

Killer whale (*Orcinus orca*) behavioral audiograms

Brian K. Branstetter,^{1,a)} Judy St. Leger,² Doug Acton,³ John Stewart,² Dorian Houser,¹ James J. Finneran,⁴ and Keith Jenkins⁴

¹National Marine Mammal Foundation, 2240 Shelter Island Drive, No. 200, San Diego, California 92106, USA

²Sea World San Diego, 500 Sea World Drive, San Diego, California 92109, USA

³Sea World San Antonio, 10500 Sea World Drive, San Antonio, Texas 78251, USA

⁴U.S. Navy Marine Mammal Program, Space and Naval Warfare Systems Center Pacific, Code 71510, 53560 Hull Street, San Diego, California 92152, USA

(Received 7 December 2016; revised 10 March 2017; accepted 10 March 2017; published online 5 April 2017)

Killer whales (*Orcinus orca*) are one of the most cosmopolitan marine mammal species with potential widespread exposure to anthropogenic noise impacts. Previous audiometric data on this species were from two adult females [Szymanski, Bain, Kiehl, Pennington, Wong, and Henry (1999). *J. Acoust. Soc. Am.* **108**, 1322–1326] and one sub-adult male [Hall and Johnson (1972). *J. Acoust. Soc. Am.* **51**, 515–517] with apparent high-frequency hearing loss. All three killer whales had best sensitivity between 15 and 20 kHz, with thresholds lower than any odontocete tested to date, suggesting this species might be particularly sensitive to acoustic disturbance. The current study reports the behavioral audiograms of eight killer whales at two different facilities. Hearing sensitivity was measured from 100 Hz to 160 kHz in killer whales ranging in age from 12 to 52 year. Previously measured low thresholds at 20 kHz were not replicated in any individual. Hearing in the killer whales was generally similar to other delphinids, with lowest threshold (49 dB re 1 μ Pa) at approximately 34 kHz, good hearing (i.e., within 20 dB of best sensitivity) from 5 to 81 kHz, and low- and high-frequency hearing cutoffs (>100 dB re μ Pa) of 600 Hz and 114 kHz, respectively.

© 2017 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4979116>]

[ANP]

Pages: 2387–2398

I. INTRODUCTION

Audiometric data for a broad range of cetacean species are critical to understanding the potential effects of anthropogenic noise on marine mammals. One of the most basic measures of hearing is the audiogram, a plot of a listener's detection threshold as a function of frequency. Aside from a few exceptions (e.g., *Tursiops truncatus* and *Delphinapterus leucas*), audiogram data for odontocetes are sparse and for mysticetes, non-existent (Mooney *et al.*, 2012).

Killer whales (*Orcinus orca*) are an apex predator, the largest of the delphinid odontocetes, and have one of the most complex social structures in the animal kingdom (Mann *et al.*, 2000). Their distribution is one of the most widespread of any marine mammal with population densities highest in cold water and at high latitudes (Hoelzel *et al.*, 2002). Because of their cosmopolitan distribution, killer whales have the potential to be impacted by a wide range of anthropogenic noise sources including petroleum exploration, naval activity, construction, and shipping (Holt, 2008). Previously, knowledge about killer whale hearing was based upon audiograms collected from only three killer whales. The whales tested consisted of two adult females (Bain and Dahlheim, 1994; Szymanski *et al.*, 1999) and one sub-adult male (Hall and Johnson, 1972) with apparent high-frequency hearing loss in the male. All three killer whales had best sensitivity between 15 kHz and 20 kHz, with behavioral hearing

thresholds significantly lower than any odontocete tested to date (e.g., 30 dB re 1 μ Pa), suggesting this species might be more sensitive to acoustic disturbance than other species (Miller *et al.*, 2014).

The extremely low thresholds measured previously in killer whales require replication and validation. Furthermore, the limited data are inadequate to provide insight into individual and population-level variability (e.g., due to age) in hearing capabilities, similar to what has been demonstrated in the bottlenose dolphin (Houser and Finneran, 2006; Popov *et al.*, 2007) and beluga whales (Castellote *et al.*, 2014). Testing the hearing of additional killer whales is also of interest in comparative hearing. Killer whales are the largest animals (aquatic or terrestrial) in which a complete audiogram has been measured (Szymanski *et al.*, 1999). Although a general relationship between animal size and high-frequency hearing cutoff has been reported (Heffner and Heffner, 2008), echolocating odontocetes appear to be a special case in which the suborder has exceptionally good high-frequency hearing capabilities. In the current study, behavioral audiograms for eight killer whales were measured. A species composite audiogram was estimated including prior behavioral audiogram measurements of Szymanski *et al.* (1999).

II. METHODS

A. Participants

Eight killer whales (*Orcinus orca*) participated in the current study. Six of these animals were housed at Sea

^{a)}Electronic mail: brian.branstetter@nmmf.org

World San Diego (SWSD) and two were housed at Sea World San Antonio (SWSA). None of the animals had prior experience in psychoacoustic related tasks. Medical histories of the animals were not available. The animals were housed in a complex of interconnected pools. The study followed a protocol approved by the Animal Care Review Committee at Sea World.

B. Hearing test procedure

1. Psychophysical methods

The pool configurations at both facilities were nearly identical. Hearing tests were conducted in a small pool (length ~ 14.6 m, width ~ 7.6 m, depth ~ 2.0 m deep) located adjacent to two larger pools that were part of the killer whales' normal habitat (Fig. 1). The pools were separated by large, underwater gates that allowed movement of water and sound from pool to pool, but leaving the test animal temporarily isolated from conspecifics.

The psychophysical methods used for conducting the behavioral hearing tests with the killer whales were based on a standard method used for testing bottlenose dolphins (Schlundt *et al.*, 2007; Branstetter and Finneran, 2008; Branstetter *et al.*, 2013) and sea lions (Mulow *et al.*, 2012) using the Hearing Test Program software package (Finneran, 2003). A go, no-go response procedure, coupled to a one-down, one-up, adaptive staircase procedure (Levitt, 1971) estimated thresholds at the 50% correct level. Each hearing test was typically composed of 4–8 dives, where the animal would receive a hand signal which instructed it to swim and station on a hearing test apparatus that was attached to a large underwater gate (Fig. 1). Stationing consisted of

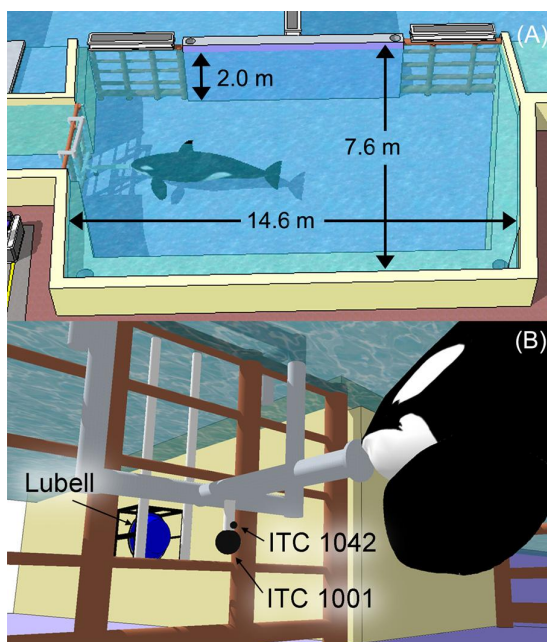


FIG. 1. (Color online) Hearing tests pool and apparatus. (A) Overhead view of a killer whale on the stationing device in the isolated small pool. (B) Underwater view of the killer whale on the stationing device with the three different sound projectors in view.

touching the tip of the animal's rostrum to an aluminum disk attached to the end of an aluminum pole on the hearing test station. The pole was designed to keep the animal stationary and at a fixed distance from the underwater sound projector. The number of trials per dive was randomized within the range of 1 to 30. During a single trial, the whale was trained to vocalize in response to a tonal sound (tone trial) or remain silent (catch trial) otherwise. The conditioned vocal response was a burst-pulse sound called a "raspberry" that is produced by expelling pressurized air from the animal's blowhole [Fig. 2(A)]. The raspberry sound is not part of the animals' normal vocal repertoire and is a very distinct, unambiguous sound to a human listener. The animal's vocal response was monitored via a speaker attached to a hydrophone placed behind the gate that the stationing device was attached to.

