Report of the Intersessional Workshop to Plan a Large-Scale Whalewatching Experiment (LaWE)

1. INTRODUCTORY ITEMS

1.1 Opening remarks

Bejder welcomed the participants (Annex A), and thanked funding organisations (the Australian Government Dept of Water, Heritage, and Arts; International Association of Oil and Gas Producers; International Fund for Animal Welfare; City of Bunbury; Perth Convention Bureau; US Dept. of Commerce/National Oceanic and Atmospheric Administration; Western Australia Department of Environment and Conservation; Bunbury Port Authority; Australia Southwest Development Commission; and Cetacean Society International).

Lusseau explained the background for the Workshop. Worldwide in 2001, whalewatching was worth approx $1 billion US (Hoyt, 2001); now it is closer to $1.5-2 billion US. This rapid growth of the industry has been difficult to follow from an impact assessment perspective. The first assessments of boat impacts came from Glacier Bay, Alaska, in the late 1970s, when an abandonment of the area by humpback whales (*Megaptera novaeangliae*) was hypothesised to be due to whalewatching from cruise ships (Baker and Herman, 1989). Since then, many assessments of varying qualities have been done, but there are many areas where no comparable work has taken place. The current Workshop originated from discussions at the Cape Town whalewatch workshop in 2004, which attempted to look for general guiding principles to understand the impacts of whalewatching interactions with whales and dolphins. The importance of the current Workshop was further necessitated by the following statement in 2006 from the Scientific Committee (SC) that 'The Committee agreed that there is new compelling evidence that the fitness of individual odontocetes repeatedly exposed to cetacean watching vessel traffic can be compromised and that this can lead to population effects. The Committee recommends that similar studies looking at individual fitness of cetaceans be carried out whereever possible. However, in the absence of these data it should be assumed that such effects are possible until indicated otherwise. The Committee strongly encouraged the development of similar studies on large whales.' The purpose of this Workshop, then, was to try to design a study that would not be seriously compromised by limitations of funding and logistics to find guiding principles to determine how whalewatching interacts with other pressures on cetaceans to affect life history parameters. The Workshop report will be presented to the Scientific Committee at IWC/60. The study topic may consider all cetaceans, not just large whales.

Bjørge thanked all sponsors on behalf of the IWC and acknowledged Murdoch University, the City of Bunbury, and the Australian Government for their hospitality. He also thanked the IWC steering group for planning the Workshop and Bejder, for organising logistics. He explained the procedures of the IWC Scientific Committee and how this Workshop related to the work of the Committee.

Dr. Sally Talbot, Parliamentary Secretary of the Environment for Western Australia, addressed the group, welcomed them, and gave a short address about the concern for cetaceans in Western Australia.

1.2 Election of Chair

Bjørge was elected Chair.

1.3 Appointment of Rapporteur

Weinrich acted as rapporteur with assistance from Williams and the Chair.

1.4 Adoption of Agenda

The amended Agenda is given in Annex B.

1.5 Documents available

The documents available for the Workshop were SC/LAWE08/1-13, as well as published documents, as needed (see Annex C).

2. WHALEWATCHING: STATUS OF KNOWLEDGE – STRENGTHS AND WEAKNESSES OF CURRENT ENVIRONMENTAL IMPACT ASSESSMENTS (EIAS)

Lusseau summarised recent advances in the understanding of interactions between cetaceans and tourist vessels, which show that the observed short-term behavioural alterations, resulting from apparent horizontal and vertical avoidance tactics, can lead to long-term consequences for the viability and fitness of individuals and their populations (SC/LAWE08/1). The Population Consequences of Acoustic Disturbances (PCAD) framework (shown in Annex D) is useful in defining the mechanistic link between disturbances and these consequences. Increased energetic challenges, either as added travelling costs or reduced foraging opportunities, can lead to reduced fitness for individuals. If such challenges occur too often, individuals may shift into long-term avoidance strategies of the degraded areas. However, such long-term decisions have to be balanced with other costs and benefits to leave a habitat degraded by whalewatching or leave a school exposed to whalewatching. Individuals that cannot leave degraded habitats may have reduced fitness that can lead to reduced reproductive success as observed in Shark Bay, Australia. Studies show that disturbances can act as a selection pressure on cetacean populations by removing individuals that are more sensitive.

