should likely be shifted to the right by an unknown amount. All estimated sound pressure levels at threshold were converted to  $L_{eq-fast}$  (see text and Table 1). Open symbol indicates one study where no reaction was seen at the indicated sound pressure level (Thompson et al., 2010). Solid line is the porpoise audiogram (Kastelein et al., 2010) and broken line the audiogram offset upwards by 45 dB.

in very little correction (0–1 dB) to convert thresholds to  $L_{eq-fast}$ . Comparing responses of different signals in this way ignores a range of important differences between signals, experimental methods and exact definitions of response threshold. Despite these caveats, the overall emerging picture is surprisingly consistent across the 11 studies, showing a decreasing threshold of reaction with increasing peak frequency of the signals, parallel to the sloping audiogram but offset by 40–50 above the hearing threshold (Fig. 7). This strongly indicates that the two most important factors determining whether porpoises react or not are the stimulus duration/repetition rate (captured by the  $L_{eq-fast}$ ) and the level above the hearing threshold, also referred to as the sensation level. As a consequence, we propose an exposure limit for negative phonotaxis to be  $L_{eq-fast}$  45 dB above the hearing threshold. This limit does not on its own provide information about possible fitness consequences of the deterrence, as the duration of the period where animals are deterred must be taken into consideration, as should the importance of the area to the animals. The onset of fitness impacts could thus be below this threshold for long-term deterrence from important habitats.

This conjecture that response distances can be predicted from  $L_{eq-fast}$  levels in combination with the audiogram, should clearly be tested in future experiments. Such experiments should determine response distances and thresholds to more complex signals, covering broader frequency ranges, of different durations and with more complex temporal structure and then compare the results with predictions of made from the model suggested here. If the model can hold up against new experimental evidence, then in line with the situation for TTS-thresholds, it becomes important to determine the most appropriate shape of a frequency weighting function, distinguishing between inversed audiograms (e.g. Terhune, 2013) and loudness-based approaches (Finneran and Schlundt, 2013 and Wensveen et al., 2014).

It remains important, however, to note that behavioural responses will occur below the levels of exposure required to generate negative phonotaxis, as could other potential fitness-related impacts (such as masking of predator signals), at least in some scenarios. However, there is not yet enough scientific information to determine exactly when fitness impacts in porpoises may start to occur. Similarly, we acknowledge that the extent to which fitness impacts and population level impacts can be tolerated by society

remains a policy-based decision. Consequently, the value of the adjustment factor needed to convert negative phonotaxis onset to fitness-based impact criterion currently relies on a blend of the scientific data presented here with the appropriate application of management objectives.

#### Acknowledgements

This review was partly funded under a contract with Energi-net.dk and forms part of a background report to the Danish Energy Authority on guidelines for conducting environmental impact assessments of sub-sea cable installations, and partly funded under the EU LIFE11 Project ENV/SE/841 BIAS. AJW conducted part of this review during his Ph.D. studies, which were co-funded by Aarhus University, the Danish Ph.D. School on Sense organs, Nerve systems, Behaviour, and Communication (SNAK) and the Carlsberg Foundation. Magnus Wahlberg, Peter Tyack, Jim Finneran and two anonymous reviewers are thanked for valuable inputs and comments to an earlier version of this manuscript. We finally thank Mark Johnson for valuable discussions.

#### References

- Ahroon, W.A., Hamernik, R.P., Lei, S.-F., 1996. The effects of reverberant blast waves on the auditory system. *J. Acoust. Soc. Am.* 100, 2247–2257. <http://dx.doi.org/10.1121/1.417934>.
- Anonymous, 1983. ANSI S1.4-1983 Specification for Sound Level Meters. American Institute of Physics, New York.
- Anonymous, 2001. ANSI S1.42-2001 Design response of weighting networks for acoustical measurements. American Institute of Physics, New York.
