Why is auditory frequency weighting so important in regulation of underwater noise?

Jakob Tougaard, and Michael Dähne

Citation: The Journal of the Acoustical Society of America **142**, EL415 (2017);
View online: https://doi.org/10.1121/1.5008901
View Table of Contents: http://asa.scitation.org/toc/jas/142/4
Published by the Acoustical Society of America

Why is auditory frequency weighting so important in regulation of underwater noise?

Jakob Tougaard[a)]
Department of Bioscience, Aarhus University, DK-4000 Roskilde, Denmark
jat@bios.au.dk

Michael Dähne
German Oceanographic Museum, Katharinenberg 14-20, 18439 Stralsund, Germany
michael.daehne@meeresmuseum.de

Abstract: A key question related to regulating noise from pile driving, air guns, and sonars is how to take into account the hearing abilities of different animals by means of auditory frequency weighting. Recordings of pile driving sounds, both in the presence and absence of a bubble curtain, were evaluated against recent thresholds for temporary threshold shift (TTS) for harbor porpoises by means of four different weighting functions. The assessed effectivity, expressed as time until TTS, depended strongly on choice of weighting function: 2 orders of magnitude larger for an audiogram-weighted TTS criterion relative to an unweighted criterion, highlighting the importance of selecting the right frequency weighting.

[WJL]
Date Received: May 10, 2017 **Date Accepted:** October 10, 2017

1. Introduction

Loud impulsive underwater sound is nowadays recognized as a significant source of impact on marine organisms. This recognition has resulted in various forms of regulation to mitigate impact on particularly marine mammals (e.g., German Federal Ministry for the Environment and Nuclear Safety, 2013). At the core of such regulation are exposure limits; thresholds above which undesired negative effects are expected to occur. These thresholds should be based on experimental data, but as the number of possible disturbing sound sources is very large, it is not possible to obtain experimental thresholds for all types. Thresholds must therefore be generalized beyond the particular experimental conditions and sounds used in experiments in order to be applicable to a wider range of sound sources and environmental conditions. The first serious attempt at deriving generalized thresholds for marine mammals was made by Southall *et al.* (2007). One of the central elements of the recommendations of Southall *et al.* (2007) was the concept of auditory frequency weighting functions. Not all frequencies are perceived equally well and therefore weighting takes the frequency specific hearing abilities of the focal species into account. This idea of weighting is far from new, as it is routinely applied to sound measurements as part of human community noise regulation (dB A-weighted levels, see, for example, Houser *et al.*, 2017). Contrary to the situation for human noise regulation, where there is an overwhelmingly large body of empirical data supporting the use of the A-weighted levels, the experimental data available to derive and support specific weighting functions for species of marine mammals is scarce. Several forms of weighting functions have been proposed (see below) and considerable controversy surrounds the question of which function(s) to use [see Wright (2015) for an example]. This controversy is far from a minor technical dispute as the choice of weighting function(s) has a significant impact on regulation of noisy activities, such as offshore pile driving, seismic surveys, and use of navy sonars. If an incorrect weighting is used it may lead to either under-regulation (significant impact on animals missed in assessments, i.e., failure to meet management objectives), or over-regulation (imposing tighter restrictions on activities than required to meet the management objectives). See Tougaard *et al.* (2015) for a further discussion of this issue.

2. An example with pile driving noise

Regulation of noise exposure from pile driving serves as a useful example of the importance of selecting the correct weighting function. This is possible because experimental

[a)]Author to whom correspondence should be addressed.

