Aquatic Mammals 2015, *41*(4), 412-433, DOI 10.1578/AM.41.4.2015.412

The Behavioural Response of Humpback Whales (*Megaptera novaeangliae*) to a 20 Cubic Inch Air Gun

Rebecca A. Dunlop,[1] Michael J. Noad,[1] Robert D. McCauley,[2] Eric Kniest,[3] David Paton,[4] and Douglas H. Cato[5]

[1]*School of Veterinary Science, The University of Queensland, Gatton, Queensland 4343, Australia*
E-mail: r.dunlop@uq.edu.au
[2]*Curtin University, GPO Box U 1987, Perth, WA 6845, Australia*
[3]*School of Engineering, University of Newcastle, Callaghan NSW, 2308, Australia*
[4]*Blue Planet Marine, PO Box 919, Canberra ACT, 2614, Australia*
[5]*Defence Science and Technology Organisation and University of Sydney, PO Box 44, Pyrmont NSW, 2009, Australia*

Abstract

Seismic surveys are widely used for exploration for oil and gas deposits below the sea floor. Despite concern they may have an impact on whale behaviour, our knowledge of marine mammal responses is limited. In the first of a series of experiments (the last one involving a full seismic array), this study tested the response of migrating humpback whale (*Megaptera novaeangliae*) groups to a 20 cubic inch air gun. Experiments were carried out during the southward migration of humpback whales along the east coast of Australia. Groups of whales were focally followed from land stations and/or small boats with observations *before*, *during*, and *after* exposure to a vessel towing the air gun. The source vessel moved either eastwards across the migratory flow or northwards into the migratory flow. In total, there were 18 *control* trials (where the source vessel ran the compressor and towed the air gun without it firing; n = 35 whale groups) and 16 *active* trials (where the air gun was firing every 11 s; n = 32 whale groups). The air gun source level was 199 dB re 1 $\mu Pa^2 \cdot s$ (Sound Exposure Level [SEL]) at 1 m, and SELs received by the whales varied from 105 to 156 dB re 1 $\mu Pa^2 \cdot s$ (modal value 128 dB re $\mu Pa^2 \cdot s$) for SELs at least 10 dB above the background noise (measured as dB re 1 µPa). Other *baseline* groups were focal followed when there was no source vessel in the area (n = 25). Results suggested that humpback whale groups responded by decreasing both dive time and speed of southwards movement though the response magnitude was not found to be related to the proximity of the source vessel, the received level of the air gun, the tow path direction, or the exposure time within the during phase. There was no evidence of orientation of the groups towards, or away from, the source vessel in the during phase. Interestingly, this behavioural response was found in the control trials as well as the active trials suggesting a response to the source vessel.

Key Words: baleen whales, anthropogenic noise, behavioural response, seismic

Introduction

The potential behavioural effect of underwater anthropogenic noise on cetacean ecology is of concern to scientists, industry, government, environmental regulators, conservationists, and other stakeholders. If operations in the ocean are to continue in the way our societies expect, methods must be found to continue these with minimum impact on the environment. This, of course, requires an understanding of the impacts, how they are caused, and the contribution of the many factors that affect the impacts. Behavioural Response Studies (BRS) are used to measure behavioural reactions of animals to various stimuli; and in the context of the effects of anthropogenic noise in the ocean, the stimulus is the underwater noise from some human activity such as the air guns used during seismic exploration.

Noise, however, is usually only one factor involved in the behavioural response to human activity. The noise is an indicator of the presence of the source, but the response may depend on the proximity of the source and the direction that it is moving relative to the subject animal. Early behavioural response experiments on northbound (breeding to feeding grounds) migrating gray whales (*Eschrichtius robustus*) defined a predictable pattern of response to a stationary and towed single air gun (track deflection, decrease in speed, and an increased likelihood of swimming into surf zone or a nearby sound shadow in migrating female-calf pairs) and related some of these behavioural changes to the received level and proximity to the source (Malme et al., 1983, 1984). Similar experiments were carried out on feeding, socializing, and migrating bowhead whales (*Balaena mysticetus*) using either a single air gun or a full-scale seismic array as the test stimulus (Richardson et al., 1985, 1986). Both the gray and bowhead whale studies found similar changes in diving behaviour. When exposed to strong seismic pulses (> 160 dB re 1 μPa), both species tended to display shorter dive and surfacing times with fewer blows per surfacing compared to unexposed whales. This dive response was noted as a common pattern of behavioural change in large whales exposed to human activities (Richardson et al., 1985, 1986). However, other studies have demonstrated that this avoidance response was not predictable in large whales. When another air gun experiment was carried out on the same population of gray whales but when they were migrating southwards (from feeding to breeding grounds) using a moving source, no response was detected (Malme et al., 1984). Further studies looked for specific avoidance behaviours in feeding humpback whales (*Megaptera novaeangliae*) exposed to a single 100 cu in air gun. No avoidance was evident up to a received level of 172 dB re 1 μPa (Malme et al., 1985), even though, on a small number of occasions, a startle response at air gun onset was noted at 150 to 169 dB re 1 μPa at ranges of up to 3 km from the source. McCauley et al. (2003) showed that the response of humpback whales to nearby air guns varied depending on the behavioural context of the whales. Southerly migrating humpback whales showed clear course changes in response to the air gun at received levels of 144 to 151 dB re $\mu Pa^2 \cdot s$ (SEL), whereas resting female-calf pairs showed avoidance responses at a considerably lower level (mean of 129 dB re $\mu Pa^2 \cdot s$).

These previous behavioural studies, though pioneering in their time, used relatively simple experimental designs and highlighted problems with low sample sizes equating to low experimental power. The earlier experiments used comparatively simple statistical analysis to test between *control* (non-exposed) and *exposed* (experimental) whales and, if the sample size allowed, between *before*, *during*, and *after* exposure periods. An appropriate before period provides one type of control, a measure of the behaviour of whales before exposure to the stimulus, to allow some comparison of behaviour to be made during and after exposure. However, in the Malme et al. (1984, 1985) experiments, a number of exposures were carried out each day. As whales analysed in subsequent experiments had potentially heard the earlier air gun trials, they could not be viewed as naive in the before period, and their behaviour may not have been indicative of pre-exposure behaviour. Behaviours such as *length of blow interval*, *surface interval*, *dive time*, or *number of blows per surfacing* were measured repeatedly on one individual and, therefore, were autocorrelated. Despite this, serial dependence issues were not accounted for in the analysis. The Richardson et al. studies (e.g. Richardson et al., 1986) acknowledged this as a problem but did not account for serial dependence because of the complexity of the analysis and the lack of suitable methods. These analyses also did not include variables that might explain some of the behavioural variability of the whales, such as season, group activity states, or whale density, and did not control for any effect due to the presence of the source vessel itself.

Improvements in modelling techniques over the years and the development of more sophisticated statistical analysis packages have allowed many of the previously mentioned analysis issues to be addressed. In addition to addressing autocorrelation issues, modelling techniques can also incorporate other environmental and social effects that may influence the general behaviour of the whales as well as their response to noise from air guns. These environmental and social effects were generally not included in the original studies. More recent studies on large whales used these modelling techniques and incorporated some environmental effects into the analysis. These studies aimed to test the effect of a full seismic survey on feeding behaviour in gray whales (measured by the frequency of visible mud plumes from bottom feeding; Yazvenko et al., 2007) as well as its effect on general abundance, behaviour, and movement (Gailey et al., 2007). Results suggested that there was no measurable effect on gray whale activity as a result of the seismic air gun survey. However, neither study accounted for the non-independence of samples in their analysis (where whales were probably repeatedly sampled between experimental conditions such as *guns off*, *post-seismic*, *strong seismic*, or *weak seismic*). Yazvenko et al. (2007), in particular, noted high variability in feeding activity and had low experimental power so that any subtle behavioural effects of the seismic survey may have been missed.

One way to account for high variability within the response variable being tested is to incorporate a random effect in the statistical model. This random effect accounts for the *within group* variance by randomly selecting a small number of test subjects from a large population. It also accounts for the repeated measures design of the experiment, wherein test subjects are measured before, during, and after the stimulus. Testing for auto-covariance (and accounting for this in the analysis) also eliminates problems with dependence on continuous data collected from an individual. Miller et al. (2009) used this approach when analysing the response of sperm whales (*Physeter macrocephalus*) to a full seismic array and were the first to use digital recording tags (Dtags) as a way of recording fine-scale continuous behaviours (3D movements of the tagged whale when submerged) as well as the sound field at the tagged animal. However, this study encountered a number of problems due to the logistics of designing such a complex experiment at sea (mainly, inadequate sample size, sampling bias towards less sensitive whales that were easier to tag, and the absence of an adequate number of controls). Due to the small sample size, only preliminary conclusions could be made. Although there were some detectable behavioural changes (e.g., possible horizontal avoidance, less pitching during diving, and lower *buzz* rates), they found no broad scale changes in behavioural state (e.g., resting or foraging), and it was suggested that whales tended to exhibit subtle changes in behaviour rather than clear avoidance reactions. Since then, other studies have adapted this modelling approach in their analyses. Robertson et al. (2013) re-analysed the original bowhead whale data from Richardson et al. (1985, 1986) using a mixed-model analysis, though they still had issues in that they sometimes did not know if successive observations came from the same, or different, whales (though individuals were tracked as much as possible). This study accounted for social, temporal, and environmental effects such as season, subject reproductive status, and whale activity and found that, though dive behaviour was significantly affected by the seismic air gun or similar stimulus, this response was context-dependent and differed with season and whale activity.

Two large populations of humpback whales migrate along the east and west coasts of Australia to and from their breeding grounds inside the Great Barrier Reef off Queensland and off the northwestern coast of Western Australia (Chittleborough, 1965). Both populations have increased substantially in the last 30 y (Bannister, 2001; Hedley et al., 2009; Noad et al., 2011a, 2011b; Salgado Kent et al., 2012). At the time of the current study, the east Australian population of whales was approximately 14,500 (Noad et al., 2011b). This population migrates close to the coast along parts of southeast Queensland, allowing land-based as well as boat-based observations. Furthermore, this population of whales would have had little previous exposure to seismic surveys because these were rare on the east coast of Australia (in contrast to the west coast where seismic surveys are common) and could be considered naive. The objectives of the study were (1) to determine if there was any measurable movement, dive, or surface behaviour response of migrating humpback whales to an hour-long presentation of shots from a 20 cu in air gun fired at a typical industry rate; (2) to determine if this response was related to measures of received level and/or the proximity of the source vessel relative to the group, while accounting for other potential social, environmental, and temporal effects in the mixed model analysis; and (3) to determine if the behavioural response was affected by towing the air gun across the migratory stream or directly into the migratory stream. The experimental design consisted of trials with three 1-h observational phases: *before*, *during,* and *after*, with the during phase being the treatment. Treatments were either *active*, wherein the air gun was towed at 7 km/h and operated for the during phase, or *control,* which were identical to the active except that the air gun was not operated. The vessel towing the air gun moved as slowly as possible in the before and after phases (staying in the same area). Observers on land and in small boats made observations of focal follow whale groups for the duration of each trial. In addition, *baseline* data were collected on migrating groups over the same general area in the absence of the source vessel or when it was at least 8 km from the whales. This was the first in a series of four experiments, with the other three involving an array of air guns, and the final one involving a full seismic array.