- Au, W.W.L., Moore, P.W.B., 1984. Receiving beam patterns and directivity indices of the atlantic bottlenose dolphin (*Tursiops truncatus*). *J. Acoust. Soc. Am.* 75, 255–262. <http://dx.doi.org/10.1121/1.390403>.
- Bailey, H., Senior, B., Simmons, D., Rusin, J., Picken, G., Thompson, P.M., 2010. Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. *Mar. Poll. Bull.* 60, 888–897. <http://dx.doi.org/10.1016/j.marpolbul.2010.01.003>.
- Barlow, J., Cameron, G.A., 2003. Field experiments show that acoustic pingers reduce marine mammal bycatch in the California drift gill net fishery. *Mar. Mamm. Sci.* 19, 265–283. <http://dx.doi.org/10.1111/j.1748-7692.2003.tb01108.x>.
- Bateson, M., 2007. Environmental noise and decision making possible implications of increases in anthropogenic noise for information processing in marine mammals. *Int. J. Comp. Psychol.* 20, 169–178.

- Beale, C.M., 2007. The behavioral ecology of disturbance responses. *Int. J. Comp. Psychol.* 20, 111–120.
- Beale, C.M., Monaghan, P., 2004. Behavioural responses to human disturbance: a matter of choice? *Anim. Behav.* 68, 1065–1069. <http://dx.doi.org/10.1016/j.anbehav.2004.07.002>.
- Bejder, L., Samuëls, A., Whitehead, H., Gales, N., 2006. Interpreting short-term behavioural responses to disturbance within a longitudinal perspective. *Anim. Behav.* 72, 1149–1158. <http://dx.doi.org/10.1016/j.anbehav.2006.04.003>.
- Berggren, P., Carlström, J., Tregenza, N., 2002. Mitigation of small cetacean bycatch. Evaluation of acoustic alarms (MISNET). Study contract 00/031. Final Report to the European Commission, Univ. of Stockholm.
- Betke, K., Matuschek, R., 2010. Messungen von Unterwasserschall beim Bau der Windenergieanlagen im Offshore-Testfeld "alpha ventus". ITAP, Hamburg.
- Brandt, M.J., Diederichs, A., Betke, K., Nehls, G., 2011. Responses of harbour porpoises to pile driving at the Horns Rev II offshore wind farm in the Danish North Sea. *Mar. Ecol. Prog. Ser.* 421, 205–216. <http://dx.doi.org/10.3354/meps08888>.
- Brandt, M.J., Höschle, C., Diederichs, A., Betke, K., Matuschek, R., Witte, S., Nehls, G., 2012. Far-reaching effects of a seal scarer on harbour porpoises, *Phocoena phocoena*. *Aquat. Cons. Mar. Freshw. Ecosyst.* 23, 222–232. <http://dx.doi.org/10.1002/aqc.2311>.
- Brandt, M.J., Höschle, C., Diederichs, A., Betke, K., Matuschek, R., Nehls, G., 2013. Seal scarers as a tool to deter harbour porpoises from offshore construction sites. *Mar. Ecol. Prog. Ser.* 475, 291–302. <http://dx.doi.org/10.3354/meps10100>.
- Carlström, J., Berggren, P., Tregenza, N.J.C., 2009. Spatial and temporal impact of pingers on porpoises. *Can. J. Fish. Aquat. Sci.* 66, 72–82. <http://dx.doi.org/10.1139/F08-186>.
- Clark, C.W., Ellison, W.T., Southall, B.L., Hatch, L., Van Parijs, S.M., Frankel, A., Ponirakis, D., 2009. Acoustic masking in marine ecosystems: intuitions, analysis, and implication. *Mar. Ecol. Prog. Ser.* 395, 201–222. <http://dx.doi.org/10.3354/meps08402>.
- Culik, B., Koschinski, S., Tregenza, N., Ellis, G.M., 2001. Reactions of harbor porpoises *Phocoena phocoena* and herring *Clupea harengus* to acoustic alarms. *Mar. Ecol. Prog. Ser.* 211, 255–260. <http://dx.doi.org/10.3354/meps211255>.
