Peer Review Report (Completed at Request of NOAA):

U.S. Navy Technical Report: Auditory Weighting Functions and TTS/PTS Exposure Functions for TAP Phase 3 Acoustic Effects Analyses (February 2015)

By
J.J. Finneran, SSC Pacific

Peer Reviewers[1]

Whitlow Au, Ph.D., University of Hawaii
Colleen Le Prell, Ph.D., University of Florida
Klaus Lucke, Ph.D., Curtin University (Australia)
John (Jack) Terhune, Ph.D., University of New Brunswick (Canada)

The intent of this NOAA-initiated independent peer review was to evaluate a new methodology proposed by the U.S. Navy (Navy) in a Navy Technical Report for consideration and incorporation into NOAA's Draft Guidance for Assessing the Effects of Anthropogenic Sound on Marine Mammal Hearing: Underwater Acoustic Threshold Levels for Onset of Permanent and Temporary Threshold Shifts (Acoustic Guidance)[2].

Thus, NOAA was responsible for conducting this independent peer review. The intent of the peer review report is to address only how NOAA plans to consider the peer reviewers' comments in the Acoustic Guidance. Peer reviewers' comments were also provided to the Navy to consider in their "TAP Phase 3" analyses. This report does not to address how the Navy will incorporate or consider this review. NOAA requested the Navy's assistance in addressing certain comments, which are specifically designated as including input from both NOAA and the Navy.

The Acoustic Guidance pertains to marine mammals, under NOAA's jurisdiction, exposed to underwater sound sources. However, the version of the Navy Technical Report that underwent independent peer review covered a broader range of marine protected species. Thus, while Reviewer's comments relating to sirenians, sea turtles, phocids (air), and otariids (air) and other non-phocid marine carnivores (air) are valuable, and included in this broader Navy Technical Report, they are not directly addressed in this Peer Review Report or in the Acoustic Guidance. As a result, the Navy provided NOAA an updated version of the Technical Report to include in the July 2015 Draft

---

[1] Note: Peer Reviewers' comments are presented as provided to NOAA. Generally, NOAA did not make any alterations (i.e., there may be spelling, grammatical, or other minor errors). If alterations were made, they were done for clarity and are indicated by brackets.

[2] Note: The peer reviewers were only asked to review the Navy's Technical Report. They were not asked to review any version of NOAA's Acoustic Guidance.

Acoustic Guidance (i.e., version made available for second public comment period), which only contained information on species under NOAA's jurisdiction.

General Comments[3]

REVIEWER 1

Comment 1: Technical and Scientific Justification**.** The US Navy's technical report entitled "Auditory Weighting Functions and TTS/PTS Exposure Functions for Navy Phase 3 Acoustic Effects" by J.J. Finneran is based on the most recent scientific data currently available. The literature is carefully reviewed and a variety of data are considered. Moreover, the authors made careful effort to assure that the final values were not biased by any single animal or database. This could have been a significant issue, as some animal subjects have served in multiple studies given the costs of obtaining, training, and maintaining marine mammal subjects. One shortcoming is the lack of data available for the low frequency cetaceans, for which there have been no direct measurements of hearing sensitivity in any mysticete. The authors relied on mathematical models for this group. There has been reasonable predictive value from the models based on basilar membrane mechanics and dimensions used by Ketten and Mountain as well as the finite element models developed by Cranford and Krysl. Because the models have been reasonably useful, this appears to be a reasonable approach in the absence of empirical data.

Response: Thank you for your comment. NOAA agrees with your assessment.

Comment 2: Summary**.** The methodology is based on data available at this time. Data are limited. TTS growth has been more carefully studied in some species than others, using a mix of [auditory evoked potentials] AEP and behavioral techniques, with the specific threshold testing procedure having the potential to impact conclusions. It is not entirely clear how AEP data are integrated, given a statement in the text that only behavioral data were used, but [there was] widespread inclusion of AEP-based data in tables and figures. For a subset of species, thresholds are unknown, and TTS growth has not been measured. The author has made a variety of assumptions, which are clearly stated throughout the document. Where speculation is required, the author has tended to err on the conservative side. This is reasonable and appropriate. Application of weighting functions to noise monitoring protocols and stated noise limits ultimately allows more noise in the environment; it could be harmful to some species if discounted noise is in fact hazardous to members of a given species. The point at which hazard begins remains a major question. The extent to which a 6-dB TTS results in biologically relevant deficits is not known and it is not clear how a regulatory agency would require noise limits be applied – i.e., different limits in different parts of the ocean based on assumed species distribution? a single limit based on an average weighting function? A single limit based on the most conservative weighting for any species such that a single broader weighting filter is applied? This document does not address those questions; it provides the data that regulatory bodies require to make such decisions. It is reasonable and appropriate to use audiometric data and TTS onset to define the TTS and PTS exposure functions as the weighting functions are essentially the inverse of the exposure functions (per equations 1 and 2). The equations, calculations, and

---

[3] Reviewer identification numbers do not correspond to the order of reviewers above.

assumptions are transparent and reproducible. As new datasets become available, it will be relatively seamless to assess whether new data support the conclusions in this report.

> ***Response***: NOAA thanks the Reviewer for the comment. Several of the points raised are addressed in subsequent responses. Specifically, your comment on clarification associated with the use of AEP data is addressed in the response to Comment 36 and your comment about weighting functions allowing more noise in the environment is addressed in response to Comment 43.
>
> NOAA agrees that the biological relevance of a 6-dB shift has yet to be determined. Nevertheless, based on limited data and uncertainty, NOAA has decided to conservatively use a 6-dB threshold shift to represent TTS onset.
>
> Regarding implementing thresholds in a regulatory context, the Reviewer's assessment is correct that this is beyond the scope of the Navy's Technical Report (i.e., intent of report is only to summarize current state of science) and that these determinations will need to be addressed by NOAA in our Acoustic Guidance. Simply, NOAA is planning on applying the thresholds and weighting functions on the basis of a functional hearing group, as depicted in Table 10 of the Navy's Technical Report (i.e., all members within a particular functional hearing group would have the same auditory weighting function and PTS/TTS threshold).

REVIEWER 2

Reviewer provided no general comments.

REVIEWER 3

Comment 3: The aim of this review is to provide comments on the scientific information and data contained within the Navy's technical report "Auditory weighting functions and TTS/PTS exposure functions for TAP Phase 3 acoustic effects analyses", J.J. Finneran (SSC Pacific, 2015). The reviewed document "describes the rationale and steps used to defined proposed numeric thresholds for predicting auditory effects on marine animals exposed to active sonars, other (non-impulsive) active acoustic sources, explosives, pile driving, and air guns for Phase 3" of the US Navy's Tactical Training Theater Assessment and Planning (TAP) Program.

In this technical report, Finneran takes the rationale described by Southall et al. (2007) as well as Finneran & Jenkins (2012) several steps further as new data allow the refining of the auditory weighting function approach. In this respect, Finneran's report clearly represents the "best available" science. The new weighting functions developed by Finneran seem to be more reliable than any of the alternative approaches (e.g. inverted audiogram, M-weighting and type I and II weighting functions). His approach is stringent, following the rationale of weighting curves and equal loudness contours determined for humans: "*Just as human damage risk criteria use auditory weighting functions to capture frequency-dependent aspects of noise, US Navy acoustic impact analyses use weighting functions to capture the frequency-dependency of TTS and PTS in marine mammals and sea turtles*" (Finneran, 2015).

Response: NOAA agrees. Thank you for your comment.

Comment 4: From a scientific and technical point of view there is no criticism to Finneran report.

Response: NOAA thanks the Reviewer for their comment.

Comment 5: There are two aspects, however, which are missing or being not fully elaborated:

One is the scientific uncertainty, the other one the definition of injury. It is debatable whether or not such aspects should be considered by the Navy in a technical report (and hence should be covered within the scope of Finneran's report).

Response: NOAA will respond to the more detailed aspects of this comment presented in further comments from this Reviewer.

REVIEWER 4

Comment 6: The task that Jim Finneran has taken on seems monumental to me since we know so little about hearing in marine mammals. There are approximately 83 species of cetaceans divided up into mysticetes and odontocetes. Of these approximately 85% of cetaceans are odontocetes. There are no audiograms for any mysticete specie and audiograms have been obtained for less than 15% of the odontocete species. Equal loudness data exist for a single *Tursiops truncatus* and a single *Phocoena phocoena* and for no other odontocetes. Temporary threshold shift was been measured for only bottlenose dolphins, beluga whales, harbor porpoise and Yangtze finless porpoise. Audiograms and equal loudness curves obtained via response latency have been measured for a small number of pinniped species, California sea lions, harbor seals, northern elephant seals and walruses. Temporary threshold shift has been measure for only three species of pinniped, harbor seals, California sea lions and northern elephant seas. Knowledge of sea turtle hearing is even more sparse.

Response: Thank you for your comment. NOAA agrees that there are data limitations. Nevertheless, we believe the available data do provide enough information to establish and support threshold levels and auditory weighting functions. Within the Acoustic Guidance, we have added an Appendix highlighting data gaps and providing areas where more research is needed. Additionally, the Acoustic Guidance is able to be updated as more data become available.

Comment 7: As a summary of my review, I would reiterate that given the tremendous lack of data of hearing in marine mammals and given the lack of some fundamental understanding of auditory principles in marine mammals, Finneran has done a extremely well in tackling this topic. I doubt if a better job can be done.

Response: Thank you for your comment.

Specific Comments (by Section)

1. INTRODUCTION

1.3. Noise-induced thresholds

REVIEWER 2

Comment 8: "TTS onset" as being the level of the sound exposure that will induce 6 dB of TTS should be defined here (and in the Executive Summary) rather than at line 612. Other authors have defined the onset of TTS as being the level that is measurably above the normal threshold and the term "onset" relates to the beginning of a phenomenon. Perhaps "TTS 6 dB" would be a more intuitive term.

> ***Response (NOAA & Navy)***: To aid the reader, the Acoustic Guidance provides a glossary defining various terms, including TTS. NOAA disagrees with the Reviewer and considers the terminology of "TTS onset" preferable and more intuitive to readers of the Acoustic Guidance compared to using "TTS 6 dB." Within the Navy Technical Report, Section 1.3 is focused on scientific definitions of NITS, TTS, and PTS and not on regulatory definitions for the onset of TTS or PTS, therefore section 1.3 is unchanged. However, the definitions for TTS onset and PTS onset have been added to the Executive Summary.