Each trial was 2.5 s in total duration and consisted of a 500 ms silent period, followed by a 500 ms stimulus period that contained either a 500 ms tone or a 500 ms silent "catch" interval. The animal was required to respond within a 2-s window after the onset of a tone. Any response outside of the response window was logged a false alarm. Tone trials and catch trials were presented in a pseudo-random order (Gellerman, 1933) where the catch trial percentage was 50%. If the animal correctly produced a vocal response to a tone, a "hit" was logged. If the animal failed to produce a vocal response to a tone, a "miss" was logged. If the animal produced a vocal response to a catch trial, a "false alarm" was logged. Animal trainers were blind to trial type (tone or catch trials) and were informed by an experimenter if the animal's response was correct or incorrect. Whistle feedback

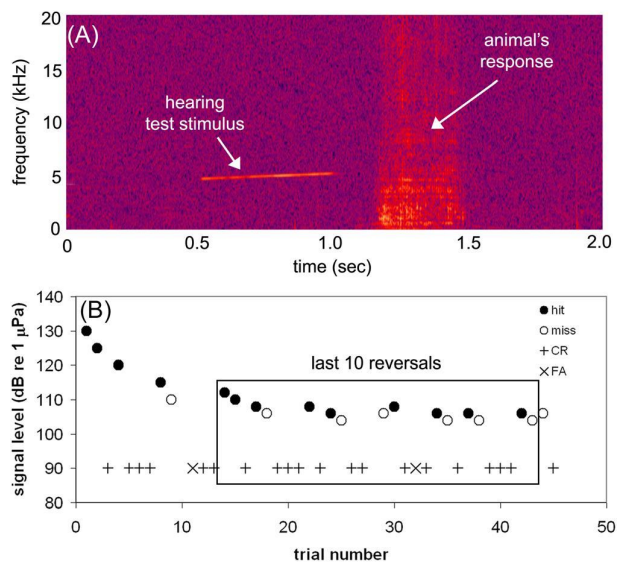


FIG. 2. (Color online) Spectrogram of a single trial and adaptive staircase example. (A) Spectrogram of a "tone" trial with a 5 kHz linear FM upsweep (bandwidth 10% of center frequency) followed by the animals "raspberry" vocal response. (B) Example staircase from a hearing test session. CR = correct rejection and FA = false alarms. Threshold (106 dB), false alarm rate (7.14%) and standard deviation (1.9 dB) were calculated from the last ten reversals (small inset box) from the staircase. CR and FA are plotted at the 90 dB level for visual purposes only, even though these were silent intervals. The signal for this particular staircase was a 1 kHz linear FM upsweep and the participant was animal G. The animal received a terminal bridge at the last CR.

(short whistle) was provided for correct responses by the trainer after each trial. Whistle feedback was withheld for incorrect responses. A longer duration whistle or a wall tap (i.e., tapping the wall with the trainer's hand or hard object) served as a terminal bridge that informed the animal to terminate the dive and return to the trainer's station to receive fish and social reinforcement. As mentioned above, the number of trials per dive was randomized between 1 and 30 and a "countdown" was displayed on the hearing test software. After the preselected number of trials for that dive was complete, a terminal bridge then occurred on the next correct trial. Terminal bridges never occurred on incorrect trials. This provided a variable-ratio schedule of reinforcement, which helped maintain the animal's motivation during the hearing tests.

In similar studies with bottlenose dolphins, one fish was rewarded to the dolphin for each correct tone trial and each correct catch trial accumulated during a dive (Branstetter and Finneran, 2008). However, killer whales consume much larger quantities of fish than dolphins, which made quantifying the exact amount of fish problematic and time consuming. Instead, "handfuls" of fish were typically rewarded to the animal after each dive, where the total quantity was "visually estimated" to be proportional to the animal's performance during the dive. An assumption was made that the animals understood the difference between "one handful" vs "five handfuls" of fish. In addition to "primary reinforcement" which was composed of food, the animals also received "secondary reinforcement" composed of non-food, appetitive stimuli (e.g., rubbing the animal's back). No feedback was given for false alarms (vocal response to a catch trial) or misses (remaining silent on a tone trial).

The sound pressure level (SPL, in dB re $1 \mu\text{Pa}$) of the first tone for each adaptive staircase began approximately 15–20 dB above the animal's estimated detection threshold (if known). The initial step size was 5 dB, followed by a 2-dB step size for the remainder of the trials after the first reversal (miss). A minimum of 11 reversals was obtained during each session and thresholds were calculated by averaging the last 10 reversals associated with the 2-dB step size [Fig. 2(B)]. Signal frequency was held constant during each staircase procedure. Each reported threshold was the average of three consecutive thresholds for that frequency, where the standard deviation for the three thresholds was 5 dB or lower. Thresholds for the same frequency were not collected on the same day for an individual animal, but were spaced out over the course of the experiment. This decreased the likelihood of averaged thresholds being potentially biased due to potential extraneous variables (e.g., lack of motivation or attention associated with a time-dependant event such as a whale in estrus in the next pool). During training, an animal's threshold will typically decrease due to learning. The 5 dB standard deviation criteria aided in determining when the effects of learning ceased, and the animal's best performance had been better approximated. All thresholds were associated with FA rates of 30% or less for each session.

2. Signal generation and calibration

Although data collection methods in San Diego and San Antonio were identical, there were slight differences in the equipment used between facilities due to equipment availability. In San Diego, a rugged notebook computer containing a National Instruments PCI-6251 multifunction data acquisition (DAQ) card (National Instruments, Austin, TX) was used. Outgoing analog tonal signals were filtered from 0.2 to 150 kHz (Krohn-Hite 3 C series) manually attenuated (RLC electronics, model AT-201-SR) and projected into the water column with an underwater sound projector. For frequencies between 80 and 160 kHz, an ITC 1042 (Santa Barbara, CA) spherical piezoelectric transducer was used. For frequencies between 10 and 56 kHz, an ITC 1001 spherical piezoelectric transducer was used. For frequencies below 10 kHz, the signal (after manual attenuation) was amplified [Peavey PVi 4B (Peavey, Meridian, MS)] and projected into the water column with a Lubell underwater speaker [Lubell LL916 (Lubell Labs Inc, Whitehall, OH)]. The ITC 1001 and ITC 1042 were connected to a PVC pole with Velcro and electrical tape. The PVC pole was then connected to the aluminum, stationing device with electrical tape (Fig. 1). The density mismatch between PVC and aluminum served to reduce potential sound conduction through the stationing apparatus to the animal's rostrum. Both the ITC 1001 and ITC 1042 were located approximately 84 cm from the tip of the whale's rostrum. The Lubell underwater speaker was mounted to a PVC stand which hung from the stainless steel, pool gate (Fig. 1), approximately 112 cm from the tip of the animal's rostrum. The aluminum stationing device also hung from the pool gate; however, foam padding was used to provide a "snug" fit, eliminate sound from the stationing device moving on the gate, and to acoustically isolate the stationing device from the gate.

In San Antonio, a Dell Inspiron laptop computer connected to a National Instruments USB-6251 DAQ device was used to generate and record signals. Outgoing signals were attenuated with a Hewlett Packard 350D manual attenuator (Palo Alto, CA). In San Antonio, the ITC 1042 was used for frequencies 10 kHz and above, while the Lubell LL916 2T (with Peavey PVi 4B amplifier) was used for all frequencies below 10 kHz.

Each tonal signal was 500 ms in duration with 10 ms linear onset and offset ramps to reduce spectral spread. The tonal signals were ascending linear FM sweeps, where the bandwidth was 10% of the center frequency. FM signals produce a more homogenous sound field in reflective pool environments than pure tones, while still resulting in comparable auditory thresholds (Finneran and Schlundt, 2007). The received SPLs (dB re $1 \mu\text{Pa}$) of the test signals were calibrated before (pre-calibrations) and after (post-calibrations) each hearing test session, with custom software (Finneran, 2003) to ensure received levels at the animal were consistent and accurate throughout each testing session. Received levels were measured with a Reson TC4013 hydrophone (Teledyne Reson Slangerup, Denmark), coupled to a Reson VP1000 preamplifier (Teledyne Reson, Slangerup, Denmark). Signals were digitized with the same DAQ devices used for sound projection.

If the difference between the pre- and post-calibrations were more than 6 dB SPL (absolute values), data from that session were rejected. If the difference between pre- and post-calibrations were greater than 3 dB, but less than 6 dB (absolute values), the threshold for that session was corrected by adding half the difference. For example, if a post-calibration value was 4 dB greater than the pre-calibration, 2 dB was added to the corresponding threshold level.

The frequency sensitivity about the head of small odontocetes is spatially complex (Brill *et al.*, 2001), but no data for spatial sensitivity of the killer whale head currently exists. All threshold levels reported are from calibrations, where the receiver was positioned at the location of the animal's rostrum tip (when the animal was not present). However, measurements were also made at all test frequencies at the location between the animal's auditory meatus when on station, but without the whale present. This measurement made an additional 57 cm from the rostrum tip was used to get an estimate of sound transmission loss. In San Diego, the mean difference in SPL between the rostrum tip and the auditory meatus was 1.8 dB (SD = 1.6 dB), while at SDSA the mean difference was 3.4 dB (SD = 2.9 dB).

During the experiment, frequencies were tested starting with 10 to 40 kHz in a pseudo random order. The lower frequencies (<10 kHz) and higher frequencies (>40 kHz) were tested last. Subjectively, animals appear to have a more difficult time with hearing tests toward the lower and higher end of their hearing ranges due to lowered sensitivity. Testing these frequencies last allowed the animals to accumulate experience to aid them in testing the more difficult frequencies. Frequency spacing was in 1-octave steps, unless there was a large difference between adjacent octave-spaced thresholds (subjectively determined by visually inspecting plotted audiograms). For example, if the average threshold for 80 kHz was significantly greater than the average threshold for 40 kHz, the half-octave between them (56 kHz) was also tested. At the upper and lower frequencies, if an animal failed to respond at the highest level that could be generated on three separate days (typically around 130 dB), an assumption was made that the animal could not hear the sound and testing was concluded for that frequency.