- Dähne, M., Gilles, A., Lucke, K., Peschko, V., Adler, S., Krügel, K., Sundermeyer, J., Siebert, U., 2013. Effects of pile-driving on harbour porpoises (*Phocoena phocoena*) at the first offshore wind farm in Germany. *Environ. Res. Lett.* 8, 025002. <http://dx.doi.org/10.1088/1748-9326/8/2/025002>.
- DeRuiter, S.L., Hansen, M., Koopman, H.N., Westgate, A.J., Tyack, P., Madsen, P.T., 2010. Propagation of narrow-band-high-frequency clicks: measured and modeled transmission loss of porpoise-like clicks in porpoise habitats. *J. Acoust. Soc. Am.* 127, 560–567. <http://dx.doi.org/10.1121/1.3257203>.
- Ellison, W.T., Southall, B.L., Clark, C.W., Frankel, A.S., 2011. A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Cons. Biol.* 26, 21–28. <http://dx.doi.org/10.1111/j.1523-1739.2011.01803.x>.
- Erbe, C., Farmer, D.M., 1998. Masked hearing thresholds of a beluga whale (*Delphinapterus leucas*) in icebreaker noise. *Deep-Sea Res. Part II* 45, 1373–1388. [http://dx.doi.org/10.1016/S0967-0645\(98\)00027-7](http://dx.doi.org/10.1016/S0967-0645(98)00027-7).
- European Commission, 1992. Council Directive 92/43/EEC of 21 May 1992 on the Conservation of Natural Habitats and of Wild Fauna and Flora.
- European Commission, 2000. Directive 2000/14/EC of the European Parliament and of the Council of 8 May 2000 on the Approximation of the Laws of the Member States Relating to the Noise Emission in the Environment by Equipment for use Outdoors.
- European Commission, 2002. Directive 2002/49/EC of the European Parliament and of the Council of 25 June 2002 Relating to the Assessment and Management of Environmental Noise.
- European Commission, 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 Establishing a Framework for Community Action in the Field of Marine Environmental Policy.
- Fair, P.A., Becker, P.R., 2000. Reviews of stress in marine mammals. *J. Aquat. Ecosyst. Stress Recov.* 7, 335–354.
- Finneran, J.J., 2008. Auditory Weighting Functions and Frequency-dependent Effects of Sound in Bottlenose Dolphins (*Tursiops truncatus*). SPAWAR Systems Center Pacific, San Diego, California.
- Finneran, J.J., Jenkins, A.K., 2012. Criteria and Thresholds for US Navy Acoustic and Explosive Effects Analysis. SPAWAR Systems Center Pacific, San Diego, California.
- Finneran, J.J., Schlundt, C.E., 2010. Frequency-dependent and longitudinal changes in noise-induced hearing loss in a bottlenose dolphin (*Tursiops truncatus*) (L.). *J. Acoust. Soc. Am.* 128, 567–570. <http://dx.doi.org/10.1121/1.3458814>.
- Finneran, J.J., Schlundt, C.E., 2013. Effects of fatiguing tone frequency on temporary threshold shift in bottlenose dolphins (*Tursiops truncatus*). *J. Acoust. Soc. Am.* 133, 1819–1826. <http://dx.doi.org/10.1121/1.4776211>.
- Finneran, J.J., Schlundt, C.E., Dear, R., Carder, D., Ridgway, S.H., 2002. Temporary shift in masked hearing thresholds in odontocetes after exposure to single underwater impulses from a seismic watergun. *J. Acoust. Soc. Am.* 111, 2929–2940. <http://dx.doi.org/10.1121/1.1479150>.
- Götz, T., Janik, V.M., 2011. Repeated elicitation of the acoustic startle reflex leads to sensitisation in subsequent avoidance behaviour and induces fear conditioning. *BMC Neuroscience* 12, 1–12. <http://www.biomedcentral.com/1471-2202/12/30>.