Methods

Study Site

Experiments were carried out in September and October 2010 and 2011 during the humpback whale southward migration as part of the BRAHSS (Behavioural Response of Australian Humpback whales to Seismic Surveys) project. The study site was located at Peregian Beach (Figure 1), 150 km north of Brisbane, on the east coast of Australia (26° 29' S, 153° 06' E) and about 800 km south of the likely main breeding grounds inside the Great Barrier Reef (Smith et al., 2012). Although humpback whales passing Peregian Beach are migrating from the breeding grounds, they show a range of behaviours typical of breeding grounds (e.g., singing, forming competitive groups, frequent joining and splitting of groups, nursing, and other maternal behaviours due to numerous newborn calves) while moving in a general southward direction. The southward migration of the whales ensured that new whales were present each day, so it is likely that no whale was sampled twice.

Land-Based Observations
Land-based behavioural observations were collected daily (0700 to 1700 h, weather permitting) from two different stations: (1) a northern station (an apartment building 10 km north of the base station at Peregian Beach; Figure 1) and (2) a southern station (Emu Mountain, a 73-m-high hill set 700 m back from the beach, 1.2 km to the south of the base station; Figure 1). Both stations had extensive, essentially unobstructed views of the ocean. The northern station had a field of view of 30° to 165° from the north, while the southern station had a field of view of 10° to 150°, with a large area of overlap between them.

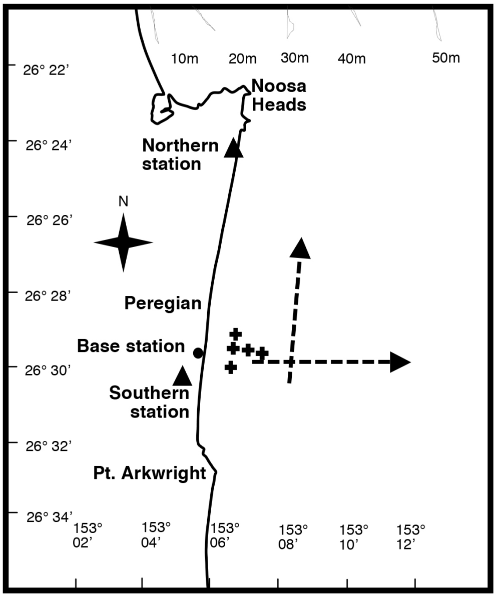

Figure 1. The Peregian study area. The hydrophone buoy array is indicated by crosses, while the two transects of the source vessel, eastwards and northwards, are indicated by the dashed arrows. The land-based observation stations are shown by the triangles.

There were five land-based observation teams operating at any one time: four focal follow teams and one scan sampling team. Two focal follow teams would operate at the northern station where focal follows were usually initiated. In this study, the sample unit was a group of whales, defined as those whales surfacing synchronously and within 100 m of each other. Once focal groups moved south within visual range of both stations, they were passed onto the two focal follow teams at the southern station who then followed the groups until they moved out of the study area or field of view. This allowed for a 3+ h land-based focal follow to be carried out on each target group. The scan sampling team operated at the southern station only, collecting less detailed tracking and behavioural observations on all groups in the study area to provide contextual data for assessing the behaviours of the focal groups. Each team used a theodolite (Leica TM 1100 [magnification 30x] or similar) connected to a notebook computer running *VADAR* software (E. Kniest, University of Newcastle, Australia). *VADAR* records the positions of whales from the theodolite readings in real time, taking into account tide, earth curvature, and refraction. Fixes were annotated with observed behaviours and group compositions. One observer in a focal follow team and three observers in a scan sampling team, each with 7×50 binoculars, were responsible for recording behaviours not captured by the theodolite. The focal follow team would attempt to record all surface behaviours of the target groups. The scan sampling team was responsible for keeping track of all visible groups in the area (including the focal groups) as *ad lib* observations. Data from the visual observers included bearing and distance from the land platform, group composition (number of adults and the presence of a calf), direction of travel, and group behaviours (blow, breach, pectoral flipper slap, tail slap, splitting apart into two groups, joining together of two groups, no blow rise or surfacing, peduncle slap, inverted tail slap, inverted pectoral flipper slap, and head lunge being the majority observed). These were recorded by the *VADAR* operator.

Boat-Based Observations
Some focal groups were followed by small boats. Behavioural observations were made at the individual level as each group member was recognisable at a close range (usually from differences in the shape of the dorsal fin). Two boat-based platforms (one 6-m rigid hulled inflatable boat and one 5.6-m centre console aluminium boat) were used for focal follow data collection, biopsy (for sex identification), and photo-identification of individuals within the focal group. Focal follows were carried out by one observer who continuously spoke behaviours of individuals in the focal group into a voice recorder (M-AUDIO Microtrack) for later transcription. Group size (number of whales in the group), group composition (whales were classified as calves, mothers of calves [females], escorts of females and calves, singers [if audible], or adults), and position estimates (distance and bearing) of the group relative to the boat were also stated regularly. The joining and splitting of animals was also noted. The positions of the boats were recorded by the on-board GPS (Garmin) every 2 to 5 min. Boats attempted to stay 100 to 200 m from their focal group in order to maintain

visibility of individuals and behaviours while minimising disturbance. Data recorded from boat-based platforms were imported into *VADAR* post-field season to allow comparability between land- and boat-based datasets.

Environmental Data

Weather conditions were retrieved from the Australian Bureau of Meteorology automatic weather station at the Sunshine Coast Airport approximately 10 km south of the study site. Data on wind speeds (km/h) were recorded half-hourly throughout the field seasons. Bathymetry and coastline data, including distances to shore, were derived using *ArcGIS 10.0* (ESRI, Redlands, CA, USA). Spatial data were imported using the WGS84 datum and re-projected into a projected coordinate system (Transverse Mercator using UTM Zone 56S) to measure distances between objects (e.g., the focal group and the shoreline or the land-based stations).

Acoustic Data

A fixed array of hydrophones was moored offshore for recording and tracking singing humpbacks as well as monitoring the air gun during trials. Each hydrophone was suspended (using a buoy) over a mooring and connected to a surface buoy which transmitted the acoustic data to a base station on shore using sonobuoy VHF transmitters. Buoys 1 through 3 were 1.5 km from the beach, moored in a line parallel to the shoreline and approximately 700 m apart. Buoys 4 and 5 were moored seaward from buoy 2 in a line perpendicular to the shore and were approximately 500 m apart (Figure 1). At least three buoys were always operational during trials, and this provided real-time fixes of the positions of vocalising (usually singing) whales using *Ishmael* software (Mellinger, 2001). (For further information on the set up and calibration of the acoustic array as well as real-time tracking of singing and vocalising whales, see Noad et al., 2004, and Dunlop et al., 2013a). The errors of a single point localization of a singer were approximately 5% of range at 2 km to 10% at 10 km and 18% at 20 km from the array but were reduced with multiple position estimates (Noad & Cato, 2001). The buoys' transmissions were monitored at the base station. One desktop computer with *Ishmael* software recorded the data to the hard drive when there was a signal of interest (e.g., a singing whale) and during all trials. A second computer also used *Ishmael* software to track vocalizing whales in real time. A third computer with *VADAR* displayed the acoustic tracks of vocalizing whales (from the tracking computer), the positions of the source vessel and small boats, and all the visual tracks of migrating whales from the five theodolite stations from the northern and southern land sites. This information was used by a *trial director* to coordinate all activities during each trial.

Four autonomous acoustic recording systems (CMST acoustic loggers; see www.cms.curtin.edu.au for specifications) were deployed at various positions on the sea floor throughout the area. The loggers were recovered every few days, the data were downloaded, and the loggers were redeployed at different positions to record the sound (levels and characteristics) of the air gun signals at various positions and propagation paths, as well as the background noise. The hydrophones were always located on the seafloor in water depths between 20 and 40 m. Recordings were made at a total of 16 positions overall, with a north-south spread of 16 km and an east-west spread of 7.5 km resulting in an air gun signal range of 100 m to 8.5 km. Loggers sampled 12 min of every 15 min at a 4 kHz sample rate to computer hard disk. Hydrophones used were High Tech Inc. HTI U90, Massa TR1026C, or Reson TC 4034. Some loggers used multiple gains from -3 to 40 dB (total system gain) with the lower gain systems used close to the air gun path to avoid saturation of short range air gun signals. The system response (gain with frequency) of each logger was calibrated before and after field deployments by recording white noise of known level onto the system (in series with the hydrophone to include its effect on response). Each logger's clock was synchronised to GPS transmitted UTC time before deployment; and the drift was read after recovery, allowing derivation of sample times to within a ± 250 ms absolute accuracy. Each logger included two Aquatech 520T temperature loggers—one on the seabed and one 11 m above the seabed.

For near field air gun signals, a High Tech Inc. HTIU90 hydrophone bolted onto the air gun towfish and connected to a 20 dB attenuator was used in 2010. A Geosys Inc. MP 8D hydrophone was set in the frame centre in 2011. The signal was continuously logged during air gun operations to a Sound Devices SD744T digital recorder sampling at 48 kHz, 10 dB gain, and 24 bits to .wav files (uncompressed).

Exposure Stimulus

A Bolt 600B air gun with 20 cu in chamber (operating at 2,000 psi) was towed at a depth of 5.6 m and speed of 7 km/h, 18 m astern of a 19-m vessel, *F/V Ash Dar S*, in 2010, and 22 m astern the 24-m vessel, *R/V Whale Song*, in 2011. The 2010 air gun gas supply setup was similar to that described in McCauley et al. (2003). In 2010, the air gun was charged by a 250 litres/min Bauer electric three-stage scuba compressor using two high-pressure, G-sized gas cylinders as reservoirs. In 2011, the same air gun was used but was in a frame as part of a larger array of air guns (though only the 20 cu in air gun was fired for the trials presented herein). The shot interval was set at 11 s. A larger four stage Atlas Copco Hurricane compressor was used powered by a large diesel engine. An Aquatech 520PT depth/temperature sensor was placed in the towfish to measure the air gun depth. A GPS Genius with purpose built hardware/software was connected to the air gun fire control system to log UTC time and GPS position for the vessel (every 1 s) and every air gun signal (sample triggered by the firing pulse).