data are available on a temporary threshold shift (TTS) induced in a harbor porpoise exposed to repeated playbacks of pile driving noise (Kastelein *et al.*, 2015). Strictly speaking, the results from Kastelein *et al.* (2015) are only valid for pile driving pulses identical to the pulses used in the experiment and further assuming that the experimental animal was representative for porpoises in general. Pulse duration, frequency spectrum, repetition rate, and duty cycle are all parameters important in determining whether a given sound can induce TTS or not [see Finneran (2015) for a recent, comprehensive review]. Pulse duration, repetition rate, and duty cycle of the stimulus used by Kastelein *et al.* (2015) were very similar to sounds from a real pile driving. Only the frequency spectrum deviated, as the underwater loudspeaker used by Kastelein *et al.* (2015) was unable to replicate the very low frequencies (below 800 Hz) present in real pile driving noise. The experimental animal was exposed to 2760 identical pulses over a period of 1 hour, each pulse with a sound exposure level (SEL_{SS}, sensu Southall *et al.*, 2007, Appendix A) of 155 dB re. 1 $\mu Pa^2 s$. This amounted to a total cumulated exposure of the animal (SEL_{cum}) of 180 dB re. 1 $\mu Pa^2 s$ and resulted in a small amount of TTS (∼4 dB at 8 kHz). Four different weighting functions, were applied to the frequency spectrum of the pile driving noise used by Kastelein *et al.* (2015) (Fig. 1). The four types were:

(1) Unweighted—all parts of the spectrum weighted equally, as used for example in German legislation (German Federal Ministry for the Environment and Nuclear Safety, 2013).
(2) M-weighting (Southall *et al.*, 2007)—Equal weighting over a broad frequency range, tapering off at low and high frequencies.
(3) Audiogram-weighting, here exemplified by the National Oceans and Atmosphere Administration (NOAA) guidance (National Marine Fisheries Service, 2016)—weighting with a function resembling an inverted audiogram, equivalent to A-weighting in human audiology (see also Nedwell *et al.*, 2007; Terhune, 2013; Tougaard *et al.*, 2015).
(4) Intermediate form (exemplified by the type II weighting of Finneran and Jenkins, 2012). Broad weighting, but with emphasis on the range of best hearing.

The y axis in Fig. 1(B) has been adjusted so the SEL_{cum} in the unweighted condition matches the level used by Kastelein *et al.* (2015), which was sufficient to induce about 4 dB of TTS. Weighting the energy in each frequency band (SEL_i) with a weighting coefficient w_i and summing across all N frequency bands provides the generalized, weighted threshold

$$SEL_w = 10 \log_{10} \sum_{N}^{1} w_i 10^{(SEL_i/10)}. \quad (1)$$

From this four different weighted thresholds (Table 1) are derived, all based on the playback signal of Kastelein *et al.* (2015). The unweighted threshold is equal to an SEL_{cum} of 180 dB re. 1 $\mu Pa^2 s$, the exposure from Kastelein *et al.* (2015). The three other thresholds should be understood as equivalent SEL_{cum} of a signal with frequency

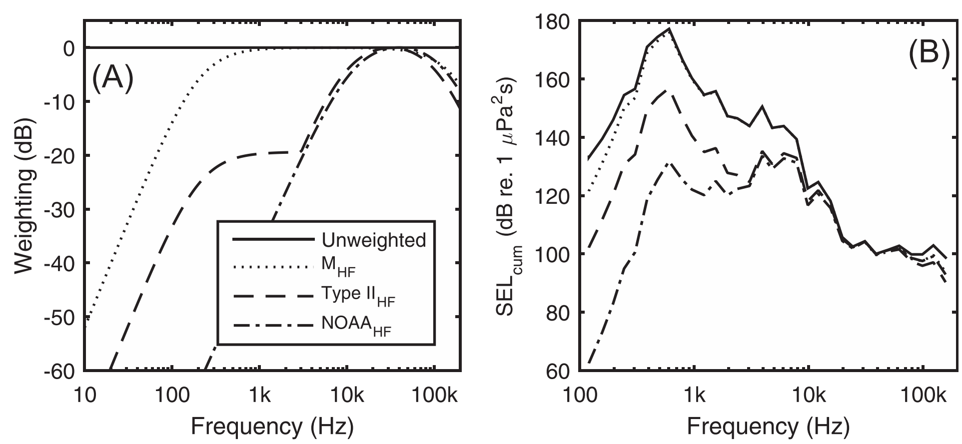

Fig. 1. (A) Four different weighting functions (filters) proposed for high frequency cetaceans: unweighted, M-weighting (Southall *et al.*, 2007), Type-II weighting (Finneran and Jenkins, 2012), and audiogram-weighting (National Marine Fisheries Service, 2016). (B) One-third octave level spectrum of the stimulus used by Kastelein *et al.* (2015), weighted according to the weighting functions in (A).