1.4. Auditory weighting functions

REVIEWER 2

Comment 9: Human dB(A) and dB(C) weighting functions reflect equal loudness curves at 40 and 100 dB above threshold at 1 kHz. The dB(C) weighting function is much flatter than the dB(A) weighting function and thus takes into account more sound energy at low and high frequencies. Thus, it would be best to use dB(A) for low and moderate sound levels and dB(C) for high amplitude levels. The "M weighting" curves defined by Southall et al. (2007) were patterned after the dB(C) weighting concept while the current proposal is adopting a dB(A) approach. That is, the low frequency rolloffs are patterned after the slopes of the composite audiograms or the equal latency curves at or below 40 dB above threshold.

The switch from a dB(C) equivalent to a dB(A) equivalent exposure function will result in relatively higher sound levels at all but the most sensitive frequency being required to induce specific levels of TTS and PTS, especially whenever lower frequencies are involved. This is undoubtedly a complex and difficult issue to deal with, especially considering the dearth of marine mammal data, but I think it should be briefly discussed or acknowledged.

> ***Response (Navy & NOAA)***: NOAA agrees that establishing marine mammal auditory weighting functions is a complex issue. Nevertheless, NOAA believes the methodology for updating marine mammal auditory weighting functions in the Navy's Technical Report (i.e., transitioning from marine mammal auditory weighting functions that have previously been more similar to human dB(C) functions to those more similar to human dB (A) functions) is

supported by the best available science presented in the Navy Technical Report. Since the development of the Navy's Phase 2 Analyses (Finneran and Jenkins 2012), new data have been obtained regarding marine mammal hearing[4] (e.g., Ghoul and Reichmuth, 2014; Ketten, 2014; Ketten and Mountain, 2014; Sills et al., 2014; Cranford and Krysl, 2015), marine mammal equal latency contours (e.g., Finneran et al., 2013; Reichmuth, 2013; Wensveen et al., 2014), and the potential effects of noise on marine mammal hearing (e.g., Kastelein et al., 2012b; Kastelein et al., 2012a; Finneran and Schlundt, 2013; Kastelein et al., 2013a; Kastelein et al., 2013b; Popov et al., 2013; Kastelein et al., 2014b; Kastelein et al., 2014a; Popov et al., 2014; Finneran et al., 2015; Kastelein et al., 2015). In particular, the recent data from Kastelein (2012a, 2014a, 2014b) for TTS in the same harbor porpoise exposed to noise at frequencies of 1–2, 4, and 6–7 kHz show large differences in TTS onset that are better fit using a "dB(A) approach" (i.e., larger TTS onset change with frequency) compared to a "dB(C) approach". This is illustrated in Figure 19 of the Navy Technical Report, which shows that the Phase 3 TTS exposure function and the composite audiogram fit the data better than the Phase 2 exposure function.

Comment 10: It might be useful to include an overview of the sequence of steps to be taken in the determination of the values for the weighting function and exposure function calculations (Table 8 and equations 1 and 2) in the Introduction. This would provide a road map for the reader and establish the rationale for the overall approach.

Response (NOAA & Navy): The text in the Navy Technical Report has been revised to help clarify the sequence in Section 3. Adding explanatory descriptions to the Introduction was not practical, since many of the required terms have not been defined.

Comment 11: My interpretation of the sequence of analyzes is:

* For each species group, the median values of the original threshold data were obtained (where available) to produce composite audiograms. These data were used to determine the detection threshold at the most sensitive frequency.

* For each individual subject, the audiogram values were normalized such that the lowest threshold was set at 0 dB. These data were used to determine the shapes of the audiograms, particularly the slopes of the lower and upper frequency rolloffs.

* equal loudness data or equal latency data were used to adjust the slopes of the lower frequency auditory weighting functions ("*a*" values of Table 8) where these slopes were lower than the slopes of the normalized audiograms.

* the "*b*" values of Table 8 were set at 2 to reflect the audiogram upper frequency slopes.

*TTS data, expressed as SEL, were obtained at different frequencies for the HF and MF groups.

---

[4] The Navy's Technical Document also describes new data available on sea turtle hearing, which will not be addressed in this Peer Review Report.

*the width of the frequency range of highest sensitivity (values *f1* and *f2* in Table 8) were estimated by varying the low and high frequency cutoff points of the normalized MF audiograms and determining the best fit to the TTS data. The delta T values determined here were used in the other species group calculations.

* The *a, b, f1* and *f2* values were applied to the generalized weighting function equation 1. The "*C*" values in the equation were then adjusted such that the weighting would be 0 dB at the most sensitive frequency.

*To determine the level of the exposure function, (equation 2, the inverse of equation 1), the "*K*" value was adjusted until the available TTS 6 dB SEL level data fell on the line. Thus, the resulting exposure function equation could be used to determine when the amplitudes and durations of sounds over the frequency range would reach the SEL level that would result in 6 dB of TTS.

* The levels associated with the onset of PTS were then set to be *x* dB above the TTS (6 dB) onset values determined in the exposure function for each group.

Alternatively, the above could be included in section 3 (Methodology to derive function parameters).

> ***Response (NOAA & Navy):*** The Reviewer's interpretation of the sequence of analysis is mostly correct. However, for the fifth bullet, TTS onset was determined for all species groups for which data exist. Although only the MF and HF data were used to determine delta T, all data were used to find the values of K that best-fit the TTS data and for Bullet 6, both the MF and HF groups were used.

2. WEIGHTING FUNCTIONS AND EXPOSURE FUNCTIONS

REVIEWER 2

Comment 12: The human dB(A) curves smooth the actual 40 phon equal loudness curves and provide a practical means (mathematical and electronic) by which dB(A) can be calculated or measured using a sound meter. Similarly, equations 1 and 2 provide a practical means to depict/calculate TTS 6 dB levels across the hearing frequency ranges of the species groups.

> ***Response:*** NOAA agrees. Thank you for your assessment.

REVIEWER 3

Comment 13: Finneran refers to "*limited nature of the underlying data*" and employs some conservative approximations in his approach, but nevertheless this report lacks a quantification and thorough discussion of the uncertainties included in these calculations. Again, this is no negative criticism to the scientific approach employed in the technical report as such, but if Finneran's report will be considered for regulatory purposes by any regulatory body, such considerations are necessary.

> ***Response***: See response to next comment regarding quantification of uncertainty.

Comment 14: There is still a significant lack of experimental data in the context of auditory effects of sound in marine species, a fact also acknowledged by Finneran: "*Ideally, these parameters would be based on experimental data describing the manner in which the onset of TTS or PTS varied as a function of exposure frequency. In other words, a weighting function for TTS should ideally be based on TTS data obtained using a range of exposure frequencies, species, and individual subjects within each species group. However, at present, there are only limited data for the frequency-dependency of TTS in marine mammals and no TTS or PTS data for sea turtles. Therefore, weighting function derivations relied upon auditory threshold measurements (audiograms), equal latency contours, and predicted audiograms from anatomically based models, as well as TTS data when available*." It is exactly this lack of experimental data and resulting requirement to use approximations, interpolations and assumptions that need to be quantified and discussed if applied in the determination of auditory weighting functions and subsequent numeric thresholds for predicting auditory effects.

> ***Response (NOAA & Navy)***: There is certainly a limit to available data and uncertainty in the resulting predictions. Unfortunately there are too few data regarding variation in TTS onset across individuals and species within the same group to properly assess the expected variability and provide confidence limits on the predictions for TTS/PTS onset. However, in using an equal energy model for sound accumulation, this approach overestimates the effects of the short duration and intermittent exposures typically produced by the most powerful sources (e.g., Navy sonar). This conservative approach therefore helps to mitigate the risk associated with the limited TTS data.
>
> The equal energy approach assumes that the effects of noise can be predicted based on SEL alone, regardless of how the SEL is distributed across time. For intermittent exposures, the equal energy approach ignores the quiet periods between sounds (where hearing can partially recover) and therefore over-estimates the effects of intermittent noise. Marine mammal TTS data have also shown larger amounts of TTS and lower TTS onset SEL values when the noise duration is longer (i.e., for the same SEL, longer duration exposures result in lower TTS onsets). Most marine mammal TTS data have been obtained using exposure durations of 16-s or more (some with durations of an hour or longer), much longer than the durations of many of the most powerful sources. The use of these longer-duration data will therefore also tend to over-estimate the effects of most typical sources.

REVIEWER 4

Comment 15: Auditory weighting function was developed for humans based on the results equal loudness experiments at different frequencies. Auditory response time can be used to approximately equal loudness results. Temporary threshold shift data as well as hearing sensitivity data can be used to determine auditory weighting function. Data from TTS and audiograms are the least favorable and most inaccurate means of determining auditory weighting function. The most critical data required are TTS onset exposure levels as a function of exposure frequency. However, TTS has been tested for a small number of frequencies although it has been common to used octave band noise. Therefore, from the small number of frequencies test, one must assume that TTS would be the same at other frequencies.

Response: NOAA agrees that understanding TTS onset as a function of exposure frequency is important and that there are limited data available from marine mammals. However, for species, such as the bottlenose dolphin, a broad range of tones have been tested (0.4 to 80 kHz), illustrating that TTS onset is frequency-dependent (Finneran and Schlundt 2010; Finneran and Schlundt 2011; Finneran and Schlundt 2013). It is specifically these studies, among others, that were used to create updated marine mammal auditory weighting functions. Based on the data used to derive auditory weighting functions (e.g., direct TTS measurement for onset and growth and composite audiograms), it is not assumed that TTS is the same all frequencies.

3. METHODOLOGY TO DERIVE FUNCTION PARAMETERS

REVIEWER 2

Comment 16: As per my comment on section 1.4, I feel that additional information there or here would be helpful, especially to a first time reader.

Response: Thank you. Please see our response in Section 1.4.

4. MARINE MAMMAL SPECIES GROUPS

REVIEWER 1

Comment 17: Audiometric Data Sets. The document does an excellent job incorporating data sets from key labs and teams around the world. The authors were careful to assure that subjects that have been described more times within the literature (i.e., they have participated in multiple studies) did not contribute multiple audiograms to this analysis and overly weight the resulting means.

The audiometric data are presented in two different ways. First, the absolute sensitivity is provided (Figure 5). Here, thresholds are plotted for all available individual animals across publications. The composite audiogram is typically the line of best fit calculated per equation 3, expect in the case of sea turtles for which equation 3 did not provide a successful fit. For sea turtles, median thresholds were used. The use of median data reduces the potential effects of outliers. This could be an issue with the sea turtle data, as there were several subjects that had higher thresholds. In the second analysis, the thresholds are normalized (Figure 6). Essentially, they are plotted as dB relative to the best threshold, such that the best threshold is set as 0, and the other thresholds are expressed as dB difference from that best threshold. The composite audiogram is then calculated for the individual normalized audiograms as either a line of best fit, or in the case of sea turtles, the median, for the individual functions. The original audiograms are useful in that both the frequency-based differences and overall differences in sensitivity are available. The normalized audiograms are particularly useful for assessing low frequency slope of the audiogram (dB change in sensitivity per Hz of frequency change) which tends to be shallower, and the high frequency slope which is much steeper.