Steps were taken to mitigate pool noise during hearing tests. In addition to measuring the spectral density (dB re 1 $\mu\text{Pa}^2/\text{Hz}$) of the pool noise during calibration of the hearing tests tones, additional noise measurements were regularly conducted with a Reson TC4032 low-noise hydrophone (Teledyne Reson Slangerup, Denmark), coupled to a VP1000 pre amplifier (Teledyne Reson Slangerup, Denmark), using the same DAQ devices for the hearing tests. Ambient noise measurements were made by averaging 100 ms clips for a total of 15 s, at a sampling rate of 300 kHz. The primary source of pool noise (highest SPL and spectral density levels) was from vocalizations (non-whistle calls) of non-participant whales. Killer whales are highly soniferous, with call peak frequencies typically below 10 kHz (Singleton and Poulter, 1967; Miller *et al.*, 2007). As a result, for test frequencies below 10 kHz, all non-participant whales were moved to locations away from the test pool, and were engaged by animal trainers for the duration of testing. Whale vocalizations tended to decrease dramatically

(typically no vocalizations) if the animals were participating in "low-energy" training tasks. For frequencies above 10 kHz, animal trainers also engaged the non-participating whales if vocalizations were subjectively judged to be of sufficient level to interfere with the hearing tests.

III. RESULTS AND DISCUSSION

Table I presents average thresholds levels for each of the eight animals in the current study, as well as for the two animals (Vigga and Yakka) reported in Szymanski *et al.* (1999). Figure 3 plots thresholds for the ten killer whales with ambient noise measurements from SWSD and SWSA. A composite audiogram model, estimated using the procedure below, is also plotted.

The current study failed to replicate the extremely low behavioral thresholds from Szymanski *et al.* (1999). The behavioral methods used in Szymanski *et al.* (1999) were variable. A variety of different catch trials types were used, thresholds were based on four reversals associated with a 6–8 dB step size, and calibration methods were not well documented. The authors stated that the calibrations were "probably accurate within 6 dB." These factors could account for some of the variability in the measured behavioral audiograms from that study. The current study used standardized hearing test methods with a 2 dB step size associated with a ten reversal threshold, strict criteria for data inclusion, pool noise mitigation measures, and rigorous calibrations before and after each threshold testing session.

TABLE I. Threshold data for each killer whale (A–H) at each frequency, with mean, median, number of subject (N) and standard deviation (SD). Killer whales A and B were not included in mean, median, N, or SD calculations due to hearing loss (see composite audiograms). Data from the killer whales Vigga and Yakka are from Szymanski *et al.* (1999). All thresholds are in dB re 1 μPa .

kHz	A	B	C	D	E	F	G	H	Vigga	Yakka	Mean	Median	N	SD
0.1								123			123	123	1	
0.25				104				104			104	104	2	0
0.3			104								104	104	1	
0.5				103	100	106	110	96			103	103	5	6
1				106	110	91	99	107	100		102	103	6	7
2.5			75	71	81	85	83	70			78	78	6	6
4								59	62		61	61	2	2
5	102	95	65	56	65	80	80	65			68	65	6	9
8									55	60	58	58	2	4
10		82	63	58	55	53	53	60			57	57	6	4
12									50	40	45	45	2	7
16									44	48	46	46	2	3
20	67	57	53	47	51	50	54	50	38	30	46	50	8	8
32									55	36	46	46	2	13
40	57	48	53	59	49	47	53	59			53	53	6	5
45									54	42	48	48	2	8
56		53	52		48	45	72				54	50	4	12
60									63	41	52	52	2	16
80		58	64	75	58	97	79	74	69	60	72	72	8	13
100		71	82	112	83		79	91	77	74	85	82	7	13
113		116		135	119			94			116	119	3	21
120										108	108	108	1	
160									122		122	122	1	

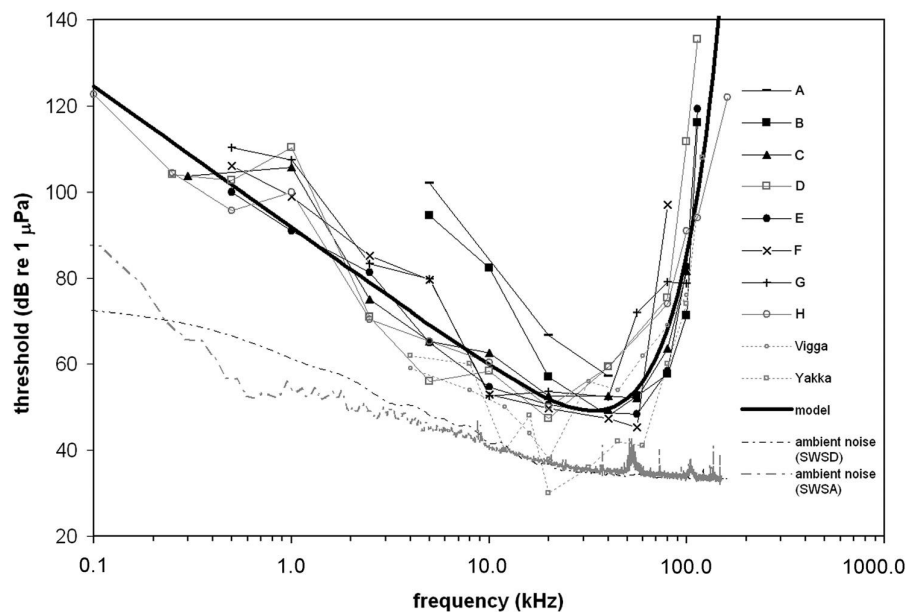


FIG. 3. Audiograms of individual killer whales. Solid black lines represent animals from SWSD and solid gray lines represent animals from SWSA, while dashed lines represent animals from Szymanski *et al.* (1999). The “model” data is a composite audiogram estimated from the procedure discussed. Average ambient noise values are in dB re $1 \mu\text{Pa}^2/\text{Hz}$. Ambient noise was measured with a TC4032 low-noise hydrophone coupled to a VP1000 preamplifier.

One hypothesis is that pool noise masked the 20 kHz tone in the current study, resulting in elevated thresholds relative to Yakka and Vigga from Szymanski *et al.* (1999). The measured pool noise at SWSD and SWSA, at 20 kHz was approximately $36 \text{ dB re } 1 \mu\text{Pa}^2/\text{Hz}$, which is the noise floor of the recording system, rather than an accurate measure of the pool noise. Critical ratios for killer whales have not been published. However, critical ratios for *T. truncatus* and *D. leucas* are approximately 30 dB (Lemonds *et al.*, 2011). The median threshold at 20 kHz from the current study (not including animals A and B due to hearing loss; see below) was 53 dB re $1 \mu\text{Pa}$ ($\text{SD} = 2.35 \text{ dB}$). If critical ratio predictions were applied, the ambient noise would have to be 23 dB re $1 \mu\text{Pa}^2/\text{Hz}$. It is doubtful that killer whales are sensitive to noise at such low spectral density levels. More complicated scenarios could be envisioned where less masking would occur [e.g., level-dependent auditory filters (Irimo and Patterson, 2001); or comodulation masking release (Branstetter and Finneran, 2008)]. However, without the ability to measure the ambient noise at extremely low levels, any masking scenario remains speculative. The lowest thresholds levels in the current study are consistent with the lowest thresholds from other odontocete behavioral audiograms (see Table IV). A more parsimonious conclusion is that the lowest thresholds from the current study reflect absolute sensitivity for killer whales.

Concrete pools can produce complex acoustic environments and are not ideally suited for hearing tests. High-resolution sound field measurements in the testing pool (i.e., SPL as a function of frequency and position within the volume of water the animal’s head would occupy) were not conducted in the current study or in the studies of Szymanski *et al.* (1999). Sound fields have been measured in other hearing tests pools (Southall *et al.*, 2000; Finneran and Schlundt, 2007; Kastelein *et al.*, 2008). An assumption is made that the measured sound field, sans animal, is representative of the sound field during a hearing test with the animal present,

thus validating measured hearing thresholds. Theoretically, confidence in this assumption is warranted when the ratio between the animal size and wavelength of the sound is small (e.g., small fish in low-frequency sound field). The presence of the animal would theoretically have a minimal interaction with the sound field. However when the animal size to wavelength ratio is high (e.g., for a massive killer whale, in a relatively high-frequency sound field) this assumption becomes untenable. Sound field measurements, with a killer whale on the stationing device, were not attempted due to concerns related to equipment damage and training time constraints. Killer whale heads are large (average circumference of test subjects’ heads = 234 cm, at the auditory meatus), and although no data exist on the hearing sensitivity about a killer whale’s head, evidence from other odontocetes (Brill *et al.*, 2001; Cranford *et al.*, 2008) strongly suggest killer whale hearing sensitivity will be frequency and spatially complex. The complex sound field produced by the pool interacting with a massive animal, combined with complex hearing sensitivity about the animals head may have had an effect on measured thresholds in the current study and Szymanski *et al.* (1999). Additional research is required to parse out the effects of these variables.