Table 1. Weighted thresholds for inducing TTS with repeated pile driving noise, derived from Kastelein *et al.* (2015), by means of four different weighting functions. Right column shows the amount subtracted by the weighting, relative to the unweighted signal (weighting factor).

Weighting function	Weighted SEL_{cum} at TTS threshold	Difference re. unweighted
Unweighted	180 dB re. 1 $\mu Pa^2 s$	
M_{HF}	179 dB re. 1 $\mu Pa^2 s$ in flat part of the weighting function	−1 dB
$TypeII_{HF}$	159 dB re. 1 $\mu Pa^2 s$ in flat part of the weighting function	−21 dB
$NOAA_{HF}$	140 dB re. 1 $\mu Pa^2 s$ in flat part of the weighting function	−40 dB

in the flat part of the weighting function. Thus, for example, the $NOAA_{HF}$ threshold of 140 dB re. 1 $\mu Pa^2 s$ should be understood such that the pile driving noise of the experiment is predicted to have the same ability to induce TTS in a porpoise as a signal of similar duration, but in the frequency range where the weighting curve is flat (above 30 kHz) and with an SEL_{cum} of 140 dB re. 1 $\mu Pa^2 s$.

Which of the four weighted levels is the best candidate for a generalized TTS-threshold for pile driving noise? Although there is some evidence that weighting with an inverse audiogram may be the appropriate choice, at least for porpoises (Tougaard *et al.*, 2015), this question is not the focus of the following. The aim is instead to illustrate, by means of application of the four weighting functions to pile driving noise, how the particular choice of weighting function has severe implications for evaluation of efficacy of mitigation measures and hence regulation of pile driving activities, and thus highlight the importance of identifying the most appropriate type of weighting.

3. Application to recordings from a real pile driving

The four different thresholds were applied to recordings of pile driving noise obtained during construction of the DanTysk offshore wind farm in the German Bight. Recordings were made with autonomous data loggers (DSG-Ocean, Loggerhead Instruments, Sarasota, FL) moored about 3 m above the sea bed, at distances of either 8.1 or 8.8 km from the piling sites and under generally similar environmental conditions. The data loggers were equipped with HTI96min hydrophones (HighTech Inc., Long Beach, MA; nominal sensitivity of −186 dB re. 1 V/μPa). Signals were sampled at 80 ksamples/s, digitally low-pass filtered, decimated, and stored as 40 ksamples/s, 16 bit uncompressed files onto SD memory cards (128 Gbyte, Sandisk, Milpitas, CA), providing 20 kHz of bandwidth of the final recordings. Additional details about the recordings and pile driving operation can be found in Dähne *et al.* (2017).

Examples are shown in Fig. 2, recorded from two different foundations; one piled without bubble curtains and one piled while a circular bubble curtain (referred to as Big Bubble Curtain; see, for example, Diederichs *et al.*, 2014) was in operation. The maximum piling energy was 1.0 and 1.4 MJ, respectively, in the two cases.