- Gray, H., Waerebeek, K.V., 2011. Postural instability and akinesia in a pantropical spotted dolphin, *Stenella attenuata*, in proximity to operating airguns of a geophysical seismic vessel. *J. Nat. Conserv.* 19, 363–367. <http://dx.doi.org/10.1016/j.jnc.2011.06.005>.
- Green, D.M., Birdsall, T.G., Tanner, W.P., 1957. Signal detection as a function of signal intensity and duration. *J. Acoust. Soc. Am.* 29, 523–531. <http://dx.doi.org/10.1121/1.1908951>.
- Hammond, P.S., Berggren, P., Benke, H., Borchers, D.L., Collet, A., Heide-Jørgensen, M.-P., Heimlich, S., Hiby, A.R., Leopold, M.F., Øien, N., 2002. Abundance of harbour porpoise and other cetaceans in the North Sea and adjacent waters. *J. Appl. Ecol.* 39, 361–376. <http://dx.doi.org/10.1046/j.1365-2664.2002.00713.x>.
- Hammond, P.S., Macleod, K., Berggren, P., Borchers, D.L., Burt, L., Canadas, A., Desportes, G., Donovan, G.P., Gilles, A., Gillespie, D., Gordon, J., Hiby, L., Kuklik, I., Leaper, R., Lehnert, K., Leopold, M., Lovell, P., Øien, N., Paxton, C.G.M., Ridoux, V., Rogan, E., Samarra, F., Scheidat, M., Sequeira, M., Siebert, U., Skov, H., Swift, R., Tasker, M.L., Teilmann, J., Van Canneyt, O., Veen, J.V.D., 2013. Cetacean abundance and distribution in European Atlantic shelf waters to inform conservation and management. *Biol. Conserv.* 164, 107–122. <http://dx.doi.org/10.1016/j.biocon.2013.04.010>.
- Hansen, M., Wahlberg, M., Madsen, P.T., 2008. Low-frequency components in harbor porpoise (*Phocoena phocoena*) clicks: communication signal, by-products, or artifacts? *J. Acoust. Soc. Am.* 124, 4059–4068. <http://dx.doi.org/10.1121/1.2945154>.
- Hermanssen, L., Beedholm, K., Tougaard, J., Madsen, P.T., 2014. High frequency components of ship noise in shallow water with a discussion of implications for harbor porpoises (*Phocoena phocoena*). *J. Acoust. Soc. Am.* 136, 1640–1653. <http://dx.doi.org/10.1121/1.4893908>.
- Jensen, F.H., Bejder, L., Wahlberg, M., Soto, N.A., Johnson, M., Madsen, P.T., 2009. Vessel noise effects on delphinid communication. *Mar. Ecol. Prog. Ser.* 395, 161–175. <http://dx.doi.org/10.3354/meps08204>.
- Johnson, C.S., 1968a. Masked tonal thresholds in the bottlenosed porpoise. *J. Acoust. Soc. Am.* 44, 965–967. <http://dx.doi.org/10.1121/1.1911236>.
- Johnson, C.S., 1968b. Relation between absolute threshold and duration-of-tone pulses in the bottlenosed porpoise. *J. Acoust. Soc. Am.* 43, 757–763. <http://dx.doi.org/10.1121/1.1910893>.
- Johnson, C.S., 1991. Hearing thresholds for periodic 60-kHz tone pulses in the beluga whale. *J. Acoust. Soc. Am.* 89, 2996–3001. <http://dx.doi.org/10.1121/1.400736>.
- Johnson, M., Soto, N.A., Madsen, P.T., 2009. Studying the behaviour and sensory ecology of marine mammals using acoustic recording tags: a review. *Mar. Ecol. Prog. Ser.* 395, 55–73.
- Johnston, D.W., 2002. The effect of acoustic harassment devices in harbour porpoises (*Phocoena phocoena*) in the Bay of Fundy, Canada. *Biol. Conserv.* 108, 113–118. [http://dx.doi.org/10.1016/S0006-3207\(02\)00099-X](http://dx.doi.org/10.1016/S0006-3207(02)00099-X).