In Milford Sound, New Zealand, the whole bottlenose dolphin population (*Tursiops* sp.) avoided the fjord during high interaction periods (both within and between seasons) but only some individuals abandoned the degraded area in the Shark Bay bottlenose dolphin population, Australia. The reduced reproductive success of the individuals remaining in the degraded area in Shark Bay was directly related to their cumulative exposure to boat interactions. The effect size of responses to boat interactions in the degraded area was smaller than in a control site. While this could be interpreted as habituation, knowing that some individuals selectively abandoned the degraded site shows instead that the difference in effect size was related to selectivity on the type of individuals that stayed in the degraded area. Either individuals that were sensitive to boat interactions left the area or individuals that had left the area were those that could energetically afford to respond to the disturbance. There was a discussion on whether to consider sound as the primary reason for responses of cetaceans to whalewatching, as exemplified by the PCAD model. Many factors may be involved, however in some highly productive areas, where underwater visibility is very poor, sound may be the primary stimuli. The discussion concluded by noting that, for now, the whalewatching boat should be considered as a 'black box' that may cause disturbance.

Bejder presented recent developments showing promising avenues for the interpretation of short-term responses (SC/LAWE08/2). These frameworks and models rely heavily on assumptions and simplifications, and ultimately provide only *predictions* of the outcomes of anthropogenic disturbances on important measures of population health, e.g. reproduction and survival. Direct, long-term information on the population measures that models predict would provide a more satisfactory approach. The Indo-Pacific bottlenose dolphin (*Tursiops aduncus*) population in Shark Bay, Australia, provides a system where possible cumulative, long-term effects of an apparent benign human activity could be tested on female reproductive success while also considering kernel home range size and characteristics. Eighty-four females were identified for whom calf survival was known and >4 years of reproductive data were available since the onset of tourism in Shark Bay. For each female, a 'cumulative exposure index' (CEI) to tour vessels was calculated. The higher cumulative exposure a female had to tour vessels, the less likely she was to produce a calf that survived to independency. Cumulative vessel exposure had a greater influence on reproductive success than all other factors considered. Home range size did not influence female reproductive success. However, the larger the proportion of 'deep water' habitat in the female's home range, the less likely she was to reproduce successfully. Similarly, the larger the proportion of 'shallow water' habitat in the female's home range, the more likely she was to reproduce successfully. There was no relationship between the reproductive success of females that reproduced successfully and their cumulative vessel exposure, i.e. if a female reproduced successfully, her reproductive rate was not affected by vessel exposure or habitat type within her home range. Reduced reproductive success in females exposed to tourism interactions may not jeopardise the large, genetically-diverse Shark Bay population. However, a similar impact might have serious consequences for small, closed or isolated cetacean populations. Cetaceans in other areas are subjected to much greater vessel exposure than in Shark Bay, but the longitudinal data needed to identify biologically significant impacts are generally unavailable.

In the study's results, research survey effort was higher in the whalewatching area because it is closer to the landing spot, but none of the data used in any part of the study came from whalewatching vessels. There was considerable discussion about whether the near-shore animals, in the whalewatch area, would be more or less likely to have a detected 'reproductive failure' because they were more likely to be sighted near the research vessel's landing point. Bejder suggested that given the number of research groups and the year-round coverage this was not a major factor, and it might actually be less likely to capture reproductive events in areas further away. The Workshop agreed that sighting rates should be included in future analyses.

The Workshop discussed possible driving mechanisms causing calf mortality. Calves that died in the study showed some evidence of malnourishment prior to disappearing. Further studies to examine whether there was evidence of differential health in different calves, and how that relates to the hypothesised vessel effects would be useful. The Workshop recommended the development of a quantifiable index of body condition and the examination of the effect of group size on reproduction to investigate the effects of anti-predator strategies. Whalewatching 'exposure' in this study was clarified as not being the actual time that each animal was observed, but the amount of time that individuals spent in whalewatching areas.

Bejder proposed that Shark Bay is a good test case where there is a wealth of background data on a coastal animal with a small home range (and therefore may be most likely to be affected). However, it was suggested that there may be a potential problem with the measure of CEI. If an individual spends 0.1 of its time in an area with 0.9 of the whalewatching, it has the same CEI as an individual that spends 0.9 of its time in an area with 0.1 of the whalewatching. The former might cause loss of a calf because the disturbance is large; the latter might because the time spent disturbed is large. On the other hand, 0.1 of the whalewatching may not reflect any disturbance of importance and 0.1 of the time in an area may not be enough exposure to cause an effect. The measure should keep these two variables separate. Hence, using this as an exemplary 'test case' may not be ideal.