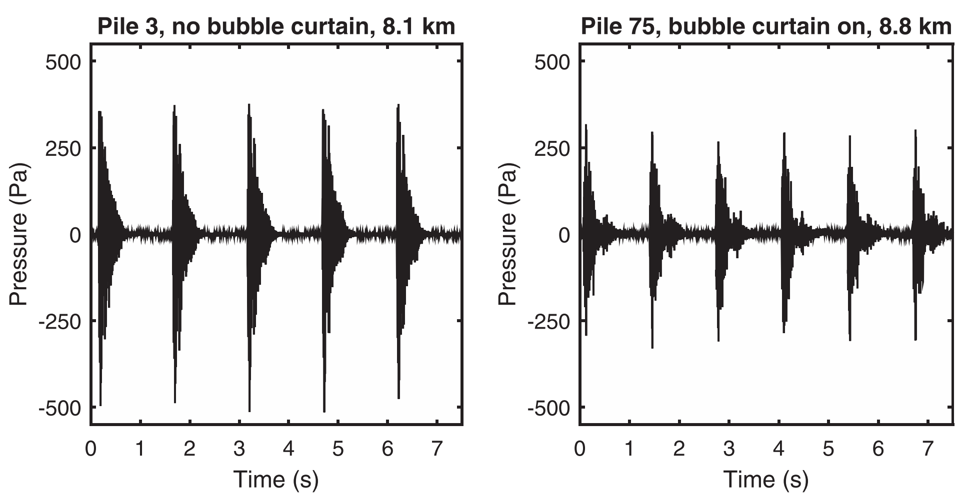

Fig. 2. Examples of acoustic pulses recorded during pile driving at the DanTysk offshore wind farm, with and without bubble curtain. Note that the stroke rate for pile no. 75 (right) was higher than for pile no. 3 (left).

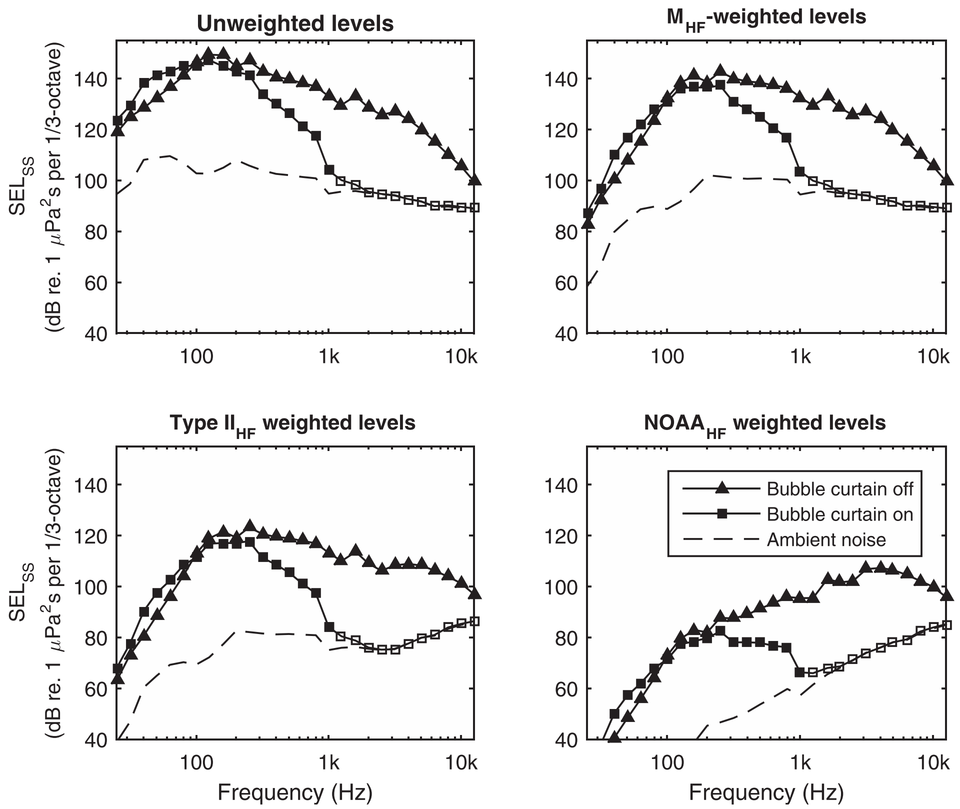

Fig. 3. Weighted spectra of the pile driving sounds from Fig. 2, including ambient noise measured immediately before start of pile driving. Ambient noises were computed with the same time window as single pulses, making the measurements directly comparable. Note that recordings with bubble curtain on were completely dominated by ambient noise above 2 kHz and thus only frequency bands up to and including 1 kHz were used in the analysis, indicated by the filled squares.