A. Composite audiogram estimation procedure

Estimating a composite audiogram from healthy listeners of a species can be useful for creating a representative audiogram for the species. One method of creating a composite audiogram is to simply take the mean or median threshold value from multiple individuals at each frequency tested. Median values are robust to outliers and are often preferable to mean values. An issue with using this approach is that different frequencies will often have different numbers of thresholds due to individual differences in hearing abilities (i.e., not all animals can hear the extreme lower and higher frequencies), or different studies measured different frequencies. As a result, some mean or median values at

specific frequencies may be heavily biased toward the hearing of a small number of individuals rather than being representative of the species.

Another technique, which is more robust to outlier data points, is to fit smooth functions to cumulative data. Castellotte *et al.* (2014) used a 4th order polynomial to fit a smooth function to beluga aggregate threshold data. The advantage of this procedure is the relative ease of application. However, inflection points inherent with higher order polynomials are present in the cumulative audiogram, which have no relation to any underlying auditory mechanisms. Popov *et al.* (2007) used piecewise regression to fit a linear function below frequencies of maximum sensitivity, and an exponential function to fit thresholds above frequencies of maximum sensitivity, using averaged threshold data. This procedure assumes there are two different mechanisms (one for lower frequencies and one for higher frequencies) affecting thresholds. Both procedures from Castellotte *et al.* and Popov *et al.* fit functions to aggregate data or central tendency data with an equal number of observations at each frequency. The data set in the current study, which included the data from Szymanski *et al.* (1999), did not have an equal number of thresholds at each frequency. As stated above, this can lead to potential biases when a frequency with a small number of observations (from only a few animals) contains potential outlier data. Finneran (2016) resolved this issue by using linear interpolation to estimate missing data points resulting in an equal number of observations at each frequency. The procedure also rejected data from animals with apparently anomalous hearing. Median values were calculated at each frequency to reduce the influence of outliers, and a five-parameter function, similar to Popov *et al.* (2007) was then fit to the median values.

A similar multi-step algorithm as Finneran (2016) was applied to the current data set. The difference between the current approach and the previous approaches is that smoothed functions were fit to each individual audiogram (rather than aggregate or central tendency data) using a four-parameter model:

$$th = b_1 e^{(b_2 x)} + b_3 x + b_4, \quad (1)$$

where b_1, \dots, b_4 are parameter coefficients, th is threshold in dB re 1 μ Pa, and x is the log-transformed frequency

$$x = \log_{10}(f), \quad (2)$$

where f is frequency in kHz.

Steps for the model-fitting algorithm were as follows:

- (1) Data from animals with anomalous hearing were excluded from the analysis. Anomalous hearing was defined as an animal having at least two or more thresholds where each threshold was 2 standard deviations or more from the mean threshold. Two animals met this criterion (animals A and B) and their data were not included in the composite audiogram.
- (2) The model [Eq. (1)] was fit (non-linear regression) to each individual audiogram (see Fig. 4).

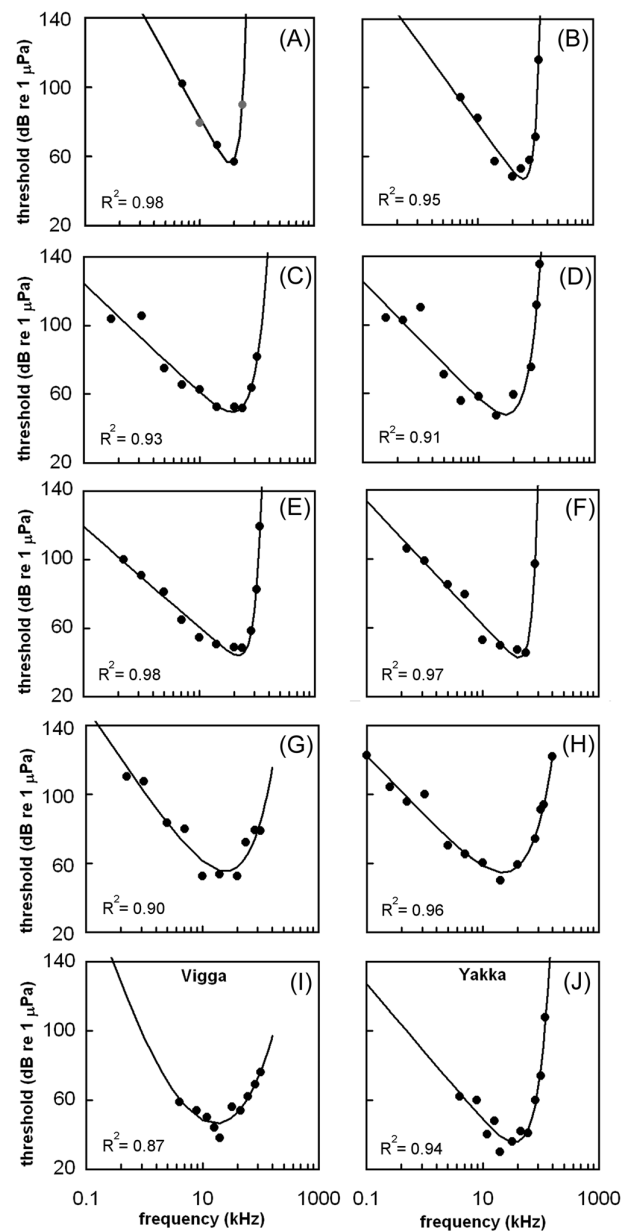


FIG. 4. Model fits to each individual audiogram using Eq. (1). Panel letters correspond to animal identification from Tables I and II. The gray data points for animal A did not meet the criteria for data threshold (i.e., $SD < 5$ dB). Data from animals A and B were not included in the composite audiogram estimation procedure because both had hearing loss. However, animal A's data were included for demographic purposes. Data for Vigga and Yakka are from Szymanski *et al.* (1999).

- (3) Median values were calculated at each frequency from all of the model fits.
- (4) A composite model [Eq. (1)] was then fit to the median values.

Each step in the algorithm is designed to remove outlier data. The fitting procedure made the assumption that “true” individual audiograms were smooth functions and that variability and outliers (i.e., peaks and valleys) in an individual's audiogram reflect abnormal hearing or

measurement error. The fitting procedure has the benefit of having an equal number of observations at each frequency, can extrapolate thresholds at frequencies above and below empirically measured frequencies, and is robust to outlier data.

The model provided good fits to each individual audiogram with R^2 values exceeding 0.90 for all audiograms except one (Fig. 4). A median audiogram was then fit from eight of the ten, modeled audiograms and is displayed in both Figs. 3 and 5. The parameter coefficients for the least squares fit for the median audiogram were $b_1 = 0.00927$, $b_2 = 4.38$, $b_3 = -32.7$, and $b_4 = 91.8$.

Figure 5 compares composite audiograms using a variety of different methods. For the mean audiogram (median values were almost identical), the values reflect biases associated with data points with a low number of observations at some of the frequencies. The mean audiogram from Fig. 5 erroneously places best sensitivity at 12 kHz, even though not a single animal had best sensitivity below 20 kHz. If the method from Popov *et al.* (2007) were applied to the current data, the composite audiogram would have also erroneously place best sensitivity at 12 kHz. This is an example of why an equal number of observations at each frequency is critical and how data from a few observation ($n = 2$) can bias conclusions. In the current example (Fig. 5) with a 4th order polynomial fit, conclusions about the low-frequency cutoff are difficult to make. The shape of a polynomial can also vary widely depending on the number of terms included. The current model is in better agreement with the individual audiogram data where the lowest thresholds for each animal from Table II result in a mean of 36.8 kHz, a median of 40 kHz, and a mode of 20 kHz.

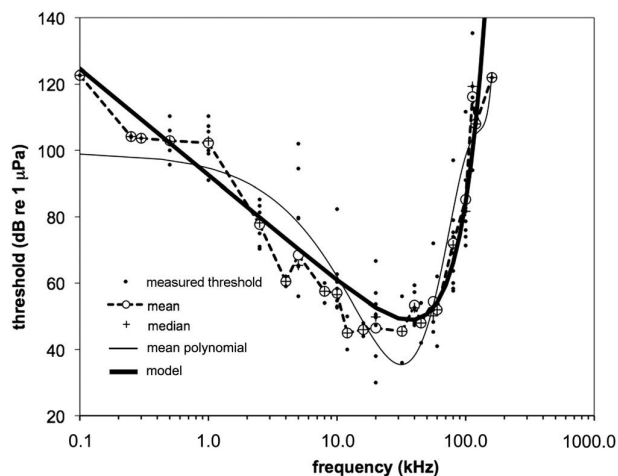


FIG. 5. Composite audiograms using three methods. Small closed circles are measured thresholds from the current study and Szymanski *et al.* (1999). Open circles and “pluses” are mean and median values, respectively. The thin line is a 4th order polynomial fit to mean values. The thick line represents model values from Eq. (1). Mean and median values are biased by an unequal number of thresholds at each frequency, erroneously concluding that best sensitivity is at 12 kHz. The shape of polynomials can vary considerably depending on the number of terms, and with the current data set an estimation of the low-frequency cutoff is difficult.