Continued monitoring of the Shark Bay population over the next four years will not allow a test of whether recent management measures to reduce whalewatching exposure reverses hypothesised effects in reproductive differences. However, it will be easier to test if the 15% decline of animals seen in the whalewatch area that occurred when a second boat was permitted (Bejder *et al*., 2006) reverses in the short term, with initial results likely to be available by 2010. It was noted that measures of prey would be an important addition to this study, and in general, continuous measures of prey abundance may be important to back-calculate effects which may not be obvious at the time.

Weinrich presented a review and combined analysis of whalewatching effects in studies that have taken place to date (SC/LAWE08/3). To find sources, a 32,000 entry Microsoft Access database of marine mammal literature containing sources from the mid-1600s to 2008 was searched for a variety of keywords (disturbance, effects, whalewatching, tour, etc.) along with additional sources. In all, 55 sources were reviewed: 43 from peer-reviewed literature, and 12 in grey literature (government reports, theses, conference abstracts). An additional seven sources could not be procured for the review (usually older thesis or government reports). Species most studied were bottlenose dolphins (17 studies), humpback whales (10), and killer whales (*Orcinus orca*; 8). In all, 14 studies were on mysticetes, and 40 on odontocetes (4 on large odontocetes, 36 on small odontocetes). Studies took place in N. America (20), Australia/New Zealand (17), S. America (10), Europe (5), Africa (2), and Asia (1). A variety of habitat types were studied: Year-round (25), Feeding grounds (11), Combined (7), Breeding/Calving (6), Migratory Corridor (4) and 'Other' (2). Of all the studies, 16 were before/after control-impact (BACI) studies, while 39 used different groups of animals as control and test groups. Each study was given a highly subjective categorical ranking of quality, based on a variety of factors including study design, sample size, analysis, etc. In total, 13 studies were rated as 'excellent', 5 'very good', 28 'good/average' and 9 'fair.' Several common comments were heard from study reviewers. They included the fact that studies had small sample sizes, authors did not present data sets clearly even when they presented test results, statistical tests were poorly presented with key values missing, and a wide variety of different measures were reported (SD, SE, variance, etc.).

From the papers, 1,138 'results', tests, or comparisons were extracted. This number is larger than it appears; sometimes results that measured closely related variables had to be extracted separately (e.g. if five respiratory variables were examined in both males and females, that would be 10 results). Among these, 334 results were based on no statistical tests, 445 were based on parametric tests, and 359 were based on non-parametric tests. Variables examined in the study varied widely. Common variables included the percentage of time in particular behavioural states, orientation in regards to a compass or a boat, and a suite of respiratory variables. Few studies used multi-variate models to look at a suite of factors, but rather often presented a simple comparison of frequencies, percents, or a statistical comparison of frequencies or means. Hedge-G tests were performed on three variables: swimming speed measured on the surface, blow rates, and inter-breath intervals. Only swimming speed showed significant differences, with baleen whales showing generally increased swimming speeds while bottlenose dolphins showed no change. However, the comparisons made were extremely limited by the large variation in studies and differences in research design.

In conclusion, there was little consistency between studies, and standardisation of techniques would be helpful, as would clearer presentation of the results in published studies. More complex analyses of what are clearly multi-variate states might be helpful, but they make inter-study comparisons difficult. More studies have been done on smaller, more accessible species, and more data are needed on large cetaceans.

The Workshop discussed the study's contention that data collected from boats were more precise than those collected from a more distant land station. Subtle behaviours (e.g. trumpet blows in humpback whales), were thought to be more likely to be missed in studies conducted from a distance. However, in boat-based studies, the design must take into account the possibility that the observation vessel itself may be part of the disturbance stimulus that could elicit the response you are trying to measure. 'Observer effect' needs to be accounted for, and attributing a response to the correct stimulus is important.

It was noted that some of the 'limitations' of published papers went far beyond whalewatching effects studies and were common problems in papers from many disciplines. Many editors require removal of big tables, and detail. The trend toward 'currency of ideas' may be limiting our ability to compare and contrast studies. Researchers were urged to insist on including their data in publications in as much detail as possible, or to set up archival datasets or on-line appendices. The Workshop **recommended** the design of a good archival system for summarised data.