The presence of the bubble curtain affected the peak-to-peak sound pressure of the pulses very little (Fig. 2), whereas there was a major effect on the frequency spectrum (Fig. 3, upper left). The dominant part of the unweighted spectrum both with and without bubble curtains was below a few hundred Hz, but the signals recorded without bubble curtains contained substantial energy also at higher frequencies. This was in contrast to the signals recorded with the bubble curtain present, where the noise level above 2 kHz was totally dominated by the ambient noise.

In order to assess how effective the bubble curtain was in preventing TTS in porpoises the number of pulses required to induce TTS (N_{TTS}) was computed for the two situations, with and without bubble curtain. The calculation is under the simplifying assumptions that the porpoise remains stationary and that TTS is determined by the total cumulated energy of all received pulses. N_{TTS} is estimated from the recorded single pulse SEL (weighted SEL_{SS}, from Table 2) and the SEL just sufficient to induce TTS (weighted SEL_{cum}, from Table 1),

$$SEL_{cum} = SEL_{SS} + 10\log_{10} N_{TTS} \Rightarrow N_{TTS} = 10^{[(SEL_{cum}-SEL_{SS})/10]}. \quad (2)$$

Both SEL_{cum} and SEL_{SS} must be weighted with the same weighting function in order to be comparable. SEL_{SS} is weighted in the same manner as SEL_{cum} [Eq. (1)],

Table 2. Weighted single pulse SEL_{SS} with and without bubble curtain. Unit is equivalent dB re. 1 $\mu Pa^2 s$ in the flat part of the weighting function. In parentheses is indicated the difference between the unweighted signal and the weighted signal.

	Unweighted	M	Type II	NOAA
No bubble curtain	156.7 dB	150.4 dB (−6.3 dB)	131.1 dB (−25.6 dB)	115.4 dB (−41.3 dB)
Bubble curtain on[a]	154.4 dB	144.5 dB (−9.9 dB)	125.2 dB (−29.2 dB)	89.4 dB (−65 dB)
Attenuation	2.3 dB	5.9 dB	5.9 dB	25.9 dB

[a]Only energy below 2 kHz has been included, as levels above 2 kHz were completely dominated by ambient noise.

i.e., by summing the weighted third-octave band levels (excluding the bands above 2 kHz for recordings with bubble curtain on, as these were dominated by ambient noise).

The efficacy of the bubble curtain can be quantified as ME, the ratio between the mitigated (M) and unmitigated (U) N_{TTS}. ME expresses how much longer a porpoise can remain exposed to pile driving pulses in the mitigated situation before TTS is induced relative to the unmitigated situation,

$$\mathrm{ME} = \frac{N_{\mathrm{TTS}}(M)}{N_{\mathrm{TTS}}(U)} = 10^{\mathrm{SEL_{SS}}(U) - \mathrm{SEL_{SS}}(M)/10}. \quad (3)$$

ME computed for the four different weighting functions are shown in Table 3.

4. Discussion

When assessing the risk of TTS in porpoises by exposure to pile driving noise we are fortunate to have a TTS threshold for a signal resembling real pile driving noise (Kastelein *et al.*, 2015). This is optimal, because no matter which weighting function we select, the amount subtracted from the threshold by the weighting (Table 1) is matched by roughly the similar amount subtracted on the pile driving signal (Table 2, no bubble curtain). Because of this similarity, the estimated number of pulses required to reach TTS, as estimated in Table 3, is remarkably similar (between 215 and 605 pulses). Most striking is the similarity between unweighted and NOAA-weighted values (215 vs 290), despite the weighted threshold being 40 dB lower than the unweighted threshold (Table 1).

In more direct terms: In the assessment of impact of pile driving signals without bubble curtains it does not matter much which weighting function is used. The reason is that the spectrum of the experimental signal from Kastelein *et al.* (2015), for all practical purposes, is similar to the spectrum of a real pile driving pulse, meaning that any weighting performed on the experimental signal is counteracted by the same weighting on the real signal.