Acoustic Analysis

The data from the acoustic loggers were analysed using purpose-built *MATLAB* programs. Air gun signals were high pass filtered above 5 Hz giving a system response of 5 Hz to 1.8 kHz and were corrected for the system frequency response and hydrophone sensitivity in the time domain. To do this, the air gun signal waveform (units of volts at this stage) was extracted with points either side which ensured the sample was a multiple of 2^n points long. An FFT of this waveform was calculated at a fine frequency resolution (< 0.5 Hz); real and imaginary parts of the FFT were corrected for system gain and phase shift across 1 to 1,800 Hz; and an inverse FFT was calculated to give the corrected waveform (units of Pa). The air gun source level was 199 dB re 1 $\mu Pa^2 \cdot s$ at 1-m range. Received levels of the air gun signals were calculated from measurements as the sound exposure level (SEL) defined as

$$SEL = 10 * \log_{10}\left(\int_0^T p_{s+n}^2(t)dt - \int_{T_1}^{T_2} p_n^2(t)dt\right)$$

where p_{s+n} is the acoustic pressure of the air gun signal plus the background noise, T is the length of the air gun signal, p_n is the background noise pressure, T_1 and T_2 specify a time period before or after the air gun signal, and $T_2 - T_1 = T$ (McCauley et al., 2003). In practice, the SEL was calculated using a technique defined by Malme et al. (1986). First, the lowest mean squared pressure value for a section of 2,000 to 4,000 samples before or after the air gun signal that was to be analysed was deemed to be the mean squared background noise. Second, the curve of the cumulative sum of the squared pressure of the air gun signal was calculated as a function of time. At each point along the curve, the product of the mean square noise and the time interval along the curve was subtracted from the sum of the squared pressure of the air gun signal. The maximum value of this difference gave the integral of the squared pressure of the air gun signal corrected for background noise, in units of $\mu Pa^2 \cdot s$. From this, the noise corrected SEL of the air gun signal was determined. The times taken for the cumulative sum curve to reach the 5 and 95% values were set as the start and end of the air gun signal and so defined its duration. Parameters of positive and negative peak, and peak-peak pressure values were also read off the waveforms (all parameters listed in McCauley et al., 2003, Table 6, were calculated for every signal analysed).

The position of the antenna used to log air gun GPS positions, the tow offsets (x-y), and the vessel heading (calculated using GPS coordinates at a 1-s time increment) were used to give the air gun position. Calculations assumed a reasonably straight tow path, which was the case. The time and GPS position of the air gun signal fired were then matched to the time of received air gun signals recorded on the loggers after correcting for travel time to the receiver to correlate each received signal with a fired signal. Received air gun signal parameters at noise loggers were then determined as a function of the distance between the air gun and direction between the air gun tow path and receiver location for each shot. From these data, empirical estimates of the relative transmission loss for the various paths available from the different air gun and receiver positions were derived. There were differences in transmission losses along different propagation paths (along slope compared with up slope) over sand which covered a large part of the area, and these could be characterized by curve fits to the received levels as a function of logarithm of range for the different propagation paths. There were also several large patches of rock seabed in the study area which caused high losses of energy in air gun signals as they crossed the rock patches (*high loss* patches). These rock patches were identified in the noise logger data in 2010 and mapped using a Hummingbird sidescan sonar in 2011, with the positions later verified using the bathymetry slope derived from high resolution (5 m) Lidar bathymetry data obtained from the Queensland Department of Transport. The loss of air gun signal energy across the rock patches was primarily driven by the path length across the rock patch, with other parameters such as seabed roughness, slope, water depth, and distance from seismic source being less significant in setting the variance measured. Using a simple linear fit to loss of normalised air gun received SEL with range across the rock patches (values of received air gun level values normalised to account for the expected loss with range over sand), the high loss rock added an extra 8 dB/km loss (r^2 = 62%)

above the loss expected from a sand-only seabed. Examples of the received 20 cu in air gun signals with no high loss rock and one run with a portion of its travel path over the high loss rock are shown in Figure 2.

Estimates of the received air gun SEL at specific whale groups were determined by using the received level at the nearest logger to the group and an empirical measure of air gun signal loss which assumed a sand-only seabed, which was then adjusted for transmission across the high loss rock patches (multiple patches may have been traversed). All measurements on which the empirical curves were derived and those for individual air gun signals used seabed-mounted hydrophones. In the water depths of the study site (< 40 m), a seabed-mounted hydrophone would be expected to have near highest levels within the water column (because of ground-borne energy); thus, the estimated measures derived are the probable highest throughout the water column at the given range. If the travel path between whale and air gun crossed the high loss rate rock patches, the total path length across the high loss rock patches was calculated and the estimated air gun SEL was reduced by 8 dB/km for this distance (i.e., the empirically derived received level for a sand-only seabed was reduced by an extra 8 dB/km according to the total path length of high loss rock traversed). For all focal whale groups, estimates of the level of every air gun signal and its time received were calculated.

A base background noise measurement was estimated close to the time the air gun signal was received using a sea noise logger dedicated for this purpose on that day. This receiver had a dynamic range suitable for ambient noise measurements (some receivers had low gains to deal with high-level air gun signals so did not accurately measure ambient noise), did not have high-level vessel noise in, and did not include the air gun measurements in the ambient noise calculations. To remove transient ambient noise sources (mostly whales, fish, vessels, and the air gun signals), the receiver with least vessel noise at an appropriate location and gain was selected; curves of ambient noise (from 5 Hz to 1.8 kHz) averaged over 9.22 s were made every 10 s across the day (excluding periods the air gun was operating and removing spikes such as vessel and nearby whale noise). This provided the base ambient noise across the day. The averaged power spectra of the noise were calculated using the 4 kHz samples, averaging across nine 4,096 sample power spectra (0.98 Hz resolution; 9.22 s average), and summing the intensity across the band 5 Hz to 2 kHz (in linear domain, converting to dB and correcting for bandwidth). Since the longer period of ambient noise is statistically stationary at the scale of an air gun signal length (mostly < 1 s here), then the mean squared (also called *rms*) noise level was independent of the averaging time. The signal to noise ratio (SNR) of the air gun signal at the whale group was calculated as the difference between the estimated SEL air gun level at the whale and the rms ambient noise level. The most appropriate measure of noise might be the SEL of the noise in the integration time of the humpback whale

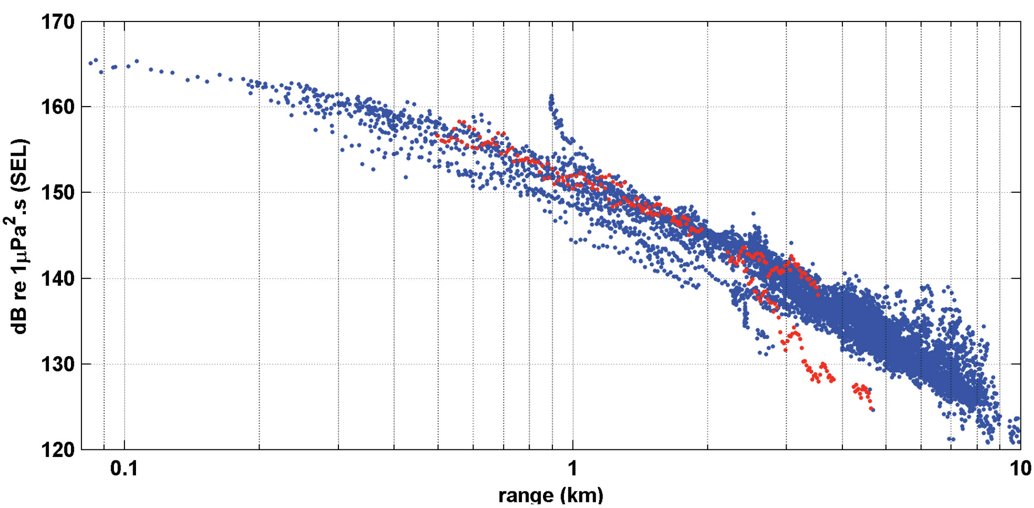

Figure 2. Decay curves of measured 20 cu in air gun SEL with log range; multiple air gun runs and multiple propagation paths are overlaid, and much of the spread in points is due to the differences between runs and paths. The blue curves are multiple runs across sand only; the red curve is one air gun run across the high loss rate rock from greater than 2.36 km on one leg.

auditory system rather than the mean square value which is equivalent to the SEL of the noise for 1 s. Since we do not know the integration time, though we expect from what is known for other mammals that it will be less than 1 s, the mean square value of the noise is a conservative estimate.

Experimental Design

Treatments were either *active* when the air gun was towed at 7 km/h and fired at 11 s intervals or *control,* which were identical to the active trials except that the air gun was not fired while towed.

Five different treatments were used (see Figure 1 which illustrate the different tow paths):

1. *Control east (CE1)* for which the *source vessel* maintained an easterly course (directly offshore and across the migratory path) with the air gun deployed but not firing (2010 only)
2. *Control north (CN)* – CE1 but with the vessel maintaining a northerly course along the coast directly into the migratory path (2010 and 2011)
3. *Active east (AE)* – Similar to CE1 except with the air gun firing at 11-s intervals (2010 and 2011)
4. *Active north (AN)* – Similar to CN except with the air gun firing at 11-s intervals (2010 and 2011)
5. *Control east 2 (CE2)* – Similar to CE but using *R/V Whale Song* (2011 only)—This second set of CE trials was carried out to account for the potential effects of using different source vessels with a different compressor.

When trials were not underway and the source vessel was not in the area or at least 8 km away, groups of whales were focally followed through the study area to provide a baseline dataset (termed *baseline* or *BA*). The 2010 baseline dataset was used in this study as the majority of the data came from this year.

A random block design approach was used to select which trial was to be carried out on each day. Within each set of four trials (AN, AE, CN, and CE1 in 2010 and AN, AE, CN, and CE2 in 2011), the treatment was randomly selected (usually by tossing a coin), but each set of four had to be completed before moving on to the next set of four to allow for a balanced sample size. On days during which two trials were conducted, one active and one control trial were carried out with the control always preceding the active trial (there were never two active trials in the one day). This was so that whales that were to the north of the study area during the morning trial but moved southwards into the study area for the afternoon trial were not pre-exposed to the air gun signals.

Each trial comprised a before, during, and after phase of 1 h each; these were the periods before, during, and after the treatment. The air gun was fired for the treatment (during phase) for active trials, and then towed at the same speed but not fired for the control trials. Prior to the start of the before phase, the source vessel moved to its start position in the southern part of the study area, deployed the air gun or array, then moved very slowly to maintain just enough way to keep the air gun astern of the vessel while staying in the vicinity of the start point. Land- and boat-based focal follow groups were picked up at the northern end of the study site. Generally, there were at least two focal follows per trial, one focally followed by one of the small boats as well as by one of the land stations, and the other left free from contact by the small boats and followed by a land station only. The beginning of the before phase was denoted when there were at least two focal follows underway.