- Kastelein, R.A., Bunskoek, P., Hagedoorn, M., Au, W.W.L., Haan, D., 2002. Audiogram of a harbor porpoise (*Phocoena phocoena*) measured with narrow-band frequency modulated signals. *J. Acoust. Soc. Am.* 112, 334–344. <http://dx.doi.org/10.1121/1.1480835>.
- Kastelein, R.A., Wensveen, P.J., Hoek, L., Verboom, W.C., Terhune, J.M., 2009. Underwater detection of tonal signals between 0.125 and 100 kHz by harbor seals (*Phoca vitulina*). *J. Acoust. Soc. Am.* 125, 1222–1229. <http://dx.doi.org/10.1121/1.3132522>.
- Kastelein, R.A., Hoek, L., de Jong, C.A.F., Wensveen, P.J., 2010. The effect of signal duration on the underwater detection thresholds of a harbor porpoise (*Phocoena phocoena*) for single frequency-modulated tonal signals between 0.25 and 160 kHz. *J. Acoust. Soc. Am.* 128, 3211–3222. <http://dx.doi.org/10.1121/1.3493435>.
- Kastelein, R.A., Gransier, R., Hoek, L., Olthuis, J., 2012. Temporary threshold shifts and recovery in a harbor porpoise (*Phocoena phocoena*) after octave-band noise at 4 kHz. *J. Acoust. Soc. Am.* 132, 3525–3537. <http://dx.doi.org/10.1121/1.4757641>.
- Kastelein, R.A., Gransier, R., Hoek, L., Rambags, M., 2013. Hearing frequency thresholds of a harbor porpoise (*Phocoena phocoena*) temporarily affected by a continuous 1.5 kHz tone. *J. Acoust. Soc. Am.* 134, 2286–2292. <http://dx.doi.org/10.1121/1.4816405>.
- Kastelein, R.A., Hoek, L., Gransier, R., Rambags, M., Clayes, N., 2014. Effect of level, duration, and inter-pulse interval of 1–2 kHz sonar signal exposures on harbor porpoise hearing. *J. Acoust. Soc. Am.* 136, 412–422. <http://dx.doi.org/10.1121/1.4883596>.
- Kujawa, S.G., Liberman, M.C., 2009. Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. *J. Neurosci.* 29, 14077–14085. <http://dx.doi.org/10.1523/JNEUROSCI.2845-09.2009>.
- Lucke, K., Siebert, U., Lepper, P.A., Blanchet, M.-A., 2009. Temporary shift in masked hearing thresholds in a harbor porpoise (*Phocoena phocoena*) after exposure to seismic airgun stimuli. *J. Acoust. Soc. Am.* 125, 4060–4070. <http://dx.doi.org/10.1121/1.3117443>.
- Lusseau, D., Bejder, L., 2007. The long-term consequences of short-term responses to disturbance: experiences from whalewatching impact assessment. *Int. J. Comp. Psychol.* 20, 228–236.
- Lusseau, D., Christiansen, F., Harwood, J., Mendes, S., Thompson, P.M., Smith, K., Hastie, G.D., 2012. Assessing the risks to marine mammal populations from renewable energy devices – an interim approach. Final Workshop Report. Joint Nature Conservation Committee, the Countryside Council for Wales and NERC Knowledge Exchange Programme.
- Madsen, P.T., 2005. Marine mammals and noise: problems with root mean square sound pressure levels for transients. *J. Acoust. Soc. Am.* 117, 3952–3957. <http://dx.doi.org/10.1121/1.1921508>.

- Madsen, P.T., Johnson, M., Miller, P.J.O., Aguilar Soto, N., Lynch, J., 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, 2366–2379. <http://dx.doi.org/10.1121/1.2229287>.
- Meloni, T., Rosenheck, A., 1995. Choice of frequency weighting for the evaluation of weapon noise. *J. Acoust. Soc. Am.* 97, 3636–3641. <http://dx.doi.org/10.1121/1.412410>.