Audiogram data for each individual, were also fit using the same procedure above, however the equation from Finneran (2016) was used instead of Eq. (1):

$$th = T_o + A \log_{10} \left(1 + \frac{F_1}{f} \right) + \left(\frac{f}{F_2} \right)^B, \quad (3)$$

where th is the threshold for a given frequency f , and T_o , F_1 , F_2 , A , and B are fitting parameters. The inclusion of an additional parameter allows for a more shallow slope near the region of best sensitivity (Finneran, 2016). Parameter values for the least squares fit for the current data set were $T_o = 44.13$, $F_1 = 12.55$, $F_2 = 26.50$, $A = 40.58$, and $B = 2.74$. The mean R^2 values from each individual audiogram, using Eqs. (1) and (3) were 0.940 and 0.941, respectively. For the current data set, the more parsimonious model [Eq. (1)] is more appropriately justified, however the parameters for Eq. (3) are provided to facilitate inclusion into future auditory weighting functions (Finneran, 2016).

B. Hearing Loss

In the current study, three animals had elevated thresholds consistent with hearing loss (Fig. 3). Animal F, (age 26, female), had an elevated threshold at 80 kHz and did not respond to any tones above that frequency. Animal A (age 52, female) had broadband hearing loss consistent with a more advanced stage of presbycusis. Only the three data points that met the criteria [standard deviation (SD) < 5 dB for three consecutive replications] are reported for animal A in Table I and plotted in Fig. 3. Additional data were collected at 10 kHz (mean threshold = 79 dB SPL, SD = 18.2) and 56 kHz (90 dB SPL, SD = 8.4), but these did not meet the criteria for inclusion. Animal A did not respond to 80 kHz presented at 130 dB SPL. Presbycusis, or age-related hearing loss, consists of hearing loss to higher frequencies with progressive spreading to lower frequencies (Corso, 1959). Presbycusis has been previously reported in bottlenose dolphins (Houser and Finneran, 2006), a false killer whale (Kloepper *et al.*, 2010), and appears to affect killer whales as well. It is considered to be part of the normal aging process, is found in most mammals, and is not otherwise considered pathological.

The killer whale with hearing loss in Hall and Johnson (1972) was reported as a sub-adult male of unknown age that weighed 1820 kg, making him the smallest (and likely the youngest) killer whale tested to date. This animal had an upper frequency cutoff of only 30 kHz. Given the animal’s age, the source of this hearing loss is inconsistent with age-related hearing loss. However, the pattern is consistent with drug-induced hearing loss (ototoxicity), which normally affects the basal end of the cochlea, impairing high-frequency hearing. A similar pattern of high-frequency hearing loss was observed in a young (8–11 years old) beluga who received aminoglycoside antibiotics at 6 months of age (Finneran *et al.*, 2005). Medical records for the animal in the Hall and Johnson (1972) study were not available so the etiology of the hearing loss could not be determined.

TABLE II. Demographic data. For “test location,” SWSD and SWSA are Sea World San Diego and Sea World San Antonio, respectively. “Estimated age” is the estimated age at the time of testing. “Best sensitivity” is the frequency at which the lowest threshold occurred. “Lowest threshold” is the lowest threshold associated with the best sensitivity. “Low-freq cutoff” is the lowest frequency where the threshold is below 100 dB. “High-freq cutoff” is the highest frequency where the threshold is below 100 dB. Both low and high-frequency cutoffs were estimated from model data in Fig. 3. All other values are from direct measurements. “Model” values are estimates from the composite model.

Study ID	Test location	Estimated age	Sex	Weight (kg)	Best sensitivity (kHz)	Lowest threshold (dB)	Low-freq cutoff (kHz)	High-freq cutoff (kHz)
A	SWSD	52	f	3767	40	57	5.1	58.4
B	SWSD	12	m	3014	40	48	3.5	108.8
C	SWSD	22	m	3756	56	52	0.6	118.7
D	SWSA	24	m	4141	20	47	0.5	93.0
E	SWSD	13	m	2697	56	48	0.5	106.3
F	SWSD	26	f	2892	56	45	0.9	81.7
G	SWSD	22	f	2248	40	51	1.1	130.6
H	SWSA	16	m	3266	20	50	0.4	122.5
Vigga	Marine World	13	f	NA	20	38	0.9	169.3
Yaka	Marine World	27	f	NA	20	30	0.5	116.0
			mean	3222	36.8	46.6	1.4	110.5
			median	3140	40.0	48.0	0.8	112.4
			mode	NA	20.0	48.0	0.5	NA
			STD	633	15.9	7.7	1.6	29.8
			model		33.7	49	0.6	113.7

Animal B had low-frequency hearing loss with elevated thresholds for frequencies of 20 kHz and below. He had normal hearing for frequencies of 40 kHz and above. To the best of our knowledge, low-frequency hearing loss has not been reported in the scientific literature for any marine mammal. Low-frequency hearing loss, although rare, is found in humans and can be associated with different types of sensorineural hearing loss such as Meniere’s disease, Wolfram’s syndrome, Mondini dysplasia and bacterial or viral infections (Gulick *et al.*, 1989). These conditions were not clinically detected in this animal and thus the origin of the altered hearing remains unknown.

C. Demographics

Statistical tests (analysis of covariance) were conducted to examine the relationship of demographic predictors (sex, age, weight) and outcome variables related to thresholds (high-frequency cutoff, low-frequency cutoff, frequency with the lowest threshold, lowest threshold level). All data values were based on empirical measurements except the

low and high-frequency cutoffs, which were estimated from the individual model function in Fig. 4. Low and high-frequency cutoffs were defined as the lowest and highest frequencies where the audiograms intersected 100 dB re 1 μ Pa. Model estimates for these values were necessary because a few of the killer whales did not have low-frequency or high-frequency thresholds above 100 dB re 1 μ Pa (Fig. 3). There is a precedence for defining low and high-frequency cutoffs at 120 dB re 1 μ Pa to facilitate comparison of hearing ranges with terrestrial mammals, which has been arbitrarily defined at 60 dB re 20 μ Pa (Heffner and Heffner, 2008) (this essential equates thresholds based on sound intensity in the two media). However, threshold data for marine mammals at or above 120 dB (re 1 μ Pa) are relatively sparse. Instead, 100 dB was selected as the cutoff level, because more empirical threshold data exists at this level (especially at lower frequencies) for killer whales (and other odontocete species). Threshold levels at 100 dB rely less on model assumptions and model extrapolation than values at 120 dB. Threshold data from animal B were not included in the demographic analysis since his data were considered abnormal. Although

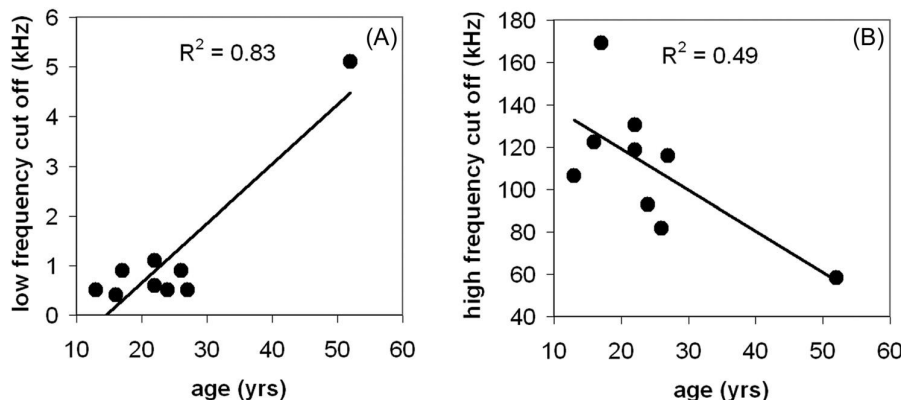


FIG. 6. Age as a predictor of low-frequency cutoff and high-frequency cutoff. The low and high-frequency cutoffs were arbitrarily defined as the lowest and highest frequency that had a threshold below 100 dB re 1 μ Pa, respectively. Although age was a significant predictor of low and high-frequency cutoffs, this result was primarily influenced by animal (A) that was 52 years of age.

TABLE III. References used to derive species composite behavioral audiograms.