After an hour of the before period, the during period was initiated regardless of where the focal groups were relative to the source vessel. The source vessel would then move along a predetermined path (either eastwards across the migration or northwards against the migration) for 1 h at 7 km/h and then reduced speed to dead slow. Note that the compressor was turned on during control trials as well as active trials. An 11-s interval was chosen as typical industry intervals vary from 8 to 15 s (depending on the water depth and target depth), and a speed of 7 km/h was chosen to match the typical speed of a seismic survey ship. The land observers and boat crews were blind to when the during phase started and finished, and to whether the treatment was an active or a control as they were focused on their focal group (and not the source vessel). After the 1-h during period, the source vessel was brought to dead slow where it stayed within the vicinity of the end mark of its run for the remainder of the trial. This initiated the start of the after period during which the focal groups continued to be followed by land and boat for another hour.

Response Variables

The land-based observations were made for distances up to 20 km and wind speeds up to 37 km/h (20 kts). Data used in the analysis, however, were limited to distances of < 15 km and wind speeds of < 28 km/h (15 kts) because significant numbers of behaviours were likely to be missed beyond these distances and wind speeds (as shown from a preliminary analysis regressing the number of captured behaviours against wind speed and distance). All boat-based data were included, but operations were limited to wind speeds up to

28 km/h, which is considered to be the upper limit for effective focal follow observations.

Dive Behaviour—The two variables used to measure dive behaviour were the length of the group's *long* dive and length of the group's surface interval. Humpback whale dive behaviour consists of a bout of *surfacing dives* (the short, shallow dives that occur during respiration bouts, usually tens of seconds in duration) followed by a *deep dive* in which the group disappears for a longer period of time (usually several minutes). A *deep dive period* is defined as the time from when the last group member disappears to when the first group member reappears, and the *surface interval* is usually defined as the time spent on or just under the surface between deep dives, incorporating all brief surfacing dives. To quantitatively distinguish between short surfacing dives and longer deep dives, a histogram was created using the log of the sighting interval (time between successive surface sightings of animals in the focal group) using land-based data. A best fit density function for the histogram was estimated by choosing the most appropriate bandwidth for display. The distribution of the data was bimodal with the trough between the two modes at a dive interval time of 75 s. This was used as the cut-off time to separate long (probable deep) dives from short (probable surfacing) dives (see Dunlop et al., 2013b, for further details). The longest dive included in this dataset was 57 min 9 s (surfacings from this group were probably missed in this period). The analysis was repeated using boat-based data, and the same cut-off time was found. The longest dive from the boat-based data was 13 min 20 s (800 s); therefore, this was used as the upper likely limit of a group's dive time. During boat-based focal follows, groups were rarely lost. Longer dive times from the land-based data were noted as a *missed surfacing*. Surface interval times for each group were calculated as the time between the end of one long dive and beginning of the next and, therefore, encompassed the entire bout or series of short respiration dives.

Movement Behaviour—Measures of movement (speed and course) were analysed in 10-min time bins. Within each 10-min time bin, the number of surfacings and correlated positions of each group was highly variable. To calculate one standardised measure of speed and course for each 10-min time bin, the position of the group at the start and end of each time bin (the *bin edge*) was estimated and used to determine speed and course made good over the 10-min period (i.e., calculated assuming straight and constant travel between those two points). As the whales were usually submerged at any given time, the time bin's start and end positions were calculated by assuming the whales swam in a straight line at a constant speed between the last measured position of the group in one time bin and the first measured position in the next time bin. If no position was available for one or two sequential time bins, either because the whales did not surface or because the whales surfaced but a theodolite shot was missed (an issue for land-based tracks only), the bin edges were interpolated from the time bins either side of those, again assuming constant speed and course. If a position was not available for more than two time bins in a row, positions were not estimated, and these time periods were excluded from the analysis (as it was assumed the group had been temporarily lost). While extrapolating across empty time bins artificially reduced the variance of course and speed estimates, the effect was small, as for each focal follow there were usually a large number of time bins, and two or more time bins without observations were rare. In baseline groups, the observations were divided into 10-min time bins starting with the first observation of the focal follow. For the experimental datasets, time 0 was when the first air gun shot was fired (active trials) or when the source vessel started to move in the during phase (controls), and 10-min time bins were generated forward and backward from this time.

In addition to speed and course, speed of net southward movement for each bin was also calculated by using only the change in latitude and ignoring longitude. A negative *speed south* indicated net northward movement over the time bin. Whales meander significantly on the southward migration, and some may move north for short distances before resuming their general southward migration. Absolute deviation from a bearing of 180° (the general direction of the migration) was also calculated. Finally, the angle between the direction of travel and the direction of the source vessel was used as a measure of orientation of the group to the source vessel in the during phase. If the group oriented toward the vessel, the angle would decrease. If the group oriented away from the source vessel, the angle would increase.

Surface Behaviour—Surface behaviours were divided into four main categories: (1) blows, (2) breaching behaviours, (3) pectoral behaviours, and (4) fluke behaviours. *Blows* included all sighted blows (plumes of condensed expired air mixed with sea water) as well as times when a whale back was sighted, but there was no visible blow plume (on the likely assumption that the animal did breathe but without an obvious blow). *Breaching behaviours* included all behaviours in which all or part of the body exited the water and forcefully re-entered the water (i.e., head slaps, breaches, half breaches, and head lunges but not pectoral or fluke behaviours). *Pectoral behaviours* included all behaviours in which just the pectoral fin exited and was slapped on the surface of the water (pectoral fin *waves* were not included). *Fluke behaviours* included all behaviours in which the tail fluke or peduncle was slapped against the surface of the water (fluke waving behaviours without a slap were not included). (The omission of waving as opposed to slapping behaviours was because slapping behaviours are likely to be heard by other whales in the area and so are probably signalling behaviours [Dunlop et al., 2010] whereas waving behaviours are not likely to perform the same role.)

The number of sighted blows, breaching, and pectoral and fluke behaviours were summed for each 10-min time bin. When comparing measured blow rates between land- and boat-based platforms (in groups that were followed by both land and boat stations), blow rate was found to be underestimated by the land-based platform compared to the boat-based platform, particularly in groups that contained a calf (calf blows being smaller and harder to spot from land). Therefore, due to the issues with using land-based data to determine blow rate, only the boat-based dataset was used in the analysis of blow rate. This eliminated baseline groups, however, as these groups were mainly followed by land only. An average blow rate per animal per group was calculated by summing the number of blows per 10-min time period and dividing by the number of animals within the group.

Predictor Variables

Predictor variables were divided into five main categories: (1) experimental manipulation, (2) social variables, (3) temporal variables, (4) environmental variables, and (5) data measurement variables. Table 1 lists the predictor variables with a description of each.

The distance to the nearest other group of whales (*nearest neighbour* and *nearest singer*) was categorised as within 1 km, 1 to 2 km, 2 to 5 km, and beyond 5 km as this was deemed to be more robust than using distance as a continuous variable due to potential measurement error. The nearest neighbour data came from the scan platform (theodolite and binocular fixes), but there were limitations in that the further away the focal group was from the scan observers, the more likely groups in the area of the focal group were missed. To account for this, if the focal group was within 15 km of the scan observers, the distance of the nearest neighbour was used. If the focal group was beyond 15 km, the data were not used unless the distance of the nearest neighbour was within 2 km of the focal group as it was unlikely that a closer nearest neighbour was missed given that any close-by neighbour would have been spotted by the focal team. This situation only arose for focal follows from the northern station (the scan team was at the southern station) or for boat-based focal follows. The nearest singer data came from acoustic tracking and, therefore, was subject to some measurement error at long range.

For the 10-min binned dataset, predictor variables were measured in one of three ways: (1) the first observation of each 10-min time bin was used, (2) the observations were averaged over the 10-min time bin, or (3) the minimum or maximum value of the observations was chosen depending on which was the most applicable (Table 1).

The received SEL data were limited such that the signal to noise ratio (SNR) at the group was ≥ 10 dB for at least some of the during phase of active trials to ensure that the air gun was clearly audible for part of the trial. For all but nine active samples (focal follows), the signal to noise levels exceeded 10 dB for the entire during phase, with the highest reaching 54 dB. Within these nine samples, received SELs would have been close to background noise, which varied from 90 to 117 dB re 1 μPa, for a small part of the trial. Received SELs for SNR ≥ 10 dB ranged from 105 to 156 dB re 1 $\mu Pa^2 \cdot s$ with a mode of 128 dB re 1 $\mu Pa^2 \cdot s$ (see Table 2 for sample size).

Statistical Analysis

Generalized linear mixed models (GLMMs) were generated using *R* (R Development Core Team, 2012). GLMMs account for issues of non-independence of data by incorporating random effects as well as issues with non-normally distributed data by specifying the sample distribution and using link functions (see review by Bolker et al., 2008).

For normally distributed response data (speed of southward movement, blow rate per animal, the log of dive time, log of surface interval, log of course deviation from 180° and *speed made good*), the *lme4* package (Bates et al., 2012) was used to compare models that included different combinations of predictor effects. *Group ID* (individual group identity) was included as a random factor. Within model t values with associated p values are reported for specific within-model comparisons. The p values were generated using the *lmerTest* package (Kuznetsova et al., 2013). Model effects (which were back-transformed from logged values if necessary) are reported along with 95% confidence intervals. Each model was inspected for collinearity between variables (e.g., distance offshore and water depth); and, if found, one term was dropped in favour of the other, with the

Table 1. Details of all predictor variables used in the analysis as well as how each predictor variable was used in the 10-min time bin.