- Miller, P.J.O., Johnson, M.P., Madsen, P.T., Biassoni, N., Quero, M., Tyack, P.L., 2009. Using at-sea experiments to study the effects of airguns on the foraging behavior of sperm whales in the Gulf of Mexico. *Deep Sea Res. Part 1* 56, 1168–1181. <http://dx.doi.org/10.1016/j.dsr.2009.02.008>.
- 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: Peterson, N.M., (Ed.), *The Question of Sound from Icebreaker Operations. Proceedings of a Workshop 23 and 24 February 1981. Arctic Pilot Project, Toronto*, pp. 259–268.
- Møhl, B., Andersen, S., 1973. Echolocation: high-frequency component in the click of the harbour porpoise (*Phocoena ph. L.*). *J. Acoust. Soc. Am.* 54, 1368–1372. <http://dx.doi.org/10.1121/1.1914435>.
- Montague, W.E., Strickland, J.F., 1961. Sensitivity of the water-immersed ear to high- and low-level tones. *J. Acoust. Soc. Am.* 33, 1376–1381. <http://dx.doi.org/10.1121/1.1908443>.
- Nabe-Nielsen, J., Sibly, R.M., Tougaard, J., Teilmann, J., Sveegaard, S., 2013. Effects of noise and by-catch on a Danish harbour porpoise population. *Ecol. Modell.* 272, 242–251. <http://dx.doi.org/10.1016/j.ecolmodel.2013.09.025>.
- Nachtigall, P.E., Pawloski, D.A., Au, W.W.L., 2003. Temporary threshold shifts and recovery following noise exposure in the Atlantic bottlenosed dolphin (*Tursiops truncatus*). *J. Acoust. Soc. Am.* 113, 3425–3429. <http://dx.doi.org/10.1121/1.1570438>.
- Nachtigall, P.E., Supin, A.Y., Pawloski, J.L., Au, W.W.L., 2004. Temporary threshold shifts after noise exposure in the bottlenose dolphin (*Tursiops truncatus*) measured using evoked auditory potentials. *Mar. Mammal. Sci.* 20, 673–687. <http://dx.doi.org/10.1111/j.1748-7692.2004.tb01187.x>.
- National Research Council, 2003. *Ocean Noise and Marine Mammals*. The National Academies Press, Washington, DC.
- National Research Council, 2005. *Marine Mammal Populations and Ocean Noise: Determining When Noise Causes Biologically Significant Effects*. National Academic Press, Washington, DC.
- Nielsen, T.P., Wahlberg, M., Heikkilä, S., Jensen, M., Sabinsky, P., Dabelsteen, T., 2012. Swimming patterns of wild harbour porpoises *Phocoena phocoena* show detection and avoidance of gillnets at very long ranges. *Mar. Ecol. Prog. Ser.* 453, 241–248. <http://dx.doi.org/10.3354/meps09630>.
- NMFS, 2003. Taking marine mammals incidental to conducting oil and gas exploration activities in the Gulf of Mexico. *Fed. Reg.* 68, 9991–9996.
- NOAA, 2013. *Draft Guidance for Assessing the Effects of Anthropogenic Sound ON Marine Mammals – Acoustic Threshold Levels for Onset of Permanent and Temporary Threshold Shifts*. US Office of Commerce, Maryland.
- Nowacek, D.P., Thorne, L.H., Johnston, D.W., Tyack, P.L., 2007. Responses of cetaceans to anthropogenic noise. *Mamm. Rev.* 37, 81–115. <http://dx.doi.org/10.1111/j.1365-2907.2007.00104.x>.
- Olesiuk, P.F., Nichol, L.M., Sowden, M.J., Ford, J.K.B., 2002. Effect of the sound generated by an acoustic harassment device on the relative abundance and distribution of harbor porpoises (*Phocoena phocoena*) in Retreat Passage, British Columbia. *Mar. Mamm. Sci.* 18, 843–862. <http://dx.doi.org/10.1111/j.1748-7692.2002.tb01077.x>.