The situation is dramatically different, however, when a bubble curtain is used. If assessed by means of unweighted levels, the bubble curtain appears to have had very little affect (2.3 dB attenuation, corresponding to an ME ratio of 1.7). This is in stark contrast to the assessment on basis of NOAA-weighted levels (25.9 dB attenuation; ME ratio 393). Assessments based on M-weighting and type II weighting are almost identical (ME ratio 3.9 and 4.0, respectively) and fall between the values for unweighted and NOAA-weighted although much closer to the unweighted condition. The explanation behind these very large differences lies in the uneven effect of the bubble curtain across the frequency spectrum. The bubble curtain almost completely removed energy above 1 kHz, whereas the effect on lower frequencies was more moderate. As especially the NOAA-weighting puts more weight to the higher frequencies, where porpoises have better hearing, this means that the bubble curtain had a much more dramatic effect on the NOAA-weighted levels than on the unweighted, M- and type II-weighted levels. Our conclusions regarding the effectivity of the bubble curtain thus strongly depends on the choice of weighting function. If assessment is based on unweighted levels, as in the German legislation regulating pile driving noise, but a NOAA-weighting is really more appropriate (as indicated by Tougaard *et al.*, 2015), then we would erroneously conclude that the use of bubble curtains is not very effective. The consequences of this could easily be an over-regulation, if legislation requires broadband levels to be further attenuated, despite the fact that hardly any energy is present above 1 kHz (at 8 km range, as in the current example).

A complicating factor, especially for the NOAA-weighted levels, is the ambient noise. In the example above, it was assumed that ambient noise does not contribute to TTS and thus the part of the noise spectrum above 1 kHz, dominated by ambient noise, was excluded. If instead, ambient noise is assumed to contribute to TTS and the

Table 3. Number of pulses required to reach the exposure from Kastelein *et al.* (2015) that was just sufficient to induce TTS, estimated for a stationary porpoise at the point where measurements were taken (8.1 and 8.8 km from the piling site, without and with bubble curtain, respectively), and assessed with four different weighting functions. The ratio (ME) between the two is the mitigation efficiency.

	Unweighted	$\mathrm{M_{HF}}$	Type $\mathrm{II_{HF}}$	$\mathrm{NOAA_{HF}}$
No bubble curtain	215	718	613	290
Bubble curtain on	361	2787	2424	113 994
Ratio (ME)	1.7	3.9	4.0	393

full spectrum up to 12.5 kHz is included, then the NOAA-weighted SEL_{cum} with bubble curtain increases from 89.4 dB re. 1 $\mu Pa^2 s$ in the flat part of the weighting function to 93.3 dB re. 1 $\mu Pa^2 s$ in the flat part of the weighting function, i.e., about 3 dB less attenuation of the bubble curtain, whereas all other weighted levels remain unaffected. The higher SEL_{cum} for the NOAA-weighted level with the bubble curtain present is reflected in a similar reduction in ME, which degreases from 393 to 162. In the end, this means that in the current example it matters for the numbers whether ambient noise is included or not, but the overall conclusion remains the same: the effectivity of the bubble curtain, as assessed by NOAA-weighted sound levels, is 2 orders of magnitude higher than if assessed on broadband (unweighted) sound levels.

The bottom line is that the weighting functions are central in the regulation of underwater noise and continued effort should go into designing, conducting, and evaluating experiments targeting this weighting. As the weighting is likely species, or species-group specific, and possibly also depends on noise sources and type of impact (injury or behavioral disturbance), this is not a quick process and there are many unresolved issues. This must be recognized in particular by regulators. Even if difficult and inconvenient, it is important to keep legislation and regulation sufficiently open and flexible to allow incorporation of new results as soon as a sufficient consensus is reached about their validity.