Response (NOAA & Navy): Thank you for your assessment of the absolute sensitivity and normalized audiometric data. All composite audiograms were based on median thresholds.

> For sea turtles, the median values were used directly. For others, a curve was fit to the median values. The median was used to reduce specifically the impact of outliers.
>
> NOAA will re-evaluate the Reviewer's comments regarding sea turtles when NOAA develops national acoustic guidance for sea turtles (i.e., comments relating to sea turtles will not be addressed in this Peer Review Report).

Comment 18: The data used to calculate composite audiograms were largely limited to behavioral data. If at least three behavioral audiograms were available, AEP data were excluded as thresholds tend to be higher, and errors increase as low frequencies, which could bias weighting functions at lower frequencies. Given adequate behavioral data were available, AEP data were not used to generate any of the composite audiograms, except for sea turtles, for which there was insufficient behavioral data. Given the argument that AEP data systematically introduce error, particularly at low frequencies, this raises questions about the sea turtle data. Sea turtles have audiograms that are more "V" shaped than "U" shaped in Figure 5, and the normalized function is essentially flat at the low frequencies in Figure 6 for the sea turtles but not for other species. Is this a quantitative difference in threshold sensitivity? Or could this in part be an artifact of the low frequency issues with AEPs used for sea turtles and not for other species? This is an important question, as the author later notes that turtles are more like fish than marine mammals with respect to auditory function and vulnerability to noise, a conclusion first suggested by a working group on sound exposure guidelines for fish and sea turtles. Given the differences in the type of threshold data that are available, the lack of fit for the current models to the sea turtle audiograms, the differences in overall sensitivity, and the differences in the anatomy of the ear in turtles and marine mammals, the current methods appear to be much more poorly suited to questions regarding effects of noise on sea turtle. It would have perhaps been helpful if composite thresholds generated using only AEP data were generated in order to assess the differences introduced by use of evoked potential testing, which is much quicker and much less costly than training animals to perform behavioral tasks to provide threshold data.

> ***Response (NOAA & Navy)***: We thank you for your assessment. General relationships between AEP and behavioral thresholds are reasonably well understood only for mammals, where the biomechanics of the cochlea cause the traveling wave to slow down as it progresses from the base (high-frequency region) to the apex (low-frequency region), resulting in a loss of synchrony in low-frequency neural responses and an inability to measure these responses using surface electrodes placed on the head. The text in Section 5 of the Navy Technical Report has been modified to clarify this point.
>
> NOAA will re-evaluate the Reviewer's comments regarding sea turtles as NOAA develops national acoustic guidance for sea turtles (i.e., comments relating to sea turtles will not be addressed in this Peer Review Report).

REVIEWER 2

Comment 19: I agree that the species group concept is the best way to proceed at this time. While there is a lot of variation in the audiograms of some groups, especially the otariids and other non-phocid marine carnivores (Figure 4), it would not be practical to attempt to derive species-specific weighting or exposure functions until much more data become available. Doing so would give the appearance of greater accuracy than is currently possible.

Response: NOAA agree with your assessment and has no intention, at this point (due to limited data, as mentioned by the Reviewer), to develop species-specific weighting functions. Our intent is to base auditory weighting functions on broader functional hearing groups.

REVIEWER 3

Comment 20: The grouping of marine species in nine functional species groups seems well supported and all data on auditory sensitivity in marine species are carefully considered to create the composite audiogram and determine the weighting function parameters. The approach used to derive the weighting functions and subsequent numeric thresholds for predicting auditory effects for all the different types of sound exposure are clearly stated (also explained graphically) in all its details, the equations used are (mostly) well explained and/or deduced.

Response: NOAA thanks the Reviewer for their comment and agrees.

REVIEWER 4

Comment 21: Given this sparse information environment, Finneran attached [attacked] the task of estimating auditory weighting function in a systematic manner collecting all the hearing data available for marine mammals. He divided marine mammals into 9 groups depending on frequency of hearing and taxa which seems logical and reasonable. Cetacean were divided into three groups, pinnipeds were divided into two subgroups (phocids and otariids) which were further subdivided for hearing underwater and in air and sirenians. Sea turtle was in a group by itself. One can quibble on the need to have two groups for pinnipeds since they could be easily thrown into the same group. Their upper frequency of hearing are approximately in the same range. Estimating (best guess) separate weighting functions for underwater and in-air hearing is a good way to go.

Response: NOAA thanks the Reviewer for their comment and agrees with their conclusion.

4.1. Low-frequency (LF) cetaceans

REVIEWER 4

Comment 22: The major objection I have of this work has to do with how baleen whales are group together which makes no sense. Minke and humpback whales produce sounds that are at least one decade higher in frequency than blue, fin and perhaps sei whale. One can guess that hearing sensitivity would have some rough correspondence to the sounds that a specie[s] produces. One decade of frequency different is a large number and to group all these whale together is not reasonable. I agree that we have no idea what the audiogram of a baleen whale should be however, any rationale analysis would indicate that they should be separated. From my perspective, estimation of hearing sensitivity by Dr. Darlene Ketten and Dr. Ted Cranford based on anatomy, density of bones, structure of the assumed auditory pathways and finite element modeling of vibration have resulted in audiograms that I believe are too biased toward high frequencies and tend to suggest that smaller baleen whales like minke and humpback whales have similar audiograms.

> ***Response***: NOAA acknowledges that, as more data become available, marine mammal hearing ranges may require future modification and that it may be necessary to divide LF cetaceans into subdivisions. However, at this time, NOAA does not believe there are enough data to support further LF cetacean divisions and subsequent auditory weighting functions, especially since no direct information on hearing is available for this group. This particular topic will be highlighted during the upcoming public comment period as an area where input will be particularly valuable.

5. COMPOSITE AUDIOGRAMS

REVIEWER 2

Comment 23: I agree with the limited use of AEP data, especially as this methodology cannot operate at low frequencies. The use of only recent airborne phocid data is appropriate as the earlier in-air audiograms likely were elevated due to a methodological problem. The available data were used appropriately to estimate the LF cectaean and turtle audiograms.

> ***Response***: NOAA thanks the Reviewer for their comment.

Comment 24: It is appropriate to use the median value at each frequency for the group audiograms depicted in Figure 5. Using the lowest threshold value per frequency present in the data set would likely result in an overestimation of the sensitivity of the hearing abilities of that group and give too much emphasis to the experimental errors associated with the individual threshold determinations.

> ***Response:*** NOAA thanks the Reviewer for their comment.

Comment 25: The audiogram data (figure 5) and composite audiograms using normalized thresholds, per subject (Figure 6) indicate the low and high frequency rolloff rates for most groups. Fitting the data to equation 3 permits accurate measures of the rolloff values. The high R^2 values shown in Tables 3 and 4 indicate a high level of agreement between the measured and calculated data sets.

> ***Response***: NOAA thanks the Reviewer for their comment.

Comment 26: In Table 5, the lowest *So* values of the slopes per decade divided by 20 were used to determine the value "*a*" in Equations 1 and 2 and in Table 8 except for LF cetaceans where the slope of 30 dB/decade was assumed. This is a conservative approach in that it increases the emphasis of lower frequency noise. See section 9.1 for corrections to some values.

> ***Response***: NOAA thanks the Reviewer for their comment. The suggested corrections will be addressed in Section 9.1.

REVIEWER 4

Comment 27: Finneran then proceeded to group audiograms for the different groups of marine mammals and determine the "average" audiogram for the animals in the different groups. A generic weighting function was used and the available audiograms of animals in any particular groups was used to adjust and determine the appropriate parameters. The weighting function is nothing more than a bandpass filters that is used for mammal auditory system including humans and is given by the equation

$$W(f) = C + 10\log_{10}\left\{\frac{(f/f_1)^{2a}}{[1+(f/f_1)^2]^a\,[1+(f/f_2)^2]^b}\right\}$$

where C is a constant affecting the level of W(f), f_1 is the low frequency cutoff and f_2 is the high frequency cutoff, a is a parameter determining the rate of increase of the W(f) in the low frequency cutoff region and b is the parameter determining the rate of decrease in the high frequency cutoff region.

In this way, Finneran was able to provide estimate weighting function parameters for the nine groups of animals. The different steps in estimating the 5 parameters in the weighting function was described in detail by Finneran. All things being equal, the approach taken by Finneran is about the best one could do given the paucity of real hearing data.

> ***Response***: NOAA agrees. Thank you for your comment.

Comment 28: On the whole, Finneran has done a tremendous job in collecting all the information on hearing in marine mammals. Just accumulating of all the marine mammal hearing data in one volume is a worthwhile accomplishment. The placement of sea turtle with marine mammals is not justifiable and this species seem to have very poor hearing that they should have never been group together with marine mammals.

> ***Response***: NOAA agrees. Thank you for your comment. The method in the Navy Technical Report is applied to sea turtles with minor adjustments based on limited data. Nevertheless, this report works to use sea turtle-specific data or data from fishes as suggested by the recent ANSI panel (Popper et al. 2014) to predict auditory weighting functions and thresholds, rather than using marine mammals as a surrogate for sea turtles. These factors will be considered as NOAA develops national acoustic guidance for sea turtles.

6. EQUAL LOUDNESS DATA

REVIEWER 2

Comment 29: The relevant data are acknowledged to be too sparse to be extrapolated to other species groups.

Response: NOAA agrees. Thank you for your comment.

REVIEWER 3

Comment 30: The current equal loudness data are based on studies in few (usually one or two) animals per functional category (Finneran & Schlundt, 2010[5]; Reichmuth, 2013) or modeling efforts (Ketten et al., 2014[6]; Cranford & Krysl, 2015). Intra- and inter-subject variability cannot sufficiently be taken into account based on this data alone[7]. While these results represent important input in the context of defining underwater noise criteria for marine species, the small sample size in itself is indicative of potential variations if repeated in a larger number of subjects. Suzuki and Takeshima (2004) demonstrated how influential defining studies (such as Finneran's report) can be: According to Suzuki and Takeshima (2004), the A-weighting which has been used over the past decades in human audiometry is based on Robinson and Dadson (1956). While this study as well as others studies (referred to by Finneran in this context)[8, 9] are already based on variable data sets, Suzuki and Takeshima's (2004) review of more recent studies on human (and non-human primates) equal loudness studies reveals differences up to 20 dB (at 40 phon) in comparison. If transferred to e.g. the data collected by Finneran et al. (2011) and the subsequent use for defining the new Phase 3 weighting approach, such differences would likely result in differences in numeric exposure thresholds by the same order of magnitude.