- Parker, D.E., Tubbs, R.L., Littlefield, V.M., 1978. Visual-field displacements in human beings evoked by acoustical transients. *J. Acoust. Soc. Am.* 63, 1912–1918. <http://dx.doi.org/10.1121/1.381894>.
- Plomp, R., Bouman, A., 1959. Relation between hearing threshold and duration for tone pulses. *J. Acoust. Soc. Am.* 31, 749–758. <http://dx.doi.org/10.1121/1.1907781>.
- Popov, V.V., Supin, A.Y., Wang, D., Wang, K., 2006. Nonconstant quality of auditory filters in the porpoises, *Phocoena phocoena* and *Neophocaena phocaenoides* (Cetacea, Phocoenidae). *J. Acoust. Soc. Am.* 119, 3173–3180. <http://dx.doi.org/10.1121/1.2184290>.
- Popov, V.V., Supin, A.Y., Wang, D., Wang, K., Dong, L., Wang, S., 2011. Noise-induced temporary threshold shift and recovery in Yangtze finless porpoises *Neophocaena phocaenoides asiatica*. *J. Acoust. Soc. Am.* 130, 574–584. <http://dx.doi.org/10.1121/1.3596470>.
- Popov, V.V., Supin, A.Y., Rozhnov, V.V., Nechaev, D.I., Sysueva, E.V., 2014. The limits of applicability of the sound exposure level (SEL) metric to temporal threshold shifts (TTS) in beluga whales, *Delphinapterus leucas*. *J. Exp. Biol.* 217, 1804–1810. <http://dx.doi.org/10.1242/jeb.098814>.
- Richardson, W.J., Greene, C.R., Malm, C.I., Thomson, D.H., 1995. *Marine Mammals and Noise*. Academic Press, San Diego.
- Schlundt, C.E., Finneran, J.J., Carder, D., Ridgway, S.H., 2000. Temporary shift in masked hearing thresholds of bottlenose dolphins, *Tursiops truncatus*, and white whales, *Delphinapterus leucas*, after exposure to intense tones. *J. Acoust. Soc. Am.* 107, 3496–3508. <http://dx.doi.org/10.1121/1.429420>.
- Simmonds, M.P., Dolman, S.J., Jasny, M., Parsons, E.C.M., Weigart, L., Wright, A.J., Leaper, R., 2014. Marine noise pollution – increasing recognition but need for more practical action. *J. Ocean Technol.* 9, 70–90.
- Southall, B.L., Bowles, A.E., Ellison, W.T., Finneran, J., Gentry, R., Green, C.R., Kastak, C.R., Ketten, D.R., Miller, J.H., Nachtigall, P.E., Richardson, W.J., Thomas, J.A., Tyack, P.L., 2007. Marine mammal noise exposure criteria. *Aquat. Mamm.* 33, 411–521. <http://dx.doi.org/10.1578/AM.33.4.2007.411>.
- Steevens, C.C., Russell, K.L., Knafelc, M.E., Smith, P.F., Hopkins, E.W., Clark, J.B., 1999. Noise induced neurologic disturbances in divers exposed to intense waterborne sound: two case reports. *Unders. Hyperb. Med.* 26, 261–265. <http://archive.rubicon-foundation.org/2333>.
- Szymanski, M.D., Bain, D.E., Kiehl, K., Pennington, S., Wong, S., Henry, K.R., 1999. Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioral audiograms. *J. Acoust. Soc. Am.* 106, 1134–1141. <http://dx.doi.org/10.1121/1.427121>.
- Terhune, J.M., 2013. A practical weighting function for harbour porpoises underwater sound level measurements (L). *J. Acoust. Soc. Am.* 134, 2405–2408. <http://dx.doi.org/10.1121/1.4816556>.