Acknowledgments

Thanks to the Vattenfall DanTysk offshore wind farm for cooperation and providing access to the wind farm. Recordings were collected in cooperation with BioconsultSH, Husum as part of the DEPONS project funded by Vattenfall, Forewind, ENECO Luchterduinen, DONG Energy, and ScottishPower Renewables. Thanks also to the constructive comments from two anonymous reviewers.

References and links

Dähne, M., Tougaard, J., Carstensen, J., Rose, A., and Nabe-Nielsen, J. (**2017**). "Bubble curtains attenuate noise from offshore wind farm construction and reduce temporary habitat loss for harbour porpoises," Mar. Ecol. Prog. Ser. **580**, 221–237.

Diederichs, A., Pehlke, H., Nehls, G., Bellmann, M., Gerke, P., Oldeland, J., Grunau, C., Witte, S., and Rose, A. (**2014**). *Entwicklung und Erprobung des Großen Blasenschleiers zur Minderung der Hydroschallemissionen bei Offshore-Rammarbeiten* (*Development and Test of Large Bubble Curtains to Reduce Underwater Noise Emissions from Offshore Pile Driving*) (Bioconsult-SH, Husum, Germany), p. 247.

Finneran, J. J. (**2015**). "Noise-induced hearing loss in marine mammals: A review of temporary threshold shift studies from 1996 to 2015," J. Acoust. Soc. Am. **138**, 1702–1726.

Finneran, J. J., and Jenkins, A. K. (**2012**). "Criteria and thresholds for U.S. Navy acoustic and explosive effects analysis" (SPAWAR Systems Center Pacific, San Diego, CA), p. 65.

German Federal Ministry for the Environment and Nuclear Safety (**2013**). "Konzept für den Schutz der Schweinswale vor Schallbelastungen bei der Errichtung von Offshore-Windparks in der deutschen Nordsee (Schallschutzkonzept)" ["Concept for protection of porpoises against noise impact from construction of offshore wind farms in the German North Sea (Noise protection concept)"], available at https://www.bfn.de/fileadmin/BfN/awz/Dokumente/schallschutzkonzept_BMU.pdf (Last viewed 10/20/2017).

Houser, D. S., Yost, W., Burkard, R., Finneran, J. J., Reichmuth, C., and Mulsow, J. (**2017**). "A review of the history, development and application of auditory weighting functions in humans and marine mammals," J. Acoust. Soc. Am. **141**, 1371–1413.

Kastelein, R. A., Gransier, R., Marijt, M. A. T., and Hoek, L. (**2015**). "Hearing frequency thresholds of harbor porpoises (*Phocoena phocoena*) temporarily affected by played back offshore pile driving sounds," J. Acoust. Soc. Am. **137**, 556–564.

National Marine Fisheries Service (**2016**). "Technical guidance for assessing the effects of anthropogenic sound on marine mammal hearing underwater acoustic thresholds for onset of permanent and temporary threshold shifts. NOAA Technical Memorandum NMFS-OPR-55," Silver Spring, MD, p. 178.

Nedwell, J. R., Turnpenny, A. W. H., Lovell, J., Parvin, S. J., Workman, R., Spinks, J. A. L., and Howell, D. (**2007**). "A validation of the dBht as a measure of the behavioural and auditory effects of underwater noise. Subacoustech Report No. 534R1231," Subacoustech, Southampton, United Kingdom.

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., and Tyack, P. L. (**2007**). "Marine mammal noise exposure criteria," Aquat. Mamm. **33**, 411–521.

Terhune, J. M. (**2013**). "A practical weighting function for harbour porpoises underwater sound level measurements (L)," J. Acoust. Soc. Am. **134**, 2405–2408.

Tougaard, J., Wright, A. J., and Madsen, P. T. (**2015**). "Cetacean noise criteria revisited in the light of proposed exposure limits for harbour porpoises," Mar. Poll. Bull. **90**, 196–208.

Wright, A. J. (**2015**). "Sound science: Maintaining numerical and statistical standards in the pursuit of noise exposure criteria for marine mammals," Front. Mar. Sci. **2**, 99.