<i>D. leucas.</i>	<i>P. phocoena</i>	<i>T. truncatus.</i>
(Finneran <i>et al.</i> , 2005)	(Kastelein <i>et al.</i> , 2002)	(Lemons, 1999)
(Johnson <i>et al.</i> , 1989)	(Kastelein <i>et al.</i> , 2010)	(Johnson, 1967)
(White <i>et al.</i> , 1978)	(Kastelein <i>et al.</i> , 2015)	(Finneran, 2010)
(Awbrey <i>et al.</i> , 1988)		(Schlundt <i>et al.</i> , 2007)
(Ridgway <i>et al.</i> , 2001)		

animals A and F also had hearing loss, presbycusis is considered to be a normal part of the aging process and their data (including the two data points at 10 kHz and 56 kHz that failed to meet data criteria) were included in the demographic analysis. Stepwise deletion (i.e., removing non-significant terms) was used to determine the most parsimonious statistical models.

Sex and weight were not significant predictors of thresholds. However, age was a significant predictor of low-frequency cutoff ($F_{1,7} = 34.69$, $p < 0.001$) and high-frequency cutoff ($F_{1,7} = 6.79$, $p = 0.035$). Caution must be exercised since this conclusion is driven primarily by two data points, from a single animal (i.e., animal A, age 52), where her threshold data points are statistical outliers within the present data set (Fig. 6). However, her threshold data points are consistent with normal, age-related hearing loss and may not be statistical outliers if additional threshold data from a broad range of whales with different ages were obtained. When data from animal A were removed from the analysis, no statistical tests were significant.

D. Comparative hearing

Within mammals, frequency of best sensitivity, and high-frequency cutoff, have been shown to correlate with body size (Heffner and Heffner, 2008). In theory, smaller animals hear higher frequencies better than larger animals due to the relationship between head size and sound wavelength. Small heads do not cast a significant acoustic shadow for lower frequencies, diminishing the utility of binaural spectral-difference cues for sound localization. Thus, a selective pressure for high-frequency hearing exists for small animals (Heffner and Heffner, 2008). Of course, size is also correlated with other acoustic-related variables such as the size of an animal's auditory and vocal anatomy (Ketten, 1992; Fitch, 1997; Ketten, 1998; May-Collado *et al.*, 2007). Smaller animals produce higher frequency sounds and thus, need to hear these sounds. Odontocetes are

relatively large animals but have, paradoxically, excellent high-frequency hearing due to the demands of echolocation. However, within odontocetes, the relationship between size and hearing range may exist. To test this, a subset of behavioral, audiometric data compiled from Finneran (2016) were selected from odontocetes, where there were at least two or more complete audiograms from different individuals for the species. Three species (in addition to killer whales) met this criterion (see Tables III and IV). Although a larger amount of AEP audiometric data were available, the current analysis was restricted to behavioral audiograms for consistency with the current behavioral study. AEP thresholds can, under certain conditions, depart considerably from behavioral thresholds depending on the stimulus type and frequencies being tested (Yuen *et al.*, 2005; Finneran and Houser, 2006). For the three species selected, composite audiograms were estimated using the method from Finneran (2016). Individual audiograms and composite audiograms are displayed in Fig. 7. From the composite audiograms, hearing metrics were calculated using the same criteria applied to the killer whales (i.e., best hearing range were the frequencies 20 dB above the lowest threshold and lower and upper cutoffs were the frequencies at 100 dB) and are displayed in Table IV. The average body masses for *O. orca* and *P. phocoena* were calculated from the animals that participated in the audiogram hearing studies. The average body mass values for *T. truncatus.* and *D. leucas.* were from Read *et al.* (1993) and National Marine Fisheries Service (2016), respectively. Body mass was highly correlated with best sensitivity [Fig. 8(A)], best hearing bandwidth [Fig. 8(C)] and upper frequency cutoff [Fig. 8(E)]. Lowest thresholds and low-frequency metrics were not well correlated with body mass. Each composite audiogram has a different shape causing the functions to converge and diverge at different levels, especially for the lower frequencies. However, the trend for the relationship between body mass and high-frequency cutoff remains consistent regardless of which dB level is arbitrarily chose as the cutoff, for SPLs up to 127 dB re 1 μ Pa [Fig. 7(D)]. Although echolocation may serve as a selective pressure for high-frequency hearing, auditory morphology that is scaled with body mass (e.g., ossicular mass) may place constraints on high frequency hearing (Hemilä *et al.*, 2001).

E. Social learning

Standard operant conditioning techniques were used for shaping the killer whales' behaviors for the hearing test

TABLE IV. Species average mass and metrics derived from species composite audiograms. Species mass for *O. orca* and *P. phocoena* were average weights from the subjects in those studies. Mass for *D. leucas.* and *T. truncatus.* were from Read *et al.* (1993) and National Marine Fisheries Service (2016). N is the number of subjects for each species composite audiogram.

Species	Mass (kg)	Best sensitivity (kHz)	Lowest threshold (dB)	Best hearing range (kHz)	Low-frequency cutoff (kHz)	High-frequency cutoff (kHz)	N
<i>O. orca</i>	3222	34	49	5–81	0.60	114	8
<i>D. leucas</i>	1500	54	52	5–103	0.75	120	8
<i>T. truncatus</i>	259	86	57	3–135	0.70	144	4
<i>P. phocoena</i>	31	118	47	3–150	0.45	160	3

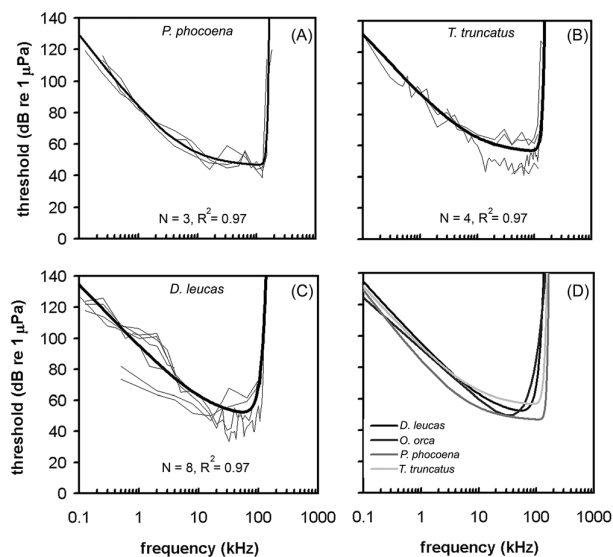


FIG. 7. Composite audiograms for four odontocete species. Composite audiograms were estimated using the procedure from Finneran (2016). Audiogram references are listed in Table III.

procedure (Skinner, 1951; Ramirez, 2012). For example, the whales received positive reinforcement (i.e., fish) for producing the raspberry sound in response to the test tones, and for remaining silent during catch trials. In addition, when a specific hand signal was provided, the animals were rewarded, for swimming and placing their rostrum on the stationing device. After tens to hundreds of repetitions, where the desired behaviors are reinforced, the whales' behaviors were successively approximated into the required behaviors for the hearing tests.

During data collection, the non-participating killer whales were moved away from the hearing test pool. However, during training, the other killer whales were free to move about in the adjacent pools. Often, many of the non-participating whales would congregate near the gate separating the pools [Fig. 1(A)], apparently observing the training sessions. Animals E and G were among the observing animals, and on conjecture from the training staff, a different training approach was used for these two animals. On their very first training session, the animals were given the hand signal to swim and station on the hearing test device "just to probe if the animals may have learned by observation." Both animals (in their respective, independent sessions) correctly swam and stationed on the hearing test stationing device. Neither animal had been previously given the hand signal from a trainer. Neither animal had previous physical contact with the hearing test stationing device. Once on the stationing device, animal G began repeatedly producing the raspberry sound, but not in response to the tone signals. Her apparent mental rule was "station on the device and repeat the raspberry sound." However, animal E stationed on the hearing tests device and only produced the raspberry sound in response to the tone signals. His mental rule, the correct one, was "station on the device, produce a raspberry sound in response to a tone, or remain silent otherwise." Interestingly, both animals correctly stationed on the hearing

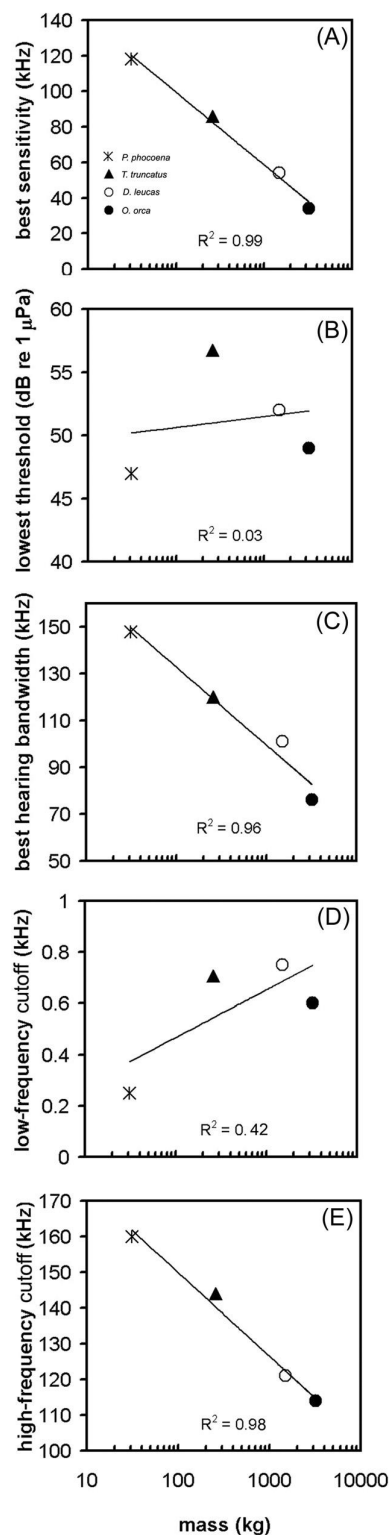


FIG. 8. Relationship between body mass of four odontocete species and hearing metrics from species composite audiograms. (A) "Best sensitivity" is the frequency at which the lowest threshold occurred. (B) "Lowest threshold" is the lowest threshold level in dB re 1 μ Pa. (C) "Best hearing bandwidth" was the difference between the upper and lower frequencies that were 20 dB above the lowest threshold. (D) "Low-frequency cutoff" is the lowest frequency where the threshold is below 100 dB. (E) "High-frequency cutoff" is the highest frequency where the threshold is below 100 dB.