- Thompson, P.M., Lusseau, D., Barton, T., Simmons, D., Rusin, J., Bailey, H., 2010. Assessing the responses of coastal cetaceans to the construction of offshore wind turbines. *Mar. Poll. Bull.* 60, 1200–1208. <http://dx.doi.org/10.1016/j.marpolbul.2010.03.030>.
- Todd, S., Stevick, P., Lien, J., Marques, F., Ketten, D., 1996. Behavioural effects to underwater explosions in Humpback Whales (*Megaptera novaeangliae*). *Can. J. Zool.* 74, 1661–1672. <http://dx.doi.org/10.1139/z96-184>.
- Tougaard, J., Carstensen, J., Teilmann, J., Skov, H., Rasmussen, P., 2009. Pile driving zone of responsiveness extends beyond 20 km for harbour porpoises (*Phocoena phocoena*, (L.)). *J. Acoust. Soc. Am.* 126, 11–14. <http://dx.doi.org/10.1121/1.3132523>.
- Tougaard, J., Kyhn, L.A., Amundin, M., Wennerberg, D., Bordin, C., 2012. Behavioral reactions of harbor porpoise to pile-driving noise. In: Popper, A.N., Hawkins, A.D. (Eds.), *Effects of Noise on Aquatic Life*. Springer, New York, pp. 277–280. doi:10.1007/978-1-4419-7311-5\_61.
- Vos, J., 2003. A- and C-weighted sound levels as predictors of the annoyance caused by shooting sounds, for various facade attenuation types. *J. Acoust. Soc. Am.* 113, 336–347. <http://dx.doi.org/10.1121/1.1527957>.
- Weigart, L.S., 2007. The need for precaution in the regulation and management of undersea noise. *J. Int. Wildlife Law Pol.* 10, 247–253. <http://dx.doi.org/10.1080/13880290701769312>.
- Wensveen, P.J., Huijser, L.A.E., Hoek, L., Kastelein, R.A., 2014. Equal latency contours and auditory weighting functions for the harbour porpoise (*Phocoena phocoena*). *J. Exp. Biol.* 217, 359–369. <http://dx.doi.org/10.1242/jeb.091983>.
- Williams, R., Lusseau, D., Hammond, P.S., 2006. Estimating relative energetic costs of human disturbance to killer whales (*Orcinus orca*). *Biol. Conserv.* 133, 301–311. <http://dx.doi.org/10.1016/j.biocon.2006.06.010>.
- Wright, A.J., Soto, N.A., Baldwin, A., Bateson, M., Beale, C., Clark, C., Deak, T., Edwards, E., Fernandez, A., Godinho, A., Hatch, L., Kakuschke, A., Lusseau, D., Martineau, D., Romero, L., Weigart, L., Wintle, B., Sciarra, G.N.D., Martin, V., 2007a. Anthropogenic noise as a stressor in animals: a multidisciplinary perspective. *Int. J. Comp. Psychol.* 20, 250–273.
- Wright, A.J., Soto, N.A., Baldwin, A., Bateson, M., Beale, C., Clark, C., Deak, T., Edwards, E., Fernandez, A., Godinho, A., Hatch, L., Kakuschke, A., Lusseau, D., Martineau, D., Romero, L., Weigart, L., Wintle, B., Sciarra, G.N.D., Martin, V., 2007b. Do marine mammals experience stress related to anthropogenic noise? *Int. J. Comp. Psychol.* 20, 274–316.
- Wright, A.J., Rose, N.A., Parsons, E.C.M., Dolman, S.J., 2009. Urging cautious policy applications of captive research data is not the same as rejecting those data. *Mar. Pollut. Bull.* 58, 314–316. <http://dx.doi.org/10.1016/j.marpolbul.2008.12.004>.
- Wright, A.J., Maar, M., Mohn, C., Nabe-Nielsen, J., Siebert, U., Jensen, L.F., Baagøe, H.J., Teilmann, J., 2013. Possible causes of a harbour porpoise mass stranding in Danish waters in 2005. *PLoS ONE* 8, e55553. <http://dx.doi.org/10.1371/journal.pone.0055553>.