Variable	Description	10-min time bin
Experimental manipulation		
Activity	3 activities: trial using active air guns, control trials, or unexposed trials	Not applicable
Treatment	Treatments: BA (baseline), CE1 (control east, ADS), CE2 (control east WS), CN (control north), AE (active east), and AN (active north)	Not applicable
Experimental phase	Before (B), During (D), or After (A) exposure to the control or air gun stimulus or, if baseline, the first, second, and third 60 min of the focal follow	First observation of 10-min time bin
Source vessel proximity	Distance from the source vessel to the focal group at the time of the observation (using GPS data for the source vessel and *VADAR* data for the group). For baseline groups, the source vessel had to be at least 8 km away (regardless of whether it was stationary or moving) for the group to be included.	Minimum (closest) distance of source vessel to group within the 10-min time bin
SEL	The received SEL at the focal group of the air gun shot immediately prior to the observed behaviour or, if there were a number of shots between successive observed behaviours, the maximum level of these shots	Maximum SEL within the 10-min time bin
SNR	The difference between the received SEL and background noise immediately prior to the observed behaviour	Maximum level within the 10-min time bin
Social variables		
Group composition	Composition of the focal group: FC (female with a calf), FCE (female with a calf and escorting adult), FCME (female with a calf and multiple escorting adults), A (lone, single adult), AA (pair), MA (multiple adults, no calf), and MFC (multiple females with calves in the group)	First observation of 10-min time bin
Group social behaviour	*Stable* (focal group not interacting with any other group at the time of the observation), *pre-join* (up to 10 min before a new animal was noted to be part of the focal group), *pre-split* (up to 10 min before an animal was noted to have left the focal group), *joining* (up to 10 min following the time at which a new animal was noted to have joined the focal group), and *splitting* (up to 10 min following the time an animal was noted to have left the group)	First observation of 10-min time bin
Nearest neighbour	The distance of the nearest group to the focal group at the time of the observation categorised into <1 km, 1 to 2 km, 2 to 5 km, and > 5 km from the group at the time of the observation (using *VADAR* fixes from the scan sampling team)	Minimum (closest) distance of nearest neighbour to group within the 10-min time bin
Nearest singer	The distance of the nearest singing whale to the focal group at the time of the observation (as determined by acoustic tracking); categorised into < 1 km, 1 to 2 km, 2 to 5 km, and > 5 km from the group at the time of the observation (using acoustic tracking).	Minimum (closest) distance of singer to group within the 10-min time bin
Density of groups	The number of groups in the study area (within 10 km of Emu Mountain as determined by the scan sampling team)	Maximum number of animals within the 10-min time bin
Density of singers	The number of singing whales in the study area (within 10 km of the array as determined by acoustic tracking)	Maximum number of singers within the 10-min time bin
Temporal variable		
Time of day	Trials were noted as *morning* or *afternoon* depending on when they took place.	Not applicable

Variable	Description	10-min time bin
Environmental variables		
Depth	The water depth that the focal group was in at the time of the observation	Averaged within the 10-min time bin
Distance from shore	The distance from shore of the focal group at the time of the observation	Averaged within the 10-min time bin
Wind speed	Wind speed at the time of the observation	Averaged within the 10-min time bin
Background noise	Measured by the nearest acoustic logger to the whale location (dB re 1 μPa)	Averaged within the 10-min time bin
Measurement variables		
Platform of observation	Named as Land-Only, Boat-only or Land/Boat depending on whether the group was followed by the land station, the research vessel, or both	Not applicable
Distance from platform	The distance of the observed group from focal follow platform; observations not used beyond 15 km from the land station	Minimum (closest) distance within the 10-min time bin
Dataset	Land focal follow or boat focal follow dataset	Not applicable

retained variable being the more significant predictor variable.

For count data such as number of breaches, pectoral slapping behaviours, or tail slapping behaviours (per group per 10 min), the *glmmADMB* package (Fournier et al., 2012) was used to generate the models. This package specifically accounts for the problems of zero-inflated count data by using Laplace approximation to estimate the parameters of the model, which is believed to be more accurate for count data. The models assumed a negative binomial distribution with zero inflation to account for the skew towards zero.

A GLMM was fitted to each response using group ID as the random effect. Within model *z* values with associated *p* values are reported for within-model comparisons. All model residuals were checked for homoscedasticity, normality, and autocorrelation.

First, a base model of normal behaviour was generated. An initial study of the baseline data for each response variable (dive variables, movement variables, and surface behaviour variables) determined which of the predictor variables (social, temporal, and environmental variables in Table 1) were important predictors of normal behaviour (Kavanagh, 2014). These variables were then retested for significance (using either the *lme4* package or *glmmADMB* package depending on the response variable). Within-model significance was set at $p < 0.05$. Predictor variables, if significant within the base model, were retained, and nonsignificant predictor variables were rejected. If the base model was deemed to contain too many significant predictor variables (due to limitations with sample size), an analysis of deviance was used to determine which of the variables to include and which to reject. Results of the analysis of deviance are reported as *F* values with associated degrees of freedom (*df*) and *p* values; significant predictor variables with the highest *F* values were included.

To test the hypothesis that humpback whale groups, after accounting for predictors of normal behaviour, significantly changed their behaviour in response to the presence of the source vessel with or without the air guns firing in the during phase of the experiment, the term *treatment*phase* (the interaction effect between treatment and phase) was added to the base model. This was termed the *experimental model*. The before phase and the baseline treatment data were set as the intercept. Base and experimental models were compared using Akaike Information Criterion (AIC) scores and checked for significant ($p < 0.05$) improvement using the maximum likelihood ratio (LR) test, where the probability distribution of the test statistic is a chi-squared distribution and the degrees of freedom equals df_1 to df_2 (where df_1 and df_2 are the degrees of freedom for the two models being compared). Significant model improvement suggested that treatment, phase, or the interaction effect of treatment and phase were significant predictors of the behavioural response variable (though only the results of the interaction effect are reported). To test if the behavioural response to the air gun (if significant) and the experimental variables differed in different social contexts, female-calf (FC) and female-calf-escort (FCE) groups, being the two most common group compositions, were selected and analysed separately.

In this study, it was assumed that the presence of the near-stationary source vessel (with engines running) would have no significant effect on behaviour in the before and after phases of the experiment (the whales were generally several kilometres away). Therefore, only the during phase of CE1, CE2, CN, AN, and AE groups was used to test the effect of vessel proximity, received level (SEL), SNR, and experimental time (time relative to the start of the during phase, i.e., the first shot in active trials or when the vessel started moving in control trials). This analysis was only performed on response variables that were found to significantly change in the during phase and included an additional response variable—the orientation of the group to the source vessel.

Results

Diving Behaviour

A preliminary analysis (using a mixed model analysis as detailed in the analysis section) comparing long dive times measured by the land- and boat-based platforms following the same group found no significant difference between the measured dive times from the two platforms. A second preliminary analysis was carried out comparing long dive times between groups followed by land only and groups followed by a research vessel to check for any behavioural effect of the presence of the boat. It was found that dive times were not significantly different between the two. Therefore, both platform datasets were included in the one analysis model. If groups were followed by both land and boat platforms, boat-based observations were used in place of the land-based to create a more even sample size between the two platforms.

Table 2. The focal follow sample size for the treatments, including the number of trials (or sampled days for the baseline groups [BA]) and total number of focal follows

Treatment	Number of trials	Total focal follows
AE	8	16
AN	8	16
CE1	5	8
CE2	5	12
CN	8	15
BA	11	25

Group long dive times ranged from 75 to 800 s with a mean of 250 s. The analysis of the baseline dataset (Kavanagh, 2014) found that water depth, group composition, group social behaviour, distance of the nearest neighbour, distance of the nearest singer, and wind speed were all significant predictor variables of group dive time. However, given the number of parameters, and the sample size with which to test them, the base model only included the three most significant predictors (based on the results of the analysis of deviance). All subsequent models testing the effect of experimental variables on group diving behaviour, therefore, included group composition ($F_{(6)} = 5.42$, $p = 0.01$), group social behaviour ($F_{(4)} = 3.82$, $p = 0.03$), and water depth ($F_{(1)} = 27.32$, $p < 0.0001$) as the important social and environmental predictor variables, but not wind speed, distance of the nearest neighbour, nor distance of the nearest singer. Table 3 summarises the base model for dive time showing all predictor variables and, of those, ones that were included in this base model (including within-model effects from this analysis), those that were significant but not included, and those that were not found to be significant in the Kavanagh (2014) study.

There was a significant improvement in the experimental model compared to the base model ($\chi^2_{(17)} = 27.51$, $p = 0.05$), suggesting a significant dive response in groups to the treatments. Baseline groups tended to increase their dive time in the second hour of the focal follow (the nominal during phase) and maintain this longer dive time in the third hour (the nominal after phase) as the majority of groups migrated in a general south-southeast direction and, therefore, into deeper water during the focal follows. In contrast, groups in the during phase of AE, AN, and CE2 trials significantly shortened their dive times ($t = -3.49$, $p = 0.0005$; $t = -3.62$, $p = 0.0003$; and $t = -3.01$, $p = 0.003$ for groups within AE [-29 s], AN [-23 s], and CE2 [-16 s] trials, respectively), suggesting a dive response to both active and to one control treatment. There was no significant difference in long dive times between AE and AN groups suggesting no effect of the direction of travel of the source vessel in active trials. Groups in the after phase of CE2 trials were also found to have significantly shorter dives (-27 s; $t = -2.23$, $p = 0.02$).

FC and FCE groups accounted for 80 of the 92 focal groups. FCE groups ($n = 31$) only had significantly shorter dives in the after phase of AE, CE2, and CN trials ($t = -2.97$, $p = 0.003$; $t = -3.02$, $p = 0.002$; and $t = -3.19$, $p = 0.002$, respectively) with no significant dive response in the during phase of trials. FC groups ($n = 49$), on the other hand, had significantly shorter dives in the during phase of AN and CE2 trials only ($t = -2.43$, $p = 0.02$ and $t = -2.04$, $p = 0.04$, respectively) suggesting that FC groups were more likely to show a dive response in the during phase of trials compared to FCE groups.

To increase the sample size, samples for active treatments and those for control treatments were pooled to give one set for active and one for control, and the analysis was repeated (giving *n*

Table 3. Results of the base model analysis for diving behaviour. Significant social, environmental, and temporal variables found using a previous analysis (Kavanagh, 2014) were retested for significance, but only the three most important were included. Those labelled "NS" were not found to be significant in Kavanagh (2014) and were not considered for inclusion in this study. Those labelled "Not included" were found to be significant in Kavanagh (2014) but were not included in this model due to sample size limitations. The within-model effects are reported as t, and the associated p values relative to the variable are indicated by *.