test device when given the novel hand signal, despite no prior reinforcement history with the hand signal. It is not clear if the animals had learned, through observation, that the specific hand signal was a cue to station on the device. An alternative hypothesis is that the small pool, with the hearing test station in the water, provided contextual information to the whales. When the novel hand signal was presented, both animals correctly inferred the correct response based off of observations of the other animals performing the correct behavior. Either interpretation is consistent with observational learning or behavioral imitation. Both animals had developed mental rules (albeit two different rules), about under what conditions they should produce the raspberry sound. Both mental rules were developed without any operant conditioning, strictly from observing the training sessions of the other animals. Unfortunately, these training sessions were impromptu and lacked video documentation and remain anecdotal observations by the first author and several of the training staff at SWSD. Both vocal and behavioral imitation, have been reported in killer whales (Abramson *et al.*, 2013; Crance *et al.*, 2014; Musser *et al.*, 2014). The current study provides an example of comprehension learning (Janik and Slater, 2000), in which a mental rule was learned through observing other animals applying the rule. Killer whales have one of the most sophisticated social structures in the animal kingdom, where groups are often composed of life-long maternal lineages (Bigg *et al.*, 1990). Social learning, including vocal and behavioral imitation, and comprehension learning, are likely important learning mechanisms, and may be particularly well developed in this social species.

IV. CONCLUSIONS

Using the composite model audiogram as a representative species audiogram, the following conclusions can be stated about killer whale hearing: (1) The frequency of best sensitivity is ~ 34 kHz where the lowest threshold is ~ 49 dB SPL, (2) the frequency range of good hearing [i.e., within 20 dB of the lowest threshold is ~ 5.0 kHz to ~ 81 kHz, and (3) the low and high-frequency cutoffs (i.e., highest and lowest frequency below 100 dB SPL) are ~ 0.6 kHz and ~ 114 kHz. Killer whale hearing is similar to other echolocating odontocetes. However, frequency of best sensitivity and upper frequency cutoff appear to be correlated with the animal's large size.

ACKNOWLEDGMENTS

The authors thank the training staff of Sea World San Diego and Sea World San Antonio, who were the backbone of this project. Data collection often required most (if not all) of the training staff to engage all non-participating animals to mitigate noise. Thank you. Special thanks go to Sea World trainers Michelle Bridwell, Summer Mathews, Deneke Wellemeier, Matt Fripp, Katie Danforth, Caitlin Andrew, Shawn Connor, Jennie Krecklow, Laura Meitin, Ann Peterson, Mary Suppa and Emily Wolfe. We would also like to thank Ariel Brewer and Dianna Samuelson Dibble for their help in data collection. Two anonymous

reviewers contributed substantially to the refinement of this publication. Financial support was provided by the Navy Living Marine Resources Program, U.S. Navy Fleet Forces Command, and Sea World Parks and Entertainment. This is scientific contribution #151 from the National Marine Mammal Foundation.

- Abramson, J. Z., Hernandez-Lloreda, V., Call, J., and Colmenares, F. (2013). "Experimental evidence for action imitation in killer whales (*Orcinus orca*)," *Anim. Cognit.* **16**, 11–22.
- Awbrey, F. T., Thomas, J. A., and Kastelein, R. A. (1988). "Low-frequency underwater hearing sensitivity in belugas, *Delphinapterus leucas*," *J. Acoust. Soc. Am.* **84**, 2273–2275.
- Bain, D. E., and Dahlheim, M. E. (1994). "Effects of masking noise on detection thresholds of killer whales," in *Marine Mammals and the Exxon Valdez*, edited by T. R. Loughlin (Academic Press, San Diego, CA), pp. 243–256.
- Bigg, M. A., Olesiuk, P. F., Ellis, G. M., Ford, J. K. B., and Balcomb, K. C. (1990). "Social organization and genealogy of resident killer whales (*Orcinus orca*) in the coastal waters of British Columbia and Washington State," in *Individual Recognition of Cetaceans: Use of Photo-Identification and Other Techniques to Estimate Population Parameters*, edited by P. S. Hammond, S. A. Mizroch, and G. P. Donovan (International Whaling Commission, Cambridge, UK), pp. 383–405.
- Branstetter, B. K., Trickey, J. S., Aihara, H., Finneran, J. J., and Liberman, T. R. (2013). "Time and frequency metrics related to auditory masking of a 10 kHz tone in bottlenose dolphins (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **134**(6), 4556–4565.
- Branstetter, B. K., and Finneran, J. J. (2008). "Comodulation masking release in bottlenose dolphins (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **124**, 625–633.
- Brill, R. L., Moore, P. W. B., Helweg, D. A., and Dankiewicz, L. A. (2001). "Investigating the dolphin's peripheral hearing system: Acoustic sensitivity about the head and lower jaw," 1865 (SSC, San Diego, CA).
- Castellote, M., Mooney, T. A., Quakenbush, L., Hoobs, R., and Goertz, C. (2014). "Baseline hearing abilities and variability in wild beluga whales (*Delphinapterus leucas*)," *J. Exp. Biol.* **217**, 1682–1691.
- Corso, J. F. (1959). "Age and sex differences in pure-tone thresholds," *J. Acoust. Soc. Am.* **31**, 498–507.
- Crance, J. L., Bowles, A. E., and Garver, A. (2014). "Evidence for vocal learning in juvenile male killer whales, *Orcinus orca*, from an adventitious cross-socializing experiment," *J. Exp. Biol.* **217**, 1229–1237.
- Cranford, T. W., Krysl, P., and Hildebrand, J. A. (2008). "Acoustic pathways revealed: Simulated sound transmission and reception in Cuvier's beaked whale (*Ziphius cavirostris*)," *Bioinsp. Biomimet.* **3**, 016001.
- Finneran, J. J. (2003). "An integrated computer-controlled system for marine mammal auditory testing" (SSC, San Diego, CA).
- Finneran, J. J. (2010). "Auditory weighting functions and frequency-dependent effects of sound in bottlenose dolphins (*Tursiops truncatus*)," Report, Office of Naval Research, Washington, DC.
- Finneran, J. J. (2016). "Auditory weighting functions and TTS/PTS exposure functions for marine mammals exposed to underwater noise," SSC Pacific TR 3026 (SSC Pacific, San Diego, CA).
- Finneran, J. J., Carder, D. A., Dear, R., Belting, T., McBain, J., Dalton, L., and Ridgway, S. H. (2005). "Pure tone audiograms and possible aminoglycoside-induced hearing loss in belugas (*Delphinapterus leucas*)," *J. Acoust. Soc. Am.* **117**, 3936–3943.
- Finneran, J. J., and Houser, D. S. (2006). "Comparison of in-air evoked potential and underwater behavioral hearing thresholds in four bottlenose dolphins (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **119**, 3181–3192.
- Finneran, J. J., and Schlundt, C. E. (2007). "Underwater sound pressure variation and bottlenose dolphin (*Tursiops truncatus*) hearing thresholds in a small pool," *J. Acoust. Soc. Am.* **122**, 606–614.
- Fitch, W. T. (1997). "Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques," *J. Acoust. Soc. Am.* **102**, 1213–1222.
- Gellerman, L. W. (1933). "Chance orders of alternating stimuli in visual discrimination experiments," *J. Genet. Psychol.* **42**, 206–208.
- Gulick, W. L., Gescheider, G. A., and Frisna, R. D. (1989). "Hearing loss and audiology," in *Hearing. Physiology Acoustics, Neural Coding, and Psychoacoustics* (Oxford University Press, New York), pp. 350–385.