Response (NOAA & Navy): NOAA agrees with the Reviewer that there are limited equal loudness data available for marine mammals, which is why the Navy did not rely on these data to estimate the shapes of loudness contours and weighting functions at lower frequencies (i.e., equal loudness data were provided for information and comparative purposes only and were not directly used in establishing any parameters of the weighting functions).

The reviewer's comments highlight the need to periodically revisit criteria/threshold development as new data become available. This is accomplished by Navy via the TAP process, where data are reviewed in each new TAP cycle and criteria/thresholds adjusted accordingly. The Navy has adjusted its methodology and criteria/thresholds from Phase 2 to Phase 3 because additional data became available; similarly, Phase 3 criteria/thresholds may

---

[5] NOAA believes the Reviewer meant to refer to Finneran and Schlundt 2011.

[6] Referred to in Finneran (2015) as "Ketten & Mountain 2014".

[7] Finneran & Schlundt (2011): "[…] confidence intervals were larger (8-10 dB) at the lower and higher frequencies, especially at the lower loudness levels (90 and 105 dB re 1 µPa).

[8] See also Stebbins (1966): "The general characteristics of the present equal-latency contours are not inconsistent with the Fletcher-Munson curves (Fletcher and Munson, 1933) although for one subject they are less flat at the higher intensities."

[9] Pfingst et al. (1975): "Dependency of RT on frequency and intensity of the stimulus is evident. RT decreases as a function of intensity, and the position with respect to intensity and slope of each L-1 function is dependent on frequency. These data are representative of those seen in all subjects tested. RTs, especially those to higher intensities of stimulation, varied among subjects and were generally faster for humans than for monkeys. However, the relative positions and the relative slopes of the L-1 functions were similar for all subjects."

be adjusted in the future as additional data become available. For NOAA, the Acoustic Guidance includes information on how it will be updated as new data become available.

Comment 31: Moreover, the rationale for using equal loudness contours as a proxy for assessing auditory effects is a first order approximation (see above). Taking the subjects' reaction time as proxy for equal loudness as proxy for auditory weighting functions to assess safe exposure limits is another order of approximation. Once more, it has to be acknowledged that making these approximations is the best available science, but is increasing the uncertainty included in the determination of the auditory weighting curves. As Finneran & Schlundt (2011) state, "*the increased variability at the higher frequencies was a result of the increasing difficulty in making the loudness comparison when the comparison tone frequency was much higher than the reference frequency of 10 kHz. For this reason and to provide equal loudness contours more suitable for populations-level application, fitting parameters were smoothed so the equal loudness contours would not exhibit large fluctuations across frequency.*" All these data are carefully considered and correctly used by Finneran in his report, but any variability in the defining parameters (low- and high-frequency roll-offs [a, b] as well as low- and high-frequency cut-off frequencies (f_1 and f_2]) could result in substantial changes in the numeric threshold values and consequently in potential auditory effects. (i.e. amount of TS elicited).

> ***Response (NOAA & Navy)***: The dolphin equal loudness data were not directly used to derive the Phase 3 functions. Equal latency data were only used if they suggested a shallower low-frequency slope compared to the composite audiograms. These data were therefore used to reduce the likelihood of significantly under-estimating TTS onset levels at the lower frequencies.

Comment 32: Further approximation and assumptions had to be made by Finneran adding more uncertainty to the resulting auditory weighting functions, e.g. in defining f_1 and f_2 by fitting ΔT, in determination of the composite audiograms (even though sensible, this method is based on various assumptions [depending on the functional species group] and averages/fitting procedure over a limited number of subjects). For turtles, the composite audiogram had to be based on median values and the estimation of lowest threshold for FL[10] cetaceans was based on ambient noise considerations.

> ***Response (NOAA & Navy)***: As stated in the Navy Technical Report, the best approach would be to fit directly functions to TTS onset data, as a function of frequency, for all species groups. Unfortunately, this is not possible due to the limited amount of data. As a result, multiple sources of audiometric data were considered to estimate TTS onset levels as a function of frequency. Additional sources of data, such as equal latency contours, were incorporated in a way to help ensure that TTS onset levels were not significantly underestimated at frequencies where data did not exist. The resulting exposure functions were fit to TTS onset data when possible to ensure the functions properly represented the existing data. At other frequencies and for species groups with no TTS data, we relied on analogies in auditory system function across mammalian species and similarities between the auditory systems of turtles and fishes. We believe this is the most appropriate approach given the underlying data.

---

[10] NOAA is assuming this is a typo by Reviewer 3 and that the intent was to refer to low-frequency (LF) cetaceans.

7. EQUAL LATENCY DATA

REVIEWER 2

Comment 33: The equal latency data are only available from four species. The line intersecting ~40 dB above the most sensitive frequency was selected for use. The 40 dB level is based on the human dB(A) weighting scale concept and values at higher dB levels were inconsistent. This inconsistency is not unexpected because the measure of latency of response to the detection of a low amplitude test signal would be expected to plateau once the test signal level is clearly above the 50% detection threshold.

Response: NOAA thanks the Reviewer for their comment.

Comment 34: The selection of the lowest slope of the composite audiogram, the normalized data composite audiogram or the equal latency lines to estimate the low frequency rolloff is a conservative move in that the greater/steeper the rolloff, the greater the reduction in the impact of low frequency sound by the weighting/exposure function.

Response: NOAA thanks the Reviewer for their comment.

8. TTS DATA

REVIEWER 1

Comment 35: TTS Onset. The author has selected 6-dB as the definition of TTS onset. This decision is based on test-retest data, with thresholds assumed to be repeatable within +/- 5 dB, in the absence of any insult that would induce hearing loss. Within any single subject, it is reasonable to assume that changes of 5-dB or less will be difficult to measure with any confidence. At the group level, it is possible to measure an average change of less than 6-dB. Because test-retest can be plus or minus 5-dB, the average difference across animals should be approximately 0-dB if the individual differences are random in nature. However, if all, or even most, animals had a small change in the same direction, i.e., slightly poorer post-noise thresholds, this would result in an average change that is different from 0. This is not an important issue with respect to interpretation of TTS onset, as it seems unlikely that TTS of less than 6-dB would have any impact on foraging, survival, or other relevant metrics of behavior. Moreover, in many cases the available data are limited to a small number of animals where smaller changes presumably cannot be reliably identified. If data from a larger number of animal subjects were available, it may be possible to identify the onset of smaller TTS changes, but given the state of the field and the current availability of TTS data, the defining the onset of TTS as beginning at 6-dB is reasonable. A key question not considered here is whether a 6-dB TTS has the potential to effect foraging success, predator detection/avoidance, or other key outcomes. If there are no deficits with a 6-dB TTS, a standard that is based on a goal of limiting sounds to levels that induce 6-dB TTS or less may be overly conservative. While this may be overly conservative, we do not know what would constitute a biologically significant TTS. Thus, the conservative approach used here, based on the smallest measurable TTS, is reasonable since the point at which overexposure becomes hazardous is not known.

Response: Thank you for your comment. NOAA agrees and has used a 6-dB threshold shift to indicate the onset of TTS for some time. The key question as to the biological significance of a 6-dB threshold to animals in the wild and whether this deficit would impact vital life functions is difficult to answer based on available data. Nevertheless, based on limited data and uncertainty, NOAA has decided to conservatively use a 6-dB threshold shift to represent TTS onset.

Comment 36: The author notes that AEP metrics result in greater TTS estimates than do behavioral studies (Figure 9). This is an interesting phenomena and raises questions about the source of the difference. AEP studies frequently use brief tone pips, repeated at a rapid rate, whereas behavioral studies often use longer duration tones. The difference between AEP and behavioral measures of TTS may perhaps reflect some effect of noise on temporal processing, and raises questions about the biological relevance of the signals in the two types of tests. If the greater deficits that are captured in AEP studies are biologically relevant, it may be more appropriate to consider the AEP data rather than the behavioral data when defining onset of risk. It would be more conservative to use the AEP data that suggest greater effects of noise on the auditory system. One reasonable approach might be replication of the analyses of the behavioral data instead using any available AEP data on TTS, in an effort to assess potential differences in conclusions about where risk begins if using AEP data, with candid discussion of unknowns regarding where risk begins.

Response (NOAA & Navy): There are insufficient data to so this. Only a single TTS onset value can be determined from the AEP data — most studies utilized a limited number of exposure levels or featured exposures too large to determine the onset of TTS. The goal of the analysis is to identify exposures likely to result in the onset of TTS (6 dB of behavioral TTS) and PTS (40 dB of behavioral TTS), not to identify the lowest exposures resulting in potential disruption of auditory system function.

Comment 37: Although the author states that calculation of TTS onset was limited to behavioral data sets (line 550 on page 23), TTS onset is shown for both behavioral and AEP data, and TTS onset is indicated for both types of data sets in Table 6. The plots are somewhat difficult to fully understand. Figure 10 plots TTS growth for mid-frequency cetaceans using behavioral methods and Figure 11 plots TTS growth for mid-frequency cetaceans using AEP methods. The legend for Figure 10 does not specify whether frequencies indicated in the plots are exposure frequency, test frequency, or both. Figure 11 specifies that frequencies indicated in the plots are exposure frequency but does not specify test frequency. The specific frequencies, and the range of frequencies, included in the two figures are different (Figure 10: 3 kHz-56.6 kHz; Figure 11: 11.2 kHz-90 kHz). In Figure 12, which plots data for high-frequency cetaceans, behavioral and AEP are both plotted (with Table 6 providing information on which plots are behavioral and which are AEP based). The legend for Figure 12 helpfully specifies that the exposure frequency is in normal font and the test frequency is in italics. Given differences in behavioral and AEP metrics, species, exposure frequency, and test frequency, it is difficult to draw any specific conclusions about the extent to which behavioral and AEP metrics influence TTS onset. Figures 18-20 perhaps help to resolve some questions as behavioral and AEP data are plotted together

Response (NOAA & Navy): The figures in the Navy Technical Report include all data for which growth curves could be generated, regardless of whether they represented AEP or behavioral data. They are organized by species group and methodology. For MF cetaceans, where more data exist, the plots were split over multiple figures rather than a single figure with 15 panels. For other groups, where fewer data exist, data were plotted in a single figure. The same specific frequencies could not be provided in Figs. 10 and 11 because the data do not exist. The intent is not to directly compare AEP and behavioral data, but to provide the reader with the TTS data and resulting growth curves for all existing data. The legend for Fig. 10 has been revised to better explain the figure contents.