	Dive time	Surface interval
Social variables		
Group composition	Longer in pairs	Longer in MFC groups; shorter in lone adults and pairs
*FC (female-calf)		$t = -2.25$, $p = 0.02$
A (single adult)		
FCE (female-calf-escort)		
FCME (FC-multiple escorts)		
MFC (multiple FCs)		$t = 2.13$, $p = 0.03$
MA (multiple As)		
AA (adult pair)	$t = 2.92$, $p = 0.001$	$t = -2.24$, $p = 0.03$
Social behaviour	Shorter in groups after a split	NS
$F_{(4)} = 2.76$		
*Stable		
Pre-join		
Post-join		
Pre-split		
Post-split	$t = -2.36$, $p = 0.02$	
Nearest neighbour	Not included in model	NS
*+5,000		
2,000-5,000		
1,000-2,000		
< 1,000		
Nearest singer	Not included in model	NS
*+5,000		
2,000-5,000		
1,000-2,000		
< 1,000		
Group density (0 to 9)	NS	NS
Singer density (0 to 3)	NS	NS
Temporal variables		
Time of day	NS	NS
Environmental variables		
Water depth (11 to 58 m)	Longer in deeper water $t = 5.34$, $p < 0.0001$	NS
Wind speed (3 to 28 km/h)	Not included in model	NS
Distance from shore (152 m to 15 km)	Not used; correlated with depth	Not used; correlated with depth
Background noise (90 to 117 dB re 1 μPa)	NS	NS
Distance of the nearest vessel	NS	NS

Table 4. Summary of the responses to the air gun sounds in terms of estimates of the effect size (back transformed if required) and 95% confidence intervals for dive time and speed of southward movement in the during and after phases of trials. Values in bold were found to be significant ($p < 0.05$) in the fitted model. The values for control and active treatments are the differences between the response values and the intercept model value shown in the line above, with the 95% confidence intervals in brackets.

	During phase	After phase
	Group dive time model intercept (s) 129 (103, 161)	
Control	**-22 (-5, -40)**	-17 (-3, -37)
Active	**-35 (-18, -52)**	-19 (-1, -39)
	Group speed of southward movement model intercept (km/h) 2.8 (1.9, 3.7)	
Control	**-0.8 (-0.1, -1.6)**	0.1 (-0.8, 0.9)
Active	**-1.0 (-0.2, -1.9)**	0.6 (-0.3, 1.5)

= 35 for active and n = 35 for control). The new experimental model was significantly improved compared to the base model ($\chi^2_{(8)} = 20.46$, $p = 0.009$) where groups significantly shortened their dives in the during phase of both active ($t = -3.99$, $p < 0.0001$) and control ($t = -2.44$, $p = 0.01$) treatments (effect sizes are reported in Table 4). Groups within the active treatment dived for significantly shorter times ($t = -2.24$, $p = 0.03$) compared to groups within the control treatment in the during phase. FC groups had significantly shorter dive times in the during phase of active trials only ($t = -1.98$, $p = 0.05$) whereas FCE groups had significantly shorter dives in the after phase of active ($t = -2.40$, $p = 0.02$) and control trials ($t = -3.32$, $p = 0.001$), suggesting FC groups were more likely to respond within the during phase of an active treatment.

The group surface interval was highly variable, ranging from 5 to 7,200 s (the longest being for logging groups that stayed at or near the surface for extended periods). Group composition was the only significant variable in predicting the length of group surface interval (Table 3), and the addition of the term *treatment*phase* did not significantly improve the base model for this response variable.

Movement Behaviour

As with the dive analysis, a preliminary analysis found that the land and boat observers measured the course and speed of groups similarly. There was no significant difference found in any measure of movement behaviour between groups followed by a boat and groups followed by land only. Therefore, it was assumed the presence of the small research vessel did not have a significant effect on the movement of the focal group. Kavanagh (2014) found that course deviation from 180° at this site was significantly related to water depth, wind speed, group social behaviour, group composition, and distance to the nearest singer during normal behaviour. The base model for course deviation from 180° (ranging from 1° to 180° with a mean magnitude of 46°, indicative of the general southward movement) in this study only included water depth ($F_{(1)} = 5.90$, $p = 0.001$) as this was the only significant effect found in this analysis (Table 5). The speed made good of the groups ranged from 0 km/h (wherein the group was usually *logging*) to 19 km/h with a mean of 4 km/h, agreeing with a previous study at the same site on swimming speed (Noad & Cato, 2007). Kavanagh (2014) found that wind speed and group composition were important predictors of speed made good although wind speed was not found to be a significant predictor variable in the within-model analysis. Group composition ($F_{(6)} = 4.48$, $p = 0.002$) was the only significant within-model predictor variable in this analysis (Table 5).

When the term *treatment*phase* was added to the base model to produce an experimental model, there was no significant improvement. Similarly, there was no significant improvement after pooling active and control treatments, suggesting groups did not significantly deviate from their general southerly course during the trials compared to baseline data and did not significantly change their travel speed.

The speed of southward movement ranged from -7 km/h (wherein the groups travelled in a northerly direction) to 17 km/h with a mean of 3 km/h. Kavanagh (2014) found that wind speed and group composition were important predictors of speed of southward movement. These variables were retested, but the effect of wind speed was

Table 5. Results of the base model analysis for movement. Significant social, environmental, and temporal variables found using a previous analysis (Kavanagh, 2014) were retested for significance. The within-model effects are reported as *t*, and associated *p* values relative to the variable are indicated by *. Those labelled "NS" were either not significant in Kavanagh (2014) or not significant in this analysis.

	Course dev. 180°	Speed made good	Speed south
Social variables			
Group composition	NS	Greater in	Greater in
*FC (female-calf)			
A (single adult)		$t = 2.70, p = 0.002$	$t = 2.45, p = 0.01$
FCE (female-calf-escort)		$t = 3.66, p = 0.0001$	$t = 2.54, p = 0.01$
FCME (FC-multiple escorts)			
MFC (multiple FCs)			
MA (multiple As)		$t = 3.66, p = 0.0001$	$t = 2.45, p = 0.02$
AA (adult pair)			
Social behaviour	NS	NS	NS
*Stable			
Pre-join			
Post-join			
Pre-split			
Post-split			
Nearest neighbour	NS		Increased with closer nearest neighbour
*+5,000			
2,000-5,000			$t = 2.67, p = 0.02$
1,000-2,000			$t = 3.84, p = 0.004$
< 1,000			$t = 1.94, p = 0.04$
Nearest singer	NS		NS
*+5,000			
2,000-5,000			
1,000-2,000			
< 1,000			
Group density (0 to 9)	NS	NS	NS
Singer density (0 to 3)	NS	NS	NS
Temporal variable			
Daytime	NS	NS	NS
Environmental variables			
Water depth (11 to 58 m)	Increased in deeper water $t = 2.02, p = 0.02$	NS	NS
Wind speed (3 to 28 km/h)	NS	NS	NS
Distance from shore (152 m to 15 km)	Not used; correlated with depth	Not used; correlated with depth	Not used; correlated with depth
Background noise (90 to 117 dB re 1 µPa)	NS	NS	NS
Distance of the nearest vessel	NS	NS	NS

not significant and so was not used in the base model. Group composition ($F_{(6)} = 2.92$, $p = 0.01$) and the distance of the nearest neighbour ($F_{(3)} = 5.31$, $p = 0.002$) were found to be significant predictor variables and so were included in the base model for this response variable (Table 5). There was a significant improvement in the model with the addition of the term *treatment*phase* ($\chi^2_{(17)} = 34.95$, $p = 0.006$) wherein groups within the during phase of CE2 ($t = -2.04$, $p = 0.04$), CE1 ($t = -3.20$, $p = 0.001$), AE ($t = -1.91$, $p = 0.05$), and AN ($t = - 2.15$, $p = 0.03$) significantly slowed their southward movement (a reduction of 0.9 to 1.8 km/h, depending on the treatment) compared to baseline groups. No significant difference in response was found between AE and AN groups, again suggesting that the direction of the source vessel (across or against the migration stream) had no effect on their response. FC groups displayed significantly slower southward movement in the during (t = -4.70, p < 0.0001) and after phases of CE1 trials (t = -2.31, p = 0.02) only; however, FCE groups did not significantly change their speed of southward movement during any treatment.

The re-analysis using pooled control and active trials also found a significant model improvement from the base model when including the term *treatment*phase* ($\chi^2_{(8)}$ = 24.57, p = 0.001). Groups in the during phase of both control (t = -1.93, p = 0.05) and active (t = -2.38, p = 0.02) trials displayed significantly slower southward movement with no significant difference between the two treatments. There was no significant response found in either FC or FCE groups separately (effect sizes are reported in Table 4).

Surface Behaviour

Blow rates ranged from 0 to 18 blows per animal per 10 min with a mean of 6 blows/10 min. Significant within-model predictor variables for blow rate included group composition ($F_{(6)}$ = 2.25, p = 0.04) and group social behaviour ($F_{(4)}$ = 1.60, p = 0.05) (Table 6). There was no significant model improvement when adding in the term *treatment*phase* to the base model, suggesting no significant change in blow rate to any treatment.

Breaching rates ranged from 0 to 42 per group per 10 min. Kavanagh (2014) found wind speed, group composition, and nearest neighbour to be significant predictors of breaching behaviour. These factors were retested, and the same three were found to be significant in this dataset (Table 6). The inclusion of the term *treatment*phase* did not significantly improve the model for breaching behaviour, suggesting no significant change in this response variable during any of the treatments. Using pooled data, the experimental model was significantly improved compared to the base model ($\chi^2_{(8)}$ = 34.94, p < 0.0001). However, this was due to a significant decrease in breaching behaviour in the after phase of control groups (t = -3.28, p = 0.001) only. There was no evidence of a significant change in breaching behaviour in the during phase of the trials.

Tail slapping behaviour rates ranged from 0 to 59 per group per 10 min. Rates were significantly related to three variables: (1) group composition, (2) group social behaviour, and (3) water depth (Kavanagh, 2014). All three variables were significant in this dataset (Table 6) and, therefore, were included in the base model. The term *treatment*phase* significantly improved the base model ($\chi^2_{(8)}$ = 39.23, p = 0.001), but the responses were variable between treatments. Tail slapping behaviours significantly increased in the during phase (z = 2.50, p = 0.01) and after phase (z = 3.27, p = 0.001) of CE2 trials and in the after phase of CN trials (z = 1.97, p = 0.05) only. Pooling the data into active and control datasets as before did not significantly improve the base model for tail slapping, suggesting no consistent change in this response variable during either treatment.

The base model for pectoral slapping behaviours (ranging from 0 to 41 per group per 10 min) included group social behaviour and depth (Table 6) but not distance of the closest singer (which was also found to be a significant variable in Kavanagh, 2014). There was significant model improvement with the addition of the term *treatment*phase* ($\chi^2_{(8)}$ = 25.64, p = 0.007) where groups displayed significantly decreased pectoral slapping behaviours in the during phase of CE1 trials (z = -2.34, p = 0.02) only. However, pooling the data into active and control as before showed there was no significant change in pectoral slapping behaviour during either treatment.

Effect of Proximity, Received Level, and Exposure Time in the During Phase

The during phase dataset (control and active trials only) was limited to groups that came within 10 km of the source vessel (n = 31 active groups and 31 control groups). Neither the dive response (wherein groups displayed a decrease in dive time in the during phase) nor movement response (wherein groups were found to significantly decrease southward movement in the during phase) were found to be significantly correlated to the proximity of the source vessel, the time from first shot, or when the vessel first started moving. The orientation of the group to the source vessel in the during phase was not found to be significantly different between groups in control and active trials, suggesting groups did not orient towards or away from the source vessel.