For tagging data, both tracks and diving behaviour are time series, and the archival process is straightforward. Behavioural data are less straightforward, and the number of variables can become problematic. Datasets that are not comparable are harder to archive. Prioritising and standardising data types is important – respiratory variables seem to be important for impact studies. Criteria for selecting variables may include, but should not be limited to, the ease of collection. Some study sites, and species, may lend themselves to certain types of data collection.

In acoustics, several workshops have been held where different people have analysed the same datasets; they often lead to widely different conclusions from different people. This can be a test to see how researchers can reach consensus on the interpretation of a dataset.

Data necessary for these studies go beyond behavioural data – they can be demographic data, or other data to address life history parameters. The ownership and access of proprietary data is another potentially important issue. In at least one case among participants, the host university owns the data and may not be willing to make them available. There may be differences in accessibility between data collected under federal funding (for instance, in the United States NSF funded research data must be made available within a year) or that privately funded. The IWC has a data availability protocol, but it may be more detailed than necessary for whalewatching studies. However, the principle that any data used for management advice should be available to those participating in the discussion (with some conditions) may be useful. The importance of combined raw data sets that can lead to much more powerful analyses than possible from the component sets was acknowledged. In many cases, data on which publications have been based are stored in notebooks, individual's computers, or other inaccessible places; at a minimum, these should be made available.

3. CETACEAN BIOLOGY AND ECOLOGY: KEY VARIABLES TO MONITOR

Hammond presented a very brief introduction to population dynamics and cetacean life history, commented on methods for estimating life history parameters and introduced a comparison of estimated life history parameters for selected species that may be subject to whalewatching (SC/LAWE08/4). Cetaceans are characterised by high survival rates and low reproductive rates, although there is considerable variation among species. Conventionally, cetaceans are seen as 'K-selected' species that have evolved to maintain population size close to the carrying capacity of their environment. Density dependent mechanisms to increase abundance are thus believed to start to act relatively rapidly as a population drops below its carrying capacity. These mechanisms may include a decrease in the age at first reproduction, a decrease in inter-birth interval, or an increase in survival rate. Studies to investigate the sensitivity of the dynamics of cetacean populations to variation in life history parameters have typically found, as expected, that growth rate and probability of extinction are most sensitive to adult survival rate, followed by calf/juvenile survival rate and fecundity. A recent study by

McDonald *et al*. (SC/LAWE08/6), however, found that dynamics were more strongly influenced by reproduction than by survival rates.

In studies of cetacean life history, age at first reproduction, inter-birth interval and survival rate have been estimated from post-mortem examination of reproductive organs from stranded or bycaught animals (e.g. Danil and Chivers, 2007; Stolen and Barlow, 2003) or mark-recapture analysis of individual recognition (photo-id) data (e.g. Barlow and Clapham, 1997; Gabriele *et al*., 2001; Mizroch *et al*., 2004; Ramp *et al*., 2006; Zeh *et al*., 2002). A recent extensive review of the literature on estimates of cetacean life history parameters has been made as part of an exercise to estimate percent mature and generation length of all cetacean populations to inform IUCN Red List assessments (Taylor *et al*., 2007). Parameters included: age at first reproduction (AFR), inter-birth interval (IBI), maximum age of reproductive females (O), calf survival rate (S_0), and non-calf survival rate (S_A). Estimates of these parameters were presented for right whale (*Eubalaena* spp.), humpback whale, bottlenose dolphin and Hector's dolphin (*Cephalorhynchus hectori*) to illustrate variation in life history characteristics in some species that are subject to whalewatching.

Ability to detect changes in life history parameters, especially adult mortality rate, depends on probability of capture – higher capture probabilities lead to a greater ability for detection. For instance, in a population of bottlenose dolphins in the Sado Estuary, Portugal that numbers only 30 animals where each animal is seen almost each year, it was easy to show differences in adult survival (Gaspar, 2003).

The Workshop discussed whether there were ecological correlates with variability in inter-birth intervals. Species with longer lives generally have longer inter-birth intervals. In baleen whales, episodic failures in reproduction, whether because of longer inter-birth intervals or lower calf survival, are common; it has been hypothesised that longevity in baleen whales may have evolved to compensate for these decadal-scale episodes (Kraus and Rolland, 2007). Individual variation is also very large and very important, and effects may relate to age class. This could be extrapolated to consideration of 'key individuals'; if these animals are forced to leave an area, the effects may be great.