8.1. Non-impulsive (steady-state) exposures – TTS

REVIEWER 2

Comment 38: The use of cumulative SEL, and thus the "equal energy" approach, is acknowledged as being a simplifying assumption. The availability of a single metric to indicate sound level, frequency and duration combinations that would be expected to result in 6 dB TTS would facilitate setting regulations and guidelines to control noise exposures. Such a measure has great value in that it would be possible to calculate the cumulative SEL from the sound characteristics alone and would not require the (often impractical) actual measurement of TTS exposures with captive animals.

> ***Response***: NOAA acknowledges the limitations and simplifying assumptions associated with relying on the $\mathrm{SEL}_{\mathrm{cum}}$ metric (i.e., equal energy hypothesis (EEH), as mentioned by the Reviewer), which is why NOAA acoustic threshold levels are also expressed as a peak pressure metric in the Acoustic Guidance. However, NOAA believes, in many cases the EEH approach functions reasonably well as a first-order approximation, especially for higher-level, short-duration sound exposures such as those that are most likely to result in a TTS in marine mammals. Additionally, there is no currently supported alternative method to accumulate exposure available, which we recognize as an important consideration for noise-induced hearing loss. If alternative methods become available, they can be evaluated when the Acoustic Guidance is updated.
>
> NOAA agrees, as the Reviewer indicates, that the $\mathrm{SEL}_{\mathrm{cum}}$ metric is valuable in terms of action proponents being able to calculate values from various source characteristics (i.e., practical consideration) vs. relying on actual measurements, from captive individuals, reflecting numerous types of available sources. We appreciate your understanding of practical considerations necessary when implementing TTS thresholds.

Comment 39: The selection of 6 dB TTS as being a biologically important level is debatable but for practical reasons some level above a just measurable amount is required. A 6 dB loss in sensitivity would equate to a halving of the acoustic communication range (assuming spherical spreading). 6 dB TTS is located above the inflection point of equation 4 (Figure 10) and above the threshold determining accuracy of the various threshold measuring techniques.

> ***Response***: NOAA thanks the Reviewer for their comment. We appreciate your understanding of practical considerations necessary when implementing TTS thresholds. NOAA agrees that the biological relevance of a 6-dB shift has yet to be determined. Nevertheless, based on limited data and uncertainty, NOAA has decided to conservatively use a 6-dB threshold shift to represent TTS onset.

Comment 40: The use of SEL is acknowledged to be a simplifying assumption (line 583) and other authors have shown that the equal energy hypothesis has problems. Kastak et al. (2005) suggest that "moderate levels of long duration sounds may have a greater impact on hearing than equal-energy sounds of greater amplitude but shorter duration". A harbor porpoise exposed to a variety of duty cycles, durations and amplitudes between 199 and 202 dB SEL, experienced TTS at 4 minutes post exposure between 7 and 32 dB (Figure 11 in Kastelein et al. 2014[11]).

> ***Response***: As mentioned in our response to the Reviewer's first comment on this section, NOAA acknowledges the limitations and simplifying assumptions associated with relying on the SEL_{cum} metric but believes it functions reasonably well as a first-order approximation. The Acoustic Guidance also highlights the factors the Reviewer mentions in several sections of the document (e.g., Qualitative Factors for Consideration and various sections relating to the importance of exposure duration).

Comment 41: The selection of a 1 second time interval in calculating SEL is an arbitrary duration, and probably is longer than the integration time of the listener. After some long duration exposure, a plateau in TTS is likely, so it may be appropriate to put some upper duration on the time over which the SEL calculation is being made. Without a description of the precise nature of the sonar signals (i.e., source level, frequency, duration, duty cycle, etc.) or how these signals change over distance, it is difficult to determine the impact on hearing just based on SEL measures. Where harmonic signals with durations greater than 5-10 cycles are defined as being steady state (line 158), the overlap between steady state sounds and impulsive sounds, and their differing SEL exposure levels, is difficult to assess.

> ***Response***: Within the Acoustic Guidance, NOAA defines cumulative SEL based on previously established definitions (EPA 1982; ANSI 1995; ANSI 2013), which are typically associated with a reference time of one second. However, NOAA intends that cumulative SEL account for accumulation over the recommended baseline accumulation period of 24 hours.
>
> Regarding the Reviewer's comment relating to overlapping sources, NOAA's recommended application of the SEL_{cum} metric is for individual activities/sources. It is not intended for accumulating sound exposure from multiple activities occurring within the same area or over the same time or to estimate the impacts of those exposures to an animal occurring over various spatial or temporal scales. Current data available for deriving acoustic threshold levels using this metric are based on exposure to only a single source and may not be appropriate for situations where exposure to multiple sources is occurring. As more data become available, the use of this metric could be re-evaluated, in terms of appropriateness, for application of exposure from multiple activities occurring in space and time.
>
> Regarding how a signal changes overs distance, NOAA recently conducted a third independent peer review in association with the Acoustic Guidance which examined how the injurious characteristics of impulsive sounds change with distance from the source (See Peer Review Report for this review at : http://www.cio.noaa.gov/services_programs/prplans/ID43.html).

---

[11] NOAA believes the Reviewer is referring to Kastelein et al. 2014a.

REVIEWER 3

Comment 42: The second aspect is the definition of injury as used in Finneran's report (as in the US regulation in general). Southall et al. (2007) states that "*Noise-induced PTS represents tissue injury, but TTS does not. Although TTS involves reduced hearing sensitivity following exposure, it results from primarily from the fatigue (as opposed to loss) of cochlear hair cells and supporting structures and is, by definition, reversible (Nordmann et al., 2000). Many mammals, including some pinnipeds (Kastak et al., 1999, 2005) and cetaceans (e.g., Schlundt et al. 2000; Nachtigall et al., 2004), demonstrate full recovery even after repeated TTS. Since TTS represents a temporary change in sensitivity without permanent damage to sensory cells or support structures, it is not considered to represent tissue injury (Ward, 1997). Instead, the onset of tissue injury from noise exposure is here considered PTS-onset.*"

This quote from Southall et al. (2007) comprises the scientific basis for the current definition of injury, but new findings (Kujawa & Liberman, 2006, 2009; see also Liberman 2013) show long-term effects of repeated overexposure of the hearing system. Even if hair cells recover, overexposure can cause synaptic loss at the hair cells and subsequent cochlear nerve degeneration. Consequently, even if a subject's hearing threshold recovers (definition of TTS) its capability for hearing in noise can be compromised in the long-term. The current damage-risk criteria assume that threshold recovery represents cochlear recovery, but Kujawa & Liberman's study (2009) shows that this might be false and overexposure to noise may be more dangerous than previously thought.

> ***Response***: NOAA currently does not consider a TTS an auditory injury. This is based on the work of a number of investigators that have measured TTS before and after exposure to intense sound.
>
> The Acoustic Guidance includes and acknowledges information from Kujawa and Liberman (2009) and Lin et al. (2011) illustrating the complexity associated with noise induced hearing loss, including the effects associated with large TTSs, and identifies this as an area where more research and examination is needed (i.e., Appendix that identifies Research Recommendations). The large threshold shifts (i.e., maximum 40 dB[12]) that led to the synaptic changes shown in these two studies are in the range of the large shifts used by Southall et al. (2007) and within the Acoustic Guidance to define PTS onset (i.e., 40 dB). It unknown whether smaller levels of TTS would lead to similar changes or the long-term implications of irreversible neural degeneration. NOAA agrees these studies are appropriate for use as qualitative considerations within a comprehensive effects analysis.
>
> More importantly, NOAA's Acoustic Guidance acknowledges the potential for injury from noise, and the protocols for examining marine mammal hearing loss data and deriving TTS onset thresholds included several conservative assumptions (e.g., using a 6 dB threshold shift to represent TTS onset, not directly accounting for exposures that did not result in threshold shifts, etc.).

---

[12] The exposure levels used in Lin et al. 2011 were "within 3 dB of the boundary between reversibility and irreversibility, at least with respect to the threshold for ABRs and DPOAEs."

8.2. Non-impulsive (steady-state) exposures – PTS

REVIEWER 1

Comment 43: PTS Onset**.** The author makes the case that a 40-50 dB TTS is likely the border at which a TTS has the potential to become a PTS. This conclusion is based on data from humans, and common laboratory rodent subjects such as rats, mice, guinea pigs, or chinchilla which typically show full recovery after even more robust TTS deficits immediately post noise. In the absence of data from marine mammal species, it is reasonable to use the suggested 40-dB as a conservative prediction of where PTS may begin. There are multiple unknowns, acknowledged by the authors, and it is appropriate to be conservative in generating weighting functions. Conceptually, weighting functions discount energy at frequencies where the energy is assumed to be less hazardous based on poorer threshold sensitivity at those frequencies. Application of weighting functions therefore ultimately serves to allow more noise in the environment. The careful use of conservative assumptions throughout the document serves to reduce the potential that weighting functions would inappropriately discount noise that is in fact hazardous to members of a given species. Because TTS growth is not linear from smaller to greater TTS values, PTS was inferred using TTS growth functions to interpolate the sound level that would result in a 40-dB TTS using only those exposures that resulted in at least 20 dB TTS. This is important; if the author had based the interpolations on small TTS values, the assumed TTS growth would be much shallower, and would lead to incorrect conclusions about the sound level at which risk of PTS begins (based on observed or expected TTS deficits of 40-dB or greater). This is highlighted in Figure 10.

> ***Response***: Thank you for your comment. NOAA agrees that a 40-dB threshold shift is a conservative prediction for the onset of PTS and the methodology provided in the Navy's technical report. For marine mammals, noise-induced thresholds shifts of 40 dB or higher have been measured and resulted in recovery (Kastelein et al. 2013a; Popov et al. 2013; Popov et al. 2014). The individuals in these studies were exposed at SEL_{cum} at levels above some of the updated PTS onset levels (i.e., even though exposed at level which would exceed the updated PTS onset acoustic threshold level, the animal still recovered completely).
>
> NOAA agrees with the Reviewer that auditory weighting functions discount sound energy at frequencies that are considered less hazardous for a particular marine mammal functional hearing group. However, NOAA does not agree with the Reviewer's assessment that "Application of weighting functions therefore ultimately serves to allow more noise into the environment." The weighting functions reflect frequencies where functional hearing groups are most sensitive to sound in terms of hearing and vulnerability to noise-induced threshold shifts. Thus, weighting functions reflect the auditory abilities and susceptibilities of an individual within that functional hearing group. An action proponent may use weighting functions when determining isopleths associated with a threshold level associated with a particular functional hearing group and activity (i.e., modeling of the area impacted around a source). Auditory weighting functions results in smaller isopleths for frequencies where a functional hearing group is less susceptible. However, invoking weighting functions do not alter any of the characteristics of a particular sound source (i.e., it is misleading to interpret the application of these functions as allowing more noise into the environment). Instead, auditory weighting functions allow for a more accurate representation of the potential for a

sound source to affect a particular functional hearing group (i.e., it does not result in any changes to the sound source or receiver).