Within the during phase of active treatments, the dataset was limited to groups receiving air gun SELs (for single pulses) of over 90 dB re 1 $\mu Pa^2 \cdot s$ (ranging from 90 to 156 with modal value of 128 dB re 1 $\mu Pa^2 \cdot s$). There was no significant relationship between dive time or speed of south movement and the received SEL (base model predictors were included in this analysis) or SNR. To reduce the variance associated with the social environment, the dataset was limited to stable FC groups (assumed to be the most sensitive cohort) with the nearest neighbour beyond 5 km (SELs

Table 6. Results of the base model analysis for surface behaviour. Significant social, environmental, and temporal variables found using a previous analysis (Kavanagh, 2014) were retested for significance. The within-model effects of significant predictor variables in this analysis are reported as z, and associated p values relative to the variable are indicated by *. Those labelled "NS" were either not significant in Kavanagh (2014) or not significant in this analysis.

	Breach	Pec slap	Tail slap	Blow rate
Social variables				
Group composition	FCME groups	NS	MFC and FCME	Increased in FC
*FC (female-calf)	breached more		groups tail slapped	compared to:
A (single adult)			more	
FCE (female-calf-				$t = -2.33, p = 0.01$
escort)				
FCME (FC-multiple	$z = 2.00, p = 0.04$		$z = 1.67, p = 0.05$	
escorts)			$z = 4.28, p < 0.001$	
MFC (multiple FCs)				$t = -2.45, p = 0.01$
MA (multiple As)				$t = -3.68, p < 0.001$
AA (adult pair)				
Social behaviour	NS	Increased after a split	Decreased before a	Increased before a join
			split and increased	
*Stable			before a join	$t = 2.44, p = 0.02$
Pre-join			$z = 3.46, p = 0.001$	
Post-join			$z = -2.06, p = 0.04$	
Pre-split		$z = 2.22, p = 0.03$		
Post-split				
Nearest neighbour	Decreased with	NS	NS	NS
*+5,000	close-by groups			
2,000-5,000				
1,000-2,000	$z = -3.96, p < 0.001$			
< 1,000				
Nearest singer	NS	NS	NS	NS
$F_{(3)} = 2.40$				
*+5,000				
2,000-5,000				
1,000-2,000				
< 1,000				
Group density (0 to 9)	NS	NS	NS	NS
Singer density (0 to 3)	NS	NS	NS	NS
Temporal variable				
Daytime	NS	NS	NS	NS
Environmental variables				
Water depth (11 to 58 m)	NS	Decreased with increasing depth $z = -3.02, p = 0.002$	Decreased with increasing depth $z = -2.52, p = 0.01$	NS
Wind speed (3 to 28 km/h)	Increased with increased wind speed $z = 3.14, p = 0.002$	NS	NS	NS
Distance from shore (152 m to 15 km)	Not used; correlated with depth			
Background noise (90 to 117 dB re 1 μPa)	NS	NS	NS	NS
Distance of the nearest vessel	NS	NS	NS	NS

ranged from 100 to 160 dB re 1 μPa^2·s). Although there was a trend in the southward movement (in that the speed of southward movement decreased with an increase in SEL and SNR), this was not significant due to the wide confidence intervals (and a small sample size of only 24 data points from 11 groups). Other variables were not tested as the analysis did not show any significant response to the active treatments.

Discussion

Measured response variables included group dive metrics (group long dive time and group surface interval), group movement (changes in course and speed), and individual respiration rate and surface behaviour rates. There was no evidence of a significant and consistent change in any of these behavioural parameters in the during phase of active or control trials apart from a short-term decrease in dive time and an about 1 km/h decrease in the speed of southward movement. However, the dive and movement responses were also found in control trials, suggesting that the responses were to the source vessel (with the compressor running) rather than specifically to the air gun shots. During the active trials, the background noise levels would have included the vessel, compressor, and air gun shot noise with air gun noise clearly audible to the whales over the vessel noise. In active trials, the noise levels received from the air gun were 10 to 54 dB higher than the background noise for at least part of the during phase in all groups. There was no measurable difference in the response between two different tow paths nor was there any evidence of a relationship of either response variable to the proximity of the source vessel, the received SEL of the air gun shots (apart from a potential trend in FC groups), or the amount of time the groups were exposed in the during phase (time from first shot or when the vessel first started moving). Therefore, the results of this study suggest that humpback whale groups showed little behavioural response that could be specifically attributed to the 20 cu in air gun stimulus as distinct from other stimuli associated with the source vessel.

Despite carrying out a power analysis to determine minimum sample size before the experiment (Dunlop et al., 2012), it is still possible that the sample size used in this study was not sufficient to detect subtle responses. One way to test for this, though controversial, is to carry out a post-experiment power analysis. For example, Robertson et al. (2013) re-analysed data on the response of bowhead whales to seismic air gun and array sounds (original study reported in Richardson et al., 1985, 1986) and found the sample size (n = 15) allowed only a 40% chance of detecting a real effect from exposure to seismic air guns. The pre-experimental power analysis for the current study was based on results of a previous study, carried out at the same study site, using the same population of whales, but using a different stimulus: an artificially generated tone sweep. It suggested that a sample size of 12 focal follow groups per treatment was appropriate, assuming a similar response to the air gun sounds. Humpback whales were found to respond to this tone stimulus by changing course as well as dive behaviour, and this response was detected with a smaller sample size than was used in the current study (Dunlop et al., 2013b). The current study aimed for, and achieved, a sample size of at least 15 per treatment. Samples for control trials and those for active trials were also pooled in some analyses to increase the sample size (and, therefore, analysis power) per treatment, and the same behavioural response results were found as when the treatments were analysed separately.

The results of this study are consistent with previous suggestions that humpback whales show little or no behavioural response to human activities such as blasting (Todd et al., 1996). However, in this study, only a small air gun was used, and received levels would have been significantly lower at any distance than from a commercial seismic array (e.g., 20 dB lower than a 2,760 cu in array; McCauley et al., 2003). The received levels per shot in our study varied from close to background noise to 156 dB re 1 μPa^2·s (mode of 128 dB re 1 μPa^2·s), and it is possible the received levels were not high enough to produce larger and more consistent behavioural responses in the tested groups (all active trials did, however, have SELs ≥ 105 dB re 1 μPa^2·s and SNRs ≥ 10 dB for at least some of the during phase and were therefore audible). The purpose of this experiment, however, was not to emulate a full commercial array but to aid in interpreting responses to a full array (in subsequent experiments), especially the early stages of ramp-up. The experimental design did not follow those of other behavioural response experiments in large whales, for which specific groups were sometimes targeted and approached in order to purposely increase the received level or force a response (e.g., McCauley et al., 2003). Rather, the source vessel in this study followed a predetermined path and did not attempt to intercept or approach any groups. Further, the source vessel started 1 h after the beginning of the trial, regardless of the proximity of the target groups to the source vessel, resulting in a spread of group distances at the start of the exposure phase. This design was chosen as it was deemed to be more realistic in terms of how most groups of whales would encounter a seismic vessel during an offshore seismic survey and as it focuses on determining whale behavioural responses to air gun exposure and their significance. Survey vessels move on a predetermined path and do not deviate (unless turning at the end of a line, and then only very slowly).

Despite the fact that there was no significant reduction in swimming speed or course variation from due south, there was a reduction in net southward migratory speed. Although this seems contradictory, it could have occurred as some groups may have slowed their swimming speed but not changed course, while others may have changed course but not slowed their swimming speed. Therefore, there may not be enough change in either contributing factor to be found to be significant whereas the interaction was. The fact that groups slowed their southward migration by a small amount during exposure may indicate a small avoidance response during which animals will slow down or move on a less southerly course to allow the vessel to pass before moving onwards. This response would also minimise the chance of groups receiving high received levels during the active trials. If humpback whales were exposed to the noise of a full source, one would predict the whale behavioural responses to increase at comparable ranges, whales to maintain a greater distance from the source, or (most likely) a combination of the two. Further experiments using a larger source should elucidate which to be true.

In addition to testing for a response to the air gun stimulus, this study took a different approach than most other behavioural response studies in that baseline groups were first analysed to determine which environmental and social effects may affect each behavioural response variable (based on Kavanagh, 2014). The fact that many social and environmental effects were found to be significant predictor variables of group behaviour suggests that humpbacks are responding to these variables despite being exposed to a moving source vessel and to air gun sounds. For example, humpback whale groups often changed behaviour when other animals joined or when there was a singer in the area regardless of whether air guns were firing or not firing. As a consequence, these social and environmental effects, which are difficult to control for from an experimental perspective (Cato et al., 2015), should be taken into account when assessing behavioural response to an anthropogenic stimulus. In this study, attempting to account for these effects by including them in the analysis model also proved to be an important step in avoiding erroneous results. For example, a preliminary analysis (without including other environmental and social effects) found a significant increase in dive time in the after phase of the trials. However, this was found to be no longer significant when water depth was included in the analysis. The baseline data showed that as groups moved through the study area, they tended to move further offshore and into slightly deeper water over time and so were usually in deeper water for the after phase. Since the whales dived longer in deeper water, the deeper dives originally found during the after phase were unrelated to the treatment. This shows not only the value of including other social and environmental effects, but it also shows the value of including baseline data in the analysis. The other advantage of using baseline data was that it was possible to show that there were responses to both the control and active trials. If a base model had not been possible, then the results, comparing just the control and active trials, would have demonstrated little or no significant difference, and it would have been assumed there was no response. From the point of view of understanding behaviour and responses to human activities, this is important information that may have been missed.

One of the original aims of the study was to test for differences in the reaction to anthropogenic sound with social context as found in previous studies (e.g., Ellison et al., 2011; Dunlop et al., 2013b). However, this study had limited power to test for social context effect other than comparing female-calf (FC) groups and female-calf pairs being escorted by another adult (FCE), which were the two most common groups and, thus, provided adequate sample size. Results indicated a difference in dive response between the two group types in that FCE groups tended to decrease their dive time in the after phase whereas FC groups decreased their dive time in the during phase. Neither group type significantly slowed their speed of southwards movement (in the pooled analysis). However, the lack of obvious response to the air gun sounds, and the small response magnitude, made it difficult to test for any consistent response differences with social context.

Although the results of this study did not find any evidence of a strong behavioural response that was specifically related to the exposure to a small-scale air gun stimulus, it does provide a framework with which to carry out further studies during which a larger source can be tested. The results illustrate the value of carrying out adequate controls (both baseline data and experimental controls) in behavioural response experiments, and these are often not carried out in large marine mammal behavioural research studies.