- Hall, J. D., and Johnson, C. S. (1972). "Auditory thresholds of a killer whale *Orcinus orca* Linnaeus," *J. Acoust. Soc. Am.* **51**, 515–517.
- Heffner, H. E., and Heffner, R. S. (2008). "High-frequency hearing," in *Handbook of the Senses: Audition*, edited by P. Dallos, D. Oertel, and R. Hoy (Elsevier, New York), pp. 55–60.
- Hemilä, S., Nummela, S., and Reuter, T. (2001). "Modeling whale audiograms: Effects of bone mass on high-frequency hearing," *Hear Res.* **151**, 221–226.
- Hoelzel, A. R., Natoli, A., Dahlheim, M. E., Olavarria, C., Baird, R. W., and Black, N. A. (2002). "Low worldwide genetic diversity in the killer whale (*Orcinus orca*): Implications for demographic history," *Proc. R. Soc. B* **269**, 1467–1473.
- Holt, M. M. (2008). "Sound exposure and southern resident killer whales (*Orcinus orca*): A review of current knowledge and data gaps," NOAA Tech. Memo NMFS-NWFSC-89, U.S. Department of Commerce, 59 pp.
- Houser, D. S., and Finneran, J. J. (2006). "Variation in the hearing sensitivity of a dolphin population obtained through the use of evoked potential audiometry," *J. Acoust. Soc. Am.* **120**, 4090–4099.
- Irino, T., and Patterson, R. D. (2001). "A compressive gammachirp auditory filter for both physiological and psychophysical data," *J. Acoust. Soc. Am.* **109**, 2008–2022.
- Janik, V. M., and Slater, P. J. B. (2000). "The different roles of social learning in vocal communication," *Anim. Behav.* **60**, 1–11.
- Johnson, C. S. (1967). "Sound detection thresholds in marine mammals," in *Marine Bioacoustics*, edited by W. N. Tavolga (Pergamon Press, Oxford, UK), pp. 247–260.
- Johnson, C. S., McManus, M. W., and Skaar, D. (1989). "Masked tonal hearing thresholds in the beluga whale," *J. Acoust. Soc. Am.* **85**, 2651–2654.
- Kastelein, R. A., Bunskoek, P., Hagedoorn, M., Au, W. W. L., and de Haan, D. (2002). "Audiogram of a harbor porpoise (*Phocoena phocoena*) measured with narrow-band frequency-modulated signals," *J. Acoust. Soc. Am.* **112**, 334–344.
- Kastelein, R. A., Heul, S.v.d., Verboom, W. C., Jennings, N., Veen, J.v.d., and Haan, D. d. (2008). "Startle response of captive North Sea fish species to underwater tones between 0.1 and 64 kHz," *Mar. Environ. Res.* **65**, 369–377.
- Kastelein, R. A., Hoek, L., de Jong, C. A. F., and Wensveen, P. J. (2010). "The effect of signal duration on the underwater detection thresholds of a harbor porpoise (*Phocoena phocoena*) for single frequency-modulated tonal signals between 0.25 and 160 kHz," *J. Acoust. Soc. Am.* **128**, 3211–3222.
- Kastelein, R. A., Schop, J., Hoek, L., and Covi, J. (2015). "Hearing thresholds of a harbor porpoise (*Phocoena phocoena*) for narrow-band sweeps," *J. Acoust. Soc. Am.* **138**, 2508–2512.
- Ketten, D. R. (1992). "The marine mammal ear: Specializations for aquatic audition and echolocation," in *The Evolutionary Biology of Hearing*, edited by D. Webster, R. Fay, and A. Popper (Springer-Verlag, New York), pp. 717–754.
- Ketten, D. R. (1998). "Marine mammal auditory systems: A summary of audiometric and anatomical data and its implications for underwater acoustic impacts," NOAA-TM-NMFS-SWFSC-256 (Dolphin-Safe Research Program, Southwest Fisheries Science Center, La Jolla, CA).
- Klopper, L. N., Nachtigall, P. E., Gisiner, R., and Breese, M. (2010). "Decreased echolocation performance following high-frequency hearing loss in the false killer whale (*Pseudorca crassidens*)," *J. Exp. Biol.* **213**, 3717–3722.
- Lemons, D. W. (1999). "Auditory filter shapes in an Atlantic bottlenose dolphin (*Tursiops truncatus*)," Ph.D. dissertation, University of Hawaii, 74 pp.
- Lemons, D. W., Klopper, L. N., Nachtigall, P. E., Au, W. W. L., Vlachos, S. A., and Branstetter, B. K. (2011). "A re-evaluation of auditory filter shape in delphinid odontocetes: Evidence of constant-bandwidth filters," *J. Acoust. Soc. Am.* **130**, 3107–3114.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Mann, J., Connor, R. C., Tyack, P. L., and Whitehead, H. (2000). *Cetacean Societies: Field Studies of Dolphins and Whales* (The University of Chicago Press, Chicago, IL), 433 pp.
- May-Collado, J. L., Agnarsson, I., and Wartzok, D. (2007). "Reexamining the relationship between body size and tonal signals frequency in whales: A comparative approach using a novel phylogeny," *Mar. Mammal Sci.* **23**, 524–552.
- Miller, P. J. O., Antunes, R. N., Wensveen, P. J., Samarra, F. I. P., Catarina Alves, A., Tyack, P. L., Kvadsheim, P. H., Kleivane, L., Lam, F.-P. A., Ainslie, M. A., and Thomas, L. (2014). "Dose-response relationships for the onset of avoidance of sonar by free-ranging killer whales," *J. Acoust. Soc. Am.* **135**, 975–993.
- Miller, P. J. O., Samarra, F. I. P., and Perthuisson, A. D. (2007). "Caller sex and orientation influence spectral characteristics of 'two-voice' stereotyped calls produced by free-ranging killer whales," *J. Acoust. Soc. Am.* **121**, 3932–3937.
- Mooney, T. A., Yamato, M., and Branstetter, B. K. (2012). "Hearing in cetaceans: From natural history to experimental biology," *Adv. Mar. Biol.* **63**, 197–246.
- Mulow, J., Houser, D. S., and Finneran, J. J. (2012). "Underwater psychophysical audiogram of a young male California sea lion (*Zalophus californianus*)," *J. Acoust. Soc. Am.* **131**, 4182–4187.
- Musser, W. B., Bowles, A. E., Grebner, D. M., and Crance, J. L. (2014). "Differences in acoustic features of vocalizations produced by killer whales cross-socialized with bottlenose dolphins," *J. Acoust. Soc. Am.* **136**, 1990–2002.
- National Marine Fisheries Service (2016). "Recovery Plan for the Cook Inlet Beluga Whale (*Delphinapterus leucas*)," National Marine Fisheries Service, Alaska Region, Protected Resources Division, Juneau, AK.
- Popov, V. V., Supin, A. Y., Pletenko, M. G., Tarakanov, M. B., Klishin, V. O., Bulgakova, T. N., and Rosanova, E. I. (2007). "Audiogram variability in normal bottlenose dolphins (*Tursiops truncatus*)," *Aquat. Mamm.* **33**, 24–33.
- Ramirez, K. (2012). "Marine mammal training: The history of training animals for medical behaviors and keys to their success," *Vet. Clin. North Am. Exot. Anim. Prat.* **15**(3), 413–423.
- Read, A. J., Wells, R. S., Hohn, A. A., and Scott, M. D. (1993). "Patterns of growth in wild bottlenose dolphins, *Tursiops truncatus*," *J. Zool.* **231**, 107–123.
- Ridgway, S. H., Carder, D. A., Kamolnick, T., Smith, R. R., Schlundt, C. E., and Elsberry, W. R. (2001). "Hearing and whistling in the deep sea: Depth influences whistle spectra but does not attenuate hearing by white whales (*Delphinapterus leucas*) (Odontoceti, Cetacea)," *J. Exp. Biol.* **204**, 3829–3841.
- Schlundt, C. E., Dear, R. L., Green, L., Houser, D. S., and Finneran, J. J. (2007). "Simultaneously measured behavioral and electrophysiological hearing thresholds in a bottlenose dolphin (*Tursiops truncatus*)," *J. Acoust. Soc. Am.* **122**, 615–622.
- Singleton, R. C., and Poulter, T. C. (1967). "Spectral analysis of the call of the male killer whale," *IEEE Trans. Audio Electroacoust.* **15**(2), 104–113.
- Skinner, B. F. (1951). "How to teach animals," *Sci. Am.* **185**(12), 26–29.
- Southall, B. L., Schusterman, R. J., and Kastak, D. (2000). "Masking in three pinnipeds: Underwater, low-frequency critical ratios," *J. Acoust. Soc. Am.* **108**, 1322–1326.
- Szymanski, M. D., Bain, D. E., Kiehl, K., Pennington, S., Wong, S., and Henry, K. R. (1999). "Killer whale (*Orcinus orca*) hearing: Auditory brainstem response and behavioral audiograms," *J. Acoust. Soc. Am.* **106**, 1134–1141.
- White, M. J., Norris, J., Ljungblad, D. K., Baron, K., and di Sciara, G. N. (1978). "Auditory thresholds of two beluga whales (*Delphinapterus leucas*)," report (Hubbs Sea World Research Institute, San Diego, CA).
- Yuen, M. M. L., Nachtigall, P. E., Breese, M., and Supin, A. Y. (2005). "Behavioral and auditory evoked potential audiograms of a false killer whale (*Pseudorca crassidens*)," *J. Acoust. Soc. Am.* **118**, 2688–2695.