Thank you for your comment relating to the use of studies with at least a 20-dB threshold shift to interpolate PTS onset. NOAA agrees with Reviewer's assessment.

Comment 44: Risk of PTS is assumed to begin with 40-dB TTS. Based on TTS growth functions for non-impulsive sounds, the author proposes a constant of 20-dB for the difference between TTS onset and PTS onset. The 20-dB constant differs from the earlier Phase 2 proposal, which included 20-dB difference for cetaceans and sea turtles, and 14-dB difference for other species. Increasing the value from 14 to 20 shifts the point at which PTS is assumed to begin to a higher SEL; i.e., 20-dB above TTS-onset, vs 14-dB above TTS onset. The less conservative approach is justified by mean differences of 25 dB, median differences of 25 dB, and a range of values from 14-37 dB, with all but one value being greater than 14 dB. It is not clear why the 14-dB value was not used for the one species with seemingly greater vulnerability to PTS, based on smaller range between where TTS onset begins and where PTS onset is assumed to begin based on TTS growth function. The single value approach is intuitively appealing in its simplicity, but if a single number is preferred, perhaps that number should be based on the most vulnerable species (the species with most rapid TTS growth) rather than the majority of species. Application of a single value to the calculation of PTS onset for impulse noise is more conservative, with a 15-dB constant added to the SEL-based TTS onset threshold, or a 6-dB constant added to the peak pressure TTS onset threshold. The 15-dB constant was selected as an arbitrary value. The data suggested a value of 46-dB, but the resulting values were deemed too high, and a smaller constant selected.

Response: The one study that indicated a difference of 13-dB was for a bottlenose dolphin and found an unusually high growth rate for exposure to a tone at 28.3 kHz (Finneran et al. 2013). Thus, this particular exposure scenario is not considered representative of most (Note: all other growth rates in this study were 1 dB of TTS/dB of noise or less). The authors attribute these results to exposures possibly exceeding the critical level (i.e., where damage switches from being primarily metabolic to more mechanical) at this frequency or exposure scenario. Having data from another individual of this species at this exposure frequency or exposure scenario would inform the trend observed in this particular individual. Until that time, NOAA will use the more supported terrestrial mammal growth rate.

NOAA disagrees with the Reviewer's evaluation that the selection of a 15-dB constant between TTS and PTS is arbitrary. For impulsive sources, Southall et al. (2007) recommended a growth rate of 2.3 dB of TTS/dB of noise. This rate was somewhere in between previously recorded rates below (range from 0.7 to 1.9 dB of TTS/dB of noise) and above (range from 2.6 to 7 dB of TTS/dB of noise) the critical levels for terrestrial mammals (Henderson and Hamernik 1982; Henderson and Hamernik 1986; Price and Wansack 1989; Levine et al. 1998; Henderson et al. 2008). Southall et al.'s (2007) recommendation resulted in a more conservative acoustic threshold level associated with PTS onset than choosing a growth rate below the critical level based on terrestrial data. Thus, NOAA accepts the recommendation made by Southall et al. (2007) as guidance for determining PTS onset levels for impulsive signals for all cetacean and underwater pinniped species, resulting in an approximate 15 dB difference between TTS and PTS onset threshold levels. NOAA has identified threshold growth rates and recovery as an area where more research is needed.

REVIEWER 2

Comment 45: The SEL at which PTS onset occurs is described in lines 640-645 as being the SEL where 40 dB of TTS would occur, which would be 34 dB of TTS above the "TTS onset" level (6 dB TTS). PTS onset must refer to the SEL level at which a small PTS could occur and not a PTS of 6 dB. Changing the term "TTS onset" (as mentioned above) would help clarify the terminology here.

> ***Response***: The interpretation provided is correct (i.e., PTS onset refers to 34 dB of thresholds shift above TTS (defined as a 6 dB threshold shift)). NOAA considers the terminology of "TTS onset" preferable and more intuitive to readers of the Acoustic Guidance compared to using "TTS 6 dB."

Comment 46: The criterion of using only data sets with TTS above 20 dB and extrapolating to 40 dB TTS is conservative in that there are a few examples of measured TTS levels of ~40 dB that have not resulted in PTS. There is a large variation in the TTS-PTS offset levels, especially as the range of one animal (13 to 37 dB) is greater than the ranges of all of the other animals listed in Table 6.

> ***Response (NOAA & Navy)***: The Reviewer's indication that there are few examples of studies where a 40-dB threshold shift has not resulted in PTS is referring to human and terrestrial mammal data (e.g., Ward 1960; Miller 1974; Hamernik et al. 1988). For marine mammals, noise-induced thresholds shifts of 40 dB or higher have resulted in recovery and were not permanent (Kastelein et al. 2013a; Popov et al. 2013; Popov et al. 2014). The individuals in these studies were exposed at SEL_{cum} at levels above some of the updated PTS onset levels (i.e., even though exposed at level which would exceed the updated PTS onset acoustic threshold level, the animal still recovered completely).
>
> The large variation in TTS-PTS offset levels is for the bottlenose dolphin BLU (Finneran et al. 2010; Finneran and Schlundt 2013). BLU is only one of two individual where multiple offset levels are available. The harbor porpoise Jerry is the other individual where two TTS-PTS offset levels are available (16 and 28 dB; Kastelein et al. 2014a, b). This variation illustrates that TTS-PTS offset levels depend on growth rates, which are frequency dependent, and typically are higher for frequencies where hearing is more sensitive (Finneran and Schlundt 2010; Finneran and Schlundt 2013). With both BLU and Jerry, TTS-PTS offsets below 20 dB were attributed to exposures possibly exceeding the critical level (i.e., where damage switches from being primarily metabolic to more mechanical) at this frequency or exposure scenario and are not considered appropriate to use as representative offsets. As more data become available, TTS-PTS offset levels can be re-examined and modified, if necessary.

8.3. Impulsive exposures

REVIEWER 2

Comment 47: See comments in Section 11.

Response: Thank you. Comments will be addressed in Section 11.

9. TTS EXPOSURE FUNCTIONS FOR SONARS

9.1. Low- and high-frequency components (a,b)

REVIEWER 2

Comment 48: There are errors in the calculations and rounding off of the "*a*" values and also with the "*C*" values in Table 8. The lowest *So* values from Table 5 were used to recalculate the "*a*" values for Table 8, with the exception that the LF value of 30 dB/decade was used. In Table 8, some "*a*" values were rounded up and some were rounded down. At least 2 significant digits should be presented for the "*a*" values.
The following recalculations of Equation 1 were performed using the revised "*a*" values below. The weighting functions values were recalculated by using the other Table 8 data in equation 1, but without the "*C*" component value (highest Eq. 1 value without "*C*") in the following table.

Group	Lowest *So*	*a* = *So*/20	Original "*a*" value in Table 8	Highest Eq. 1 value without "*C*" (dB)	Frequency of highest value (kHz)	Revised "*C*" value (dB)	Original "*C*" value (dB)
LF	30	1.50	1.5	-0.03	2.5	0.03	0.53
MF	24	1.20	1.2	-0.35	40	0.35	1.6
HF	35	1.75	1.8	-0.13	50	0.13	1.7
SI	37	1.85	1.8	-0.23	10	0.23	2.3
OW	26	1.30	2.0	-0.04	6.3	0.04	0.5
PW	11	0.55	0.8	-0.42	16	0.42	0.4
TU	28	1.40	1.4	-0.35	0.2	0.35	2.0
OA	28	1.40	1.4	-0.17	8	0.17	1.3
PA	41	2.05	2.0	-0.05	2.5	0.05	1.4

The revised "*C*" values, when applied to equation 1, will result in the highest weighting function value being 0 dB (line 835) at the frequency indicated in the above table. The values are all low and some may reach 0 at frequencies between the 1/3 octave center values used in these calculations. Where the "*a*" values were revised, the only significant changes are that both OW and PW would have higher exposures below 0.5 kHz.

> ***Response (NOAA & Navy):*** Numbers were rounded using unbiased rounding rules: (1) round up when the fractional part is > 0.5 (e.g., 1.46 rounds to 1.5), (2) round down when the fractional part is < 0.5 (e.g., 1.44 rounds to 1.4), (3) if the fractional part is exactly equal to 0.5, round to the nearest *even* digit; e.g., 1.45 rounds to 1.4, 1.55 rounds to 1.6, 1.65 rounds to 1.6, etc.). This methodology is consistent with the IEEE Standard 754.

The reviewer's calculations for the PW and OW groups are incorrect. The reviewer used the lower of the two s_0 values from the normalized and non-normalized composite audiograms. However, since the normalized composite audiograms were used to define f_1 and f_2, only the values of s_0 for the normalized composite audiograms are relevant. For the PW group, the correct value for s_0 is 15, $a = 15/20 = 0.75$ which rounds to 0.8. For the OW group, s_0 is 39, $a = 39/20 = 1.95$ which rounds to 2.0.

9.2. Frequency cutoffs (f_1 and f_2)

REVIEWER 2

Comment 49: The cutoff frequencies essentially reflect the region of best hearing, which is often defined as the frequencies where the audiogram is <20 dB above the lowest frequency. This value is arbitrary however. The procedures used here, the matching of the frequency and the SEL that causes 6 dB of TTS to the variously shaped curves (Figure 15), seem to be valid in that the final outcome is meant to be a measure of the frequency-based exposure function for TTS 6 dB levels.

> ***Response***: NOAA agrees that the Navy's procedure is more rigorous than previously used methods[13] to define the frequency range of best hearing.

9.3. Gain parameters K and C

REVIEWER 2

Comment 50: The gain parameters "*K*" are derived by raising the exposure functions up to a level where they will match the TTS 6 dB SEL values at the respective frequencies. For the MF group, there are a number of data points that fall on the line, less so for the HF group (Figures 18 and 19). For the PW group, (Figure 20) the situation is not as clear. The R^2 values for PW are very low (0.229). Further, in Kastelein et al. (2012) Table II, Seal 01 exhibited >6 dB TTS after 1-4 minutes post exposure with SELs of 174.6 and 183.6. Seal 02 exhibited >6 dB TTS after 12-16 minutes post exposure with SELs of 177.6 and 183.6. The two lowest values are not depicted in Figure 20 and if the TTS 6 dB was assumed to occur at 175 dB SEL, the *K* value for PW would drop to 115 dB and the estimated difference (median) for LF and SI would drop to 120.5 dB.