Acknowledgments

Funding was provided as part of the Joint Industry Programme on E&P Sound and Marine Life (JIP), managed by the International Association of Oil & Gas Producers (IOGP). The principal contributing companies to the programme are BG group, BHP Billiton, Chevron, ConocoPhillips, Eni, ExxonMobil, IAGC, Santos, Statoil, and Woodside. The U.S. Bureau of Ocean Energy Management (BOEM), Origin Energy, Beach Energy, and AWE provided support specifically for the BRAHSS study. We would like to thank all the BRAHSS team members, particularly Louise Bennett and Verity Steptoe, and the BRAHSS volunteers for their assistance with data collection. We would also like to thank Ailbhe Kavanagh, Michael Williamson, Elise Godwin, and Sasha Collecutt for their assistance with preliminary data analysis, and Simon Blomberg for his advice on statistical analysis.

Literature Cited

Bannister, J. L., & Hedley, S. L. (2001). Southern hemisphere Group IV humpback whales: Their status from recent aerial surveys. *Memoirs of the Queensland Museum*, *47*(2), 587-598.

Bates, D., Maechler, M., & Bolker, B. (2012). *Package lme4* [Documentation file, online]. Document last modified on 23 June 2012. Available at http://cran.stat.sfu.ca/web/packages/lme4/lme4.pdf.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Henry, M., . . . White, J-S. S. (2008). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology and Evolution*, *24*(3), 127-135. http://dx.doi.org/10.1016/j.tree.2008.10.008

Cato, D. H., Dunlop, R. A., Noad, M. J., McCauley, R. D., Kniest, E., Paton, D., & Kavanagh, A. S. (2015). Addressing challenges in studies of behavioral responses of whales to noise. In A. N. Popper & A. Hawkins (Eds.), *The effects of noise on aquatic life II*. *Advances in Experimental Medicine and Biology* series. New York: Springer-Verlag.

Chittleborough, R. G. (1965). Dynamics of two populations of the humpback whale, *Megaptera novaeangliae* (Borowski). *Australian Journal of Marine and Freshwater Research*, *16*, 33-128. http://dx.doi.org/10.1071/MF9650033

Dunlop, R. A., Cato, D. H., & Noad, M. J. (2010). Your attention please: Increasing ambient noise levels elicits a change in communication behaviour in humpback whales (*Megaptera novaeangliae*). *Proceedings of the Royal Society B: Biological Sciences*, *277*(1693), 2521-2529. http://dx.doi.org/10.1098/rspb.2009.2319

Dunlop, R. A., Cato, D. H., & Noad, M. J. (2012). Behavioral response studies: Problems with statistical power. In A. N. Popper & A. Hawkins (Eds.), *The effects of noise on aquatic life* (pp. 293-297). *Advances in Experimental Medicine and Biology*, *730*. New York: Springer Science + Business Media, LLC. http://dx.doi.org/10.1007/978-1-4419-7311-5

Dunlop, R. A., Noad, M. J., & Cato, D. H. (2013a). Source levels of social sounds in migrating humpback whales (*Megaptera novaeangliae*). *The Journal of the Acoustical Society of America*, *134*(1), 706-714. http://dx.doi.org/10.1121/1.4807828

Dunlop, R. A., Noad, M. J., Cato, D. H., Kniest, E., Miller, P., Smith, J. N., & Stokes, D. M. (2013b). Multivariate analysis of behavioural response experiments in humpback whales (*Megaptera novaeangliae*). *Journal of Experimental Biology*, *216*, 759-770. http://dx.doi.org/10.1242/jeb.071498

Ellison, W. T., Southall, B. L., Clark, C. W., & Frankel, A. S. (2011). A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Conservation Biology*, *26*(1), 21-28. http://dx.doi.org/10.1111/j.1523-1739.2011.01803.x

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., . . . Siber, J. (2012). *AD model builder*: Using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods and Software*, *27*(2), 233-249. http://dx.doi.org/10.1080/10556788.2011.597854

Gailey, G., Würsig, B., & McDonald, T. L. (2007). Abundance, behavior, and movement patterns of western gray whales in relation to a 3-D seismic survey, Northeast Sakhalin Island, Russia. *Environmental Monitoring Assessment*, *134*, 75-91. http://dx.doi.org/10.1007/s10661-007-9812-1

Hedley, S. L., Bannister, J. L., & Dunlop, R. A. (2009). *Group IV humpback whales: Abundance estimates from aerial and land-based surveys off Shark Bay, Western Australia, 2008* (Report SC/61/SH23). Paper presented to the International Whaling Commission.

Kavanagh, A. (2014). *The behaviour of humpback whales: An analysis of the social and environmental context variables affecting their behaviour on migration* (PhD dissertation). School of Veterinary Science, The University of Queensland, Australia.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). *lmerTest*: Tests for random and fixed effects for linear mixed effect models (lmer objects of *lme4* package). *R package, Version 2.0-3*. Retrieved 15 July 2015 from http://CRAN.R-project.org/package=lmerTest.

Malme, C. I., Miles, P. R., Clark, C. W., Tyack, P., & Bird, J. E. (1983). *Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behaviour* (Report 5366). Prepared by Bolt Beranek & Newman Inc., Cambridge, MA, for U.S. Minerals Management Service, Anchorage, AK.

Malme, C. I., Miles, P. R., Clark, C. W., Tyack, P., & Bird, J. E. (1984). *Investigations of the potential effects of underwater noise from petroleum industry activities on migrating gray whale behavior – Phase II* (Report 5586). Prepared by Bolt Beranek & Newman Inc., Cambridge, MA, for U.S. Minerals Management Service, Anchorage, AK.

Malme, C. I., Miles, P. R., Tyack, P., Clark, C. W., & Bird, J. E. (1985). *Investigations of the potential effects of underwater noise from petroleum industry activities on feeding humpback whale behavior* (Report 5851). Prepared by Bolt Beranek & Newman Inc., Cambridge, MA, for U.S. Minerals Management Service, Anchorage, AK.

Malme, C. I., Würsig, B., Miles, P. R., Bird, J. E., & Tyack, P. (1986). *Behavioral responses of gray whales to industrial noise: Feeding observations and predictive modeling* (Report No. 6265). Prepared by BBN Laboratories, Cambridge, MA, for National Oceanic and Atmospheric Administration, Anchorage, AK.

McCauley, R. D., Fewtrell, J., Duncan, A. J., Jenner, K. C. S., Jenner, M. N., Penrose, J. D., . . . McCabe, K. (2003). Marine seismic surveys: Analysis and propagation of air-gun signals; and effects of exposure on humpback whales, sea turtles, fishes and squid. In *Environmental implications of offshore oil and gas development in Australia: Further research* (pp. 364-521). Canberra: Australian Petroleum Production Exploration. Retrieved 22 July 2015 from www.cmst.curtin.edu.au/publicat/index.html#2000.

Mellinger, D. K. (2001). *Ishmael 1.0 user's guide* (NOAA Technical Memorandum OAR PMEL-120). Washington, DC: National Oceanic and Atmospheric Administration.

Miller, P. J. O., Johnson, M. P., Madsen, P. T., Biassoni, N., Quero, M., & Tyack, P. L. (2009). Using at-sea experiments to study the effects of air guns on the foraging behavior of sperm whales in the Gulf of Mexico. *Deep-Sea Research I*, *56*, 1168-1181. http://dx.doi.org/10.1016/j.dsr.2009.02.008

Noad, M. J., & Cato, D. H. (2001). A combined acoustic and visual survey of humpback whales off southeast Queensland. *Memoirs of the Queensland Museum (Special Issue on Humpback Whales)*, *47*(2), 507-523.

Noad, M. J., & Cato, D. H. (2007). Swimming speeds of singing and non-singing humpback whales during migration. *Marine Mammal Science*, *23*, 481-495.

Noad, M. J., Cato, D. H., & Stokes, M. D. (2004). Acoustic tracking of humpback whales: Measuring interactions with the acoustic environment. In *Proceedings of Acoustics 2004, Annual Conference of the Australian Acoustical Society, Gold Coast, 3–5 November* (pp. 353-358).

Noad, M. J., Dunlop, R. A., Paton, D., & Cato, D. H. (2011a). Absolute and relative abundance estimates of Australian east coast humpback whales (*Megaptera novaeangliae*). *Journal of Cetacean Research and Management (Special Issue)*, *3*, 243-252.

Noad, M. J., Dunlop, R. A., Paton, D., & Kniest, H. (2011b). *Abundance estimates of the east Australian humpback whale population: 2010 survey and update* (Report SC/63/SH22). Paper submitted to the International Whaling Commission Scientific Committee, Tromsø, Norway.

Richardson, W. J., Fraker, M. A., & Würsig, B. (1985). Behaviour of bowhead whales (*Balaena mysticetus*) summering in the Beaufort Sea: Reactions to industrial activities. *Biological Conservation*, *32*(3), 195-230. http://dx.doi.org/10.1016/0006-3207(85)90111-9

Richardson, W. J., Würsig, B., & Greene, C. R., Jr. (1986). Reactions of bowhead whales, *Balaena mysticetus*, to seismic exploration in the Canadian Beaufort Sea. *The Journal of the Acoustical Society of America*, *79*(1), 1117-1128. http://dx.doi.org/10.1121/1.393384

Robertson, F. C., Koski, W. R., Thomas, T. A., Richardson, W. J., Würsig, B., & Trites, A. W. (2013). Seismic operations have variable effects on dive-cycle behavior of bowhead whales in the Beaufort Sea. *Endangered Species Research*, *21*, 143-160. http://dx.doi.org/10.3354/esr00515

Salgado Kent, C. P., Jenner, K. C. S., Jenner, M., Bouchet, P., & Rexstad, E. (2012). Southern hemisphere breeding stock D humpback whale population estimates from North West Cape, Western Australia. *Journal of Cetacean Research and Management*, *12*(1), 29-38.

Smith, J. N., Grantham, H. S., Gales, N., Double, M. C., Noad, M. J., & Paton, D. (2012). Identification of humpback whale breeding and calving habitat in the Great Barrier Reef. *Marine Ecology Progress Series*, *447*, 259-272. http://dx.doi.org/10.3354/meps09462

Todd, S., Stevick, P., Lien, J., Marques, F., & Ketten, D. (1996). Behavioural effects of exposure to underwater explosions in humpback whales (*Megaptera novaeangliae*). *Canadian Journal of Zoology*, *74*, 1661-1672. http://dx.doi.org/10.1139/z96-184

Yazvenko, S. B., McDonald, T. L., Blokhin, S. A., Johnson, S. R., Melton, H. R., Newcomer, M. W., . . . Wainwright, P. W. (2007). Feeding of western gray whales during a seismic survey near Sakhalin Island, Russia. *Environmental Monitoring Assessment*, *134*, 93-106. http://dx.doi.org/10.1007/s10661-007-9810-3