> ***Response (NOAA & Navy)***: NOAA recognizes that the R^2 values associated with weighting functions and TTS onset for phocid pinnipeds are low. The R^2 value for the PW group is low because the only data exist over a region where the curve is nearly flat. When R^2 equals 0.0, the best-fit curve fits the data no better than a horizontal line going through the mean of all *y*-values. This is similar to the case for the PW group data, where the fitting function changes little from 2.5 to 4 kHz and the TTS data points differ by only ~3 dB. This

---

[13] Note: The Reviewer's reference to the arbitrary nature associated with the definition of best hearing range does not reflect any past or current procedure incorporated by NOAA (i.e., NOAA has never used 20 dB above the lowest hearing threshold to define the region of best hearing).

results in a low R^2 even though the function actually fits the data very well: The deviations between the Phase 3 TTS exposure function and the actual TTS values are only ±1.3 dB.

TTS onset values are determined by plotting TTS amounts as a function of SEL, fitting Eq. (4) to the data, then interpolating to find the SEL at which the function predicts 6 dB of TTS. This is illustrated in Fig. 13(c). The resulting TTS onset values are 180 dB SEL for seal 01 and 183 dB SEL for seal 02. The process is based on fitting a function to all of the data, not just the individual data points with TTS $\geq$ 6 dB, and results in a single TTS onset value for each individual animal.

Comment 51: The concept of matching the exposure functions to available TTS data is valuable in that it will easily permit lowering the K values as new data become available.

Response: Thanks for your comment. NOAA agrees.

Comment 52: The data from Table 6 for PA has the TTS 6 dB level at 134 dB SEL, which matches the value given in Table 8 but the values for OA are 159 and 157 dB SEL respectively. Similarly, the TTS Onset level for HF in Table 7 is 151 dB SEL while the value in Table 8 is 149 dB SEL. The differences are not explained.

Response (NOAA & Navy): Tables 6 and 8 in the Navy Technical Report do not show the same parameters. Table 6 shows the TTS onset values from the various data sets. Table 8 shows the minimum value of the TTS exposure function that best fits the TTS data points. The values would only agree if there was a single TTS data point that occurred at or very close to the frequency at which the TTS exposure function reached a minimum value. These conditions occur for the PA group but not for the others (i.e., for OA the exposure function minimum is at a higher frequency than the frequency of the TTS data, therefore K < the TTS onset value (157 < 159).

Table 7 indicates the TTS onset value at f_0 — the frequency at which the hearing threshold (from the composite audiogram) is lowest. This is not necessarily the same frequency at which the exposure function reaches a minimum (where K is defined). For examples, compare the TTS exposure functions and composite audiograms in Fig. 17.

Comment 53: In Table 6 there are a number of values in the TTS Onset (dB SEL) column that are marked with an asterisk indicating that they were not used in subsequent calculations but no explanations about why these values were not considered are given.

Response: The asterisks meant to direct the reader to see the Notes column of the Table. The table footnotes have been expanded to list the specific reasons.

10. PTS EXPOSURE FUNCTIONS FOR SONARS

REVIEWER 2

Comment 54: The PTS threshold is set at 20 dB above the TTS 6 dB SEL values. PTS onset must refer to the SEL level at which a small PTS could occur and not a PTS of 6 dB. Changing the term "TTS onset" (as mentioned above) would help clarify the terminology here. The 20 dB SEL increase above the SEL for TTS 6 dB as a possible PTS damage level seems to be conservative, given that greater amounts of SEL have been experienced by a number of animals from different species groups without PTS having been detected.

> ***Response***: The Reviewer's assumption is correct that PTS does not refer to a PTS of 6 dB. NOAA provides a definition of PTS in the Acoustic Guidance's Glossary. As mentioned by the Reviewer, NOAA has invoked several conservative assumptions into the derivation of our PTS onset thresholds to account for uncertainty (e.g., defining TTS onset as the level just above where individual variability in hearing occurs, not accounting for exposures where TTS onset did not occur, using conservative growth rates). In response to the Reviewer's previous comment, NOAA considers the terminology of "TTS onset" preferable and more intuitive to readers of the Acoustic Guidance compared to using "TTS 6 dB."

11. TTS/PTS EXPOSURE FUNCTIONS FOR EXPLOSIVES

REVIEWER 2

Comment 55: I agree with the concept that using weighted SEL thresholds in conjunction with an unweighted peak SPL threshold is a conservative approach (lines 904-906). The TTS 6 dB threshold calculations follow on from those used for the steady state sounds and the few peak SPL data are used to support those levels. The levels of peak SPL associated with the proposed limits for TTS (6 dB) and PTS thresholds given in Table 9 seem rather high however.

> ***Response***: The marine mammal peak pressure TTS onset peak pressure thresholds in Table 9 are derived directly from a harbor porpoise, representing high-frequency (HF) cetaceans, exposed to a single airgun (Lucke et al. 2009) or a beluga, representing mid-frequency (MF) cetaceans[14], exposed to a watergun (Finneran et al. 2002). A 6-dB difference between TTS and PTS onset was used based on recommendations from Southall et al. 2007 and terrestrial mammal growth rate data, since PTS has not been directly studied in marine mammals.

Comment 56: With respect to humans listening in air, Clifford and Rogers (2009) report that impulse noise damage is more extensive than that produced by continuous noise with the same acoustic energy. They cite a source stating that "the Equal Energy Hypothesis does not apply to impulse noise, although it may apply to continuous noise." They also state that "The cutoff for mechanical versus biomolecular damage appears to be somewhere around 125 dB, at which an

---

[14] The peak pressure threshold for MF cetaceans is also applied to low-frequency (LF) cetaceans and pinniped (both otariids and phocids) functional hearing groups, since there are no data currently available for these groups in this metric.

abrupt increase in permanent threshold shift and hair cell loss is seen." This suggests that hearing damage from impulses is more damaging than that from continuous noise with the same SEL.

> ***Response***: We agree that impulsive sounds are more injurious than non-impulsive sources, which is why we have divided sources into these two broad categories within the NOAA Acoustic Guidance, with impulsive sources having lower PTS/TTS onset thresholds (expressed as cumulative sound exposure level) compared to non-impulsive sources. As mentioned by the Reviewer, the risk of damage from these impulsive sounds does not only depend on the duration of exposure (e.g., concept of "critical level," where damage switches from being primarily metabolic to more mechanical). Thus, $\mathrm{SEL}_{\mathrm{cum}}$ is not an appropriate metric to capture these effects (i.e., often violates the Equal Energy Hypothesis; NIOSH 1998), which is why instantaneous peak sound pressure level has also been chosen as part of NOAA's dual metric acoustic threshold levels.

Comment 57: Stark et al. (2003) cite various regulatory sources that place the maximum limit of exposure to humans to 135 dB (A-weighted peak) or 140 dB (C-weighted peak). Assuming a similar upper limit of 125 or 140 dB peak, relative to the lowest threshold in air or water for the 9 hearing groups, the data in the following table show that the TTS 6 dB peak SPL pressures in Table 9 all exceed the threshold plus 140 dB peak levels. The dynamic range of the PTS peak SPL thresholds of Table 9 (far right column below) range from 157 to 177 dB, all well above 140 dB. A conservative approach would be to use the threshold plus 140 dB peak as the threshold for the onset of impulse PTS until further data on humans and terrestrial mammals becomes available.

Group	Threshold dB re 1 µPa or 20 µPa	T + 125 dB peak SPL	Table 9. TTS 6 dB peak SPL	T + 140 dB peak SPL	Table 9. PTS dB peak SPL	Table 9. Dynamic range (dB)
LF	65	190	224	205	230	165
MF	53	178	224	193	230	177
HF	45	170	196	185	202	157
SI	61	186	224	201	230	169
OW	67	192	224	207	230	163
PW	62	187	224	202	230	170
TU	65	190	224	205	230	165
OA	12	137	163	152	169	157
PA	0	125	163	140	169	169

> ***Response***: For peak pressure acoustic threshold levels, MF cetaceans were used as surrogates for all cetaceans, as well as pinnipeds, to ensure that the threshold level did not unrealistically exceed the cavitation threshold of water. The peak pressure thresholds for MF cetaceans were obtained from direct measurements, which seem to support a larger dynamic range than what has been observed in terrestrial mammals, and NOAA believes are proposed thresholds are not unrealistic. As direct pinniped data becomes available, NOAA can re-evaluate these acoustic threshold levels.

Literature Cited (as referenced within Peer Review Report)

ANSI (American National Standards Institute). 1995. Bioacoustical Terminology (ANSI S3.20-1995).New York: Acoustical Society of America.

ANSI (American National Standards Institute). 2013. Acoustic Terminology (ANSI S1.1-2013). New York: Acoustical Society of America.

Clifford, R.E. and R.A. Rogers. 2009. Impulse noise: theoretical solutions to the quandary of cochlear protection. Annals of Otology, Rhinology and Laryngology, 118:417-427.

Cranford, T.W., and P. Krysl. 2015. Fin whale sound reception mechanisms: Skull vibration enables low-frequency hearing. PLOS ONE 10(3): e0122298.

EPA (Environmental Protection Agency).1982. Guidelines for Noise Impact Analysis (EPA Report Number 550/9-82-105).Washington, D.C.: Office of Noise Abatement and Control.

Finneran, J.J. 2015. U.S. Navy Technical Report: Auditory Weighting Functions and TTS/PTS Exposure Functions for TAP Phase 3 Acoustic Effects Analyses (February 2015). San Diego, California: SPAWAR Systems Center Pacific (SSC PAC).

Finneran, J.J. and A.K. Jenkins.2012. Criteria and thresholds for U.S. Navy acoustic and explosive effects analysis. San Diego, California: SPAWAR Systems Center Pacific.

Finneran, J.J., and C.E. Schlundt. 2010. Frequency-dependent and longitudinal changes in noise-induced hearing loss in a bottlenose dolphin (*Tursiops truncatus*). Journal of the Acoustical Society of America 128:567-570.

Finneran, J.J., and C.E. Schlundt. 2011. Subjective loudness level measurements and equal loudness contours in a bottlenose dolphin (*Tursiops truncatus*). Journal of the Acoustical Society of America 130:3124-3136.

Finneran, J.J., and C.E. Schlundt. 2013. Effects of fatiguing tone frequency on temporary threshold shift in bottlenose dolphins (*Tursiops truncatus*). Journal of the Acoustical Society of America 133:1819-1826.

Finneran, J.J., C.E. Schlundt, R. Dear, D.A. Carder, and S.H. Ridgway. 2002. Temporary shift in masked hearing thresholds in odontocetes after exposure to single underwater impulses from a seismic watergun. Journal of the Acoustical Society of America 111:2929-2940.

Finneran, J.J., D.A. Carder, C.E. Schlundt, and R.L. Dear. 2010. Growth and recovery of temporary threshold shift at 3 kHz in bottlenose dolphins: Experimental data and mathematical models. Journal of the Acoustical Society of America 127:3256-3266.

Finneran, J.J., Mulsow, J., and Schlundt, C.E. 2013. Auditory weighting functions in sea lions and dolphins. Presentation at the 3rd International Conference on The Effects of Noise on Aquatic Life, Budapest, Hungary.

Finneran, J.J., C.E. Schlundt, B.K. Branstetter, J.S. Trickey, V. Bowman, and K. Jenkins. 2015. Effects of multiple impulses from as seismic airgun on bottlenose dolphin hearing and behavior. Journal of the Acoustical Society of America 137:1634-1646.

Fletcher, H., and W.A. Munson. 1933. Loudness, its definition, measurement and calculation. Journal of the Acoustical Society of America 5:82-108.

Ghoul, A., and Reichmuth, C. 2014. Hearing in the sea otter (*Enhydra lutris*): Auditory profiles for an amphibious marine carnivore. Journal of Comparative Physiology A 200:967-981.

Hamernik, R.P., W.A. Ahroon, and J.A. Patterson. 1988. Threshold recovery functions following impulse noise trauma. Journal of the Acoustical Society of America 84:941-950.

Henderson, D., and R.P. Hamernik.1982. Asymptotic threshold shift from impulse noise. Pages 265-298 in Hamernik, R.P., D. Henderson, and R. Salvi, eds. New Perspectives on Noise-Induced Hearing Loss. New York: Raven Press.

Henderson, D., and R.P. Hamernik. 1986. Impulse noise: Critical review. Journal of the Acoustical Society of America 80:569-584.

Henderson, D., B. Hu, and E. Bielefeld.2008. Patterns and mechanisms of noise-induced cochlear pathology. Pages 195-217 in Schacht, J., A.N. Popper, and R.R Fay, eds. Auditory Trauma, Protection, and Repair. New York: Springer.

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

Kastak, D., B.L. Southall, R.J. Schusterman, and C. Reichmuth Kastak.2005. Underwater temporary threshold shift in pinnipeds: Effects of noise level and duration. Journal of the Acoustical Society of America 118:3154-3163.

Kastak, D., C. Reichmuth, M.M. Holt, J. Mulsow, B.L. Southall, and R.J. Schusterman. 2007. Onset, growth, and recovery of in-air temporary threshold shift in a California sea lion (*Zalophus californianus*). Journal of the Acoustical Society of America, 122:2916- 2924.

Kastelein, R.A., R. Gransier, L. Hoek, A. Macleod, and J.M. Terhune. 2012a. Hearing threshold shifts and recovery in harbor seals (*Phoca vitulina*) after octave-band noise exposure at 4 kHz. Journal of the Acoustical Society of America 132:2745-2761.

Kastelein, R.A., R. Gransier, L. Hoek, and J. Olthuis. 2012b. Temporary hearing threshold shifts and recovery in a harbor porpoise (*Phocoena phocoena*) after octave-band noise at 4 kHz. Journal of the Acoustical Society of America 132:3525-3537.

Kastelein, R.A., R. Gransier, and L. Hoek. 2013a. Comparative temporary threshold shifts in a harbor porpoise and harbor seal, and severe shift in a seal (L). Journal of the Acoustical Society of America 134:13-16.

Kastelein, R.A. R. Gransier, L. Hoek, and M. Rambags. 2013b. Hearing frequency thresholds of harbor porpoise (*Phocoena phocoena*) temporarily affected by a continuous 1.5 kHz tone. Journal of the Acoustical Society of America 134:2286-2292.

Kastelein, R.A., L. Hoek, R. Gransier, M. Rambags, and N. Claeys. 2014a. Effects of level, duration, and inter-pulse interval of 1-2 kHz sonar signal exposures on harbor porpoise hearing. Journal of the Acoustical Society of America 136:412-422.

Kastelein, R.A., J. Schop, R. Gransier, and L. Hoek. 2014b. Frequency of greatest temporary hearing threshold shift in harbor porpoise (*Phocoena phocoena*) depends on the noise level. Journal of the Acoustical Society of America 136:1410-1418.

Kastelein, R.A., R. Gransier, M.A.T. Marijt, and L. Hoek. 2015. Hearing frequency thresholds of harbor porpoises (*Phocoena phocoena*) temporarily affected by play back offshore pile driving sounds. Journal of the Acoustical Society of America 137:556-564.

Ketten, D.R. 2014. Expert evidence: Chatham Rock Phosphate Ltd Application for Marine Consent. New Zealand: Environmental Protection Agency. Available at: http://www.epa.govt.nz/EEZ/EEZ000006/EEZ000006_13_04_PowerPoint_Ketten.pdf.

Ketten, D.R., and D.C. Mountain. 2014. Inner ear frequency maps: First stage audiograms of low to infrasonic hearing in mysticetes. Presentation at ESOMM 2014, Amsterdam, Netherlands.

Kujawa, S.G., and M.C. Liberman. 2006. Acceleration of age-related hearing loss by early exposure: Evidence of a misspent youth. The Journal of Neuroscience 26:2115-2123.

Kujawa, S.G., and M.C. Liberman.2009. Adding insult to injury: Cochlear nerve degeneration after "temporary" noise-induced hearing loss. The Journal of Neuroscience 29:14077-14085.

Levine, S., P. Hofstetter, X.Y. Zheng, and D. Henderson. 1998. Duration and peak level as co-factors in hearing loss from exposure to impact noise. Scandinavian Audiology Supplementum 48:27-36.

Liberman, M.C. 2013. New perspectives on noise damage. Presentation at the 3rd International Conference on The Effects of Noise on Aquatic Life, Budapest, Hungary.

Lucke, K., U. Siebert, P.A. Lepper, and M-A.Blanchet.2009. Temporary shift in masked hearing thresholds in a harbor porpoise (*Phocoena phocoena*) after exposure to seismic airgun stimuli. Journal of the Acoustical Society of America 125:4060-4070.

Miller, J.D. 1974. Effects of noise on people. Journal of the Acoustical Society of America 56:729 764.

Nachtigall, P.E., A. Ya. Supin, J.L. Pawloski, and W.W.L. Au. 2004. Temporary threshold shifts after noise exposure in the bottlenose dolphin (*Tursiops truncatus*) measured using auditory evoked potentials. Marine Mammal Science 20:673-687.

NIOSH (National Institute for Occupational Safety and Health). 1998. Criteria for a recommended standard: Occupational noise exposure. Cincinnati, Ohio: United States Department of Health and Human Services.

Nordmann, A.S., B.A. Bohne, and G.W. Harding. 2000. Histopathological differences between temporary and permanent threshold shift. Hearing Research 139-13-30.

Plingst, B.E., R. Hienz, J. Kimm, and J. Miller. 1975. Reaction-time procedure for measurement of hearing. I. Suprathreshold functions. Journal of the Acoustical Society of America 57:421-430.

Popper, A.N., A.D. Hawkins, R.R. Fay, D.A. Mann, S. Bartol, T.J. Carlson, S. Coombs, W.T. Ellison, R.L. Gentry, M.B. Halvorsen, S. Løkkeborg, P.H. Rogers, B.L. Southall, D.G. Zeddies, and W.N. Tavolga. 2014. Sound Exposure Guidelines for Fishes and Sea Turtles: A Technical Report prepared by ANSI-Accredited Standards Committee S3/SC1. New York: Springer.

Popov, V.V., A. Ya Supin, V. V Rozhnov, D.I. Nechaev, E.V. Sysuyeva, V.O. Klishin, M.G. Pletenko, and M.B. Tarakanov. 2013. Hearing threshold shifts and recovery after noise exposure in beluga whales, *Delphinapterus leucas*. The Journal of Experimental Biology 216:1587-1596.

Popov, V.V., A.Ya Supin, V.V. Rozhnov, D.I. Nechaev, and E.V. Sysueva. 2014. The limits of applicability of the sound exposure level (SEL) metric to temporal threshold shifts (TTS) in beluga whales, *Dephinapterus leucas*. The Journal of Experimental Biology 217:1804-1810.

Price, G.R., and S. Wansack.1989. Hazard from intense midrange impulses. Journal of the Acoustical Society of America 86:2185-2191.

Reichmuth, C. 2013. Equal loudness contours and possible weighting functions for pinnipeds. Journal of the Acoustical Society of America 134:4210.

Robinson, D.W., and R.S. Dadon. 1956. A re-determination of the equal-loudness relations for pure tones. British Journal of Applied Physics 7:166-181.

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

Sills, J.M., B.L. Southall, and C. Reichmuth. 2014. Amphibious hearing in spotted seals (*Phoca largha*): Underwater audiograms, aerial audiograms and critical ratio measurements. The Journal of Experimental Biology 217:726-734.

Southall, B.L., A.E. Bowles, W.T. Ellison, J.J. Finneran, R.L. Gentry, C.R. Greene, Jr., D. Kastak, D.R. Ketten, J.H. Miller, P.E. Nachtigall, W.J. Richardson, J.A. Thomas, and P.L. Tyack. 2007. Marine mammal noise exposure criteria: Initial scientific recommendations. Aquatic Mammals 33:411-521.

Starck, J., E. Toppila, and I. Pyykko. 2003. Impulse noise and risk criteria. Noise and Health, 5:63-73.

Stebbins, W.C. 1966. Auditory reaction time and the derivation of equal loudness contours for the monkey. Journal of the Experimental Analysis of Behavior 9:135-142.

Suzuki, H., and Y. Takeshima. 2004. Equal-loudness-level contours for pure tones. Journal of the Acoustical Society of America 116: 918–933.

Ward, W.D. 1960. Recovery from high values of temporary threshold shift. Journal of the Acoustical Society of America 32:497-500.

Ward, W.D. 1997. Effects of high-intensity sound. Pages 1497-1507 In M.J. Crocker (ed.) Encyclopedia of Acoustics, Volume III. New York: John Wiley & Sons.

Wensveen, P.J., L.A.E. Huijser, L. Hoek, and R.A. Kastelein. 2014. Equal latency contours and auditory weighting functions for the harbour porpoise (*Phocoena phocoena*). The Journal of Experimental Biology 217:359-369.