There was discussion about the influence of sociality on the vulnerability of individuals to disturbance. The complexity of the mother-offspring bond and of the rearing network, as well as the prolonged time to weaning in odontocetes and some mysticetes, adds to the potential for disturbances to have indirect effects on calf survival. Such information could, and should, be used in a consideration of whalewatching impacts. Further, the diversity of social systems within and between these groups could be another variable which is considered in a larger analysis. The disturbance of social networks might have effects on reproductive rates, but ultimately reproductive rate is the key parameter. However, if a relationship is established, it may be possible to use impact on social systems as a proxy for reproductive rates.

The Workshop **agreed** that it was essential to take account of variation in life history between and within species and for individuals.

Costa emphasised the dynamics of the marine environment and underlined that availability and quality of prey is a critical component of any study that attempts to identify whether whalewatching is having an effect on cetacean populations (SC/LAWE08/5). Ultimately, it would be ideal to assess the distribution, quality and availability of prey and the potential abundance of predators. Such information is necessary to determine whether cetacean populations are changing in response to whalewatching activities or due to changes in prey availability or quality (nutritional and/or energy content). These range from year to year variations due to oceanic processes such as the ENSO (El Niño Southern Oscillation), to the multiyear regime shifts associated with the NAO (North Atlantic Oscillation) or the PDO (Pacific Decadal Oscillation). Long lived species such as cetaceans are adapted to these natural variations in prey. Further, study design must take into account the appropriate scale over which the animal operates. Thus measures of the environment should also be taken at the appropriate scale. Finally, any study that examines the potential effects of whalewatching needs to incorporate measures that relate potential changes in prey availability or quality. While prey availability and quality are difficult to measure they are the only way to truly understand the energy budget of an animal. Oceanographic parameters can be used to determine whether the animal's environment has changed or is changing. In some situations these indices may be linked with prey availability, but this linkage will vary across species and habitats.

Examples of such oceanographic indices include, but are not limited to:

Remotely sensed measures	Direct measures	Derived indices
SST (sea surface temp.)	CTD profiles	Upwelling index
SSH (sea surface height)	Chlorophyll profiles	ENSO index
Chlorophyll	MLD (mixed layer depth)	PDO
Frontal structures	Prey abundance and nutritional value	NAO

The Workshop noted that with many species, even good predictors often explain the distribution of only 50% of sightings or less. Further, measuring the prey base can be difficult, and, in the case of detailed feeding studies of North Atlantic right whales (*Eubalaena glacialis*) in Cape Cod Bay, whales appear to leave Cape Cod Bay at peak prey times (S. Mayo, pers. comm. to M. Weinrich). The relevance of these widespread prey measurements on odontocetes is not as clear as it is with mysticetes, where among planktivores there is a clear, simple, low trophic level tie to productivity. This is especially true of epipelagic (surface feeding) species; the relationship is not as strong in benthic (deep) feeding species. However, the distribution of elephant seals, sea lions, and many odontocetes is related to productivity. In Cape fur seals, effects are seen, but with a time lag of months. Certainly, measurements of prey or indices of productivity should be included in multi-variate models investigating life history parameters.

Many remote sensing data sets are available, but collecting sea surface temperature will help considerably in fine-scale analysis (satellite imagery may have scales of up to 20 miles). Spatial scale is very important in experimental design.

The Workshop discussed how to deal with variability in ecological factors and the 'stress' the population can take based on prey availability. In good years, a population may be able to take far more stress from something like tour vessels than in poor years. One could manage for the 'stress' level that a population can take in poor years, or one could establish a baseline to 'buffer' good years against bad years. Managers would like a predictive capacity, but it was acknowledged that such predictability may not be possible.

A clear relationship between the amount of 'stress' in good or bad years and effects on animals has not been established. For instance, in a 'bad' year, a whale calf may be lost regardless, and the effect of whalewatching may not affect the calf survival regardless of how much 'stress' is added. Most of the findings presented come from pinnipeds, and cetaceans are much harder to study. However, unpublished data on D-tagged humpbacks in Greenland have shown that sub-surface feeding rate decreased when a whalewatcher aggressively approached the whale closely. This leads to cumulative effects models, which are commonly used in ecology.

The Workshop noted that the same set of parameters will not be useful for all cetaceans, e.g. mysticetes vs odontocetes, transient vs resident cetaceans, or animals with different life spans. While broad phylogenetic groups (i.e. mysticetes and odontocetes) show very different life history patterns, there is similarly a large variation of patterns within these groups. Ranging patterns and home range size are common ways of grouping animals. In localised species with small home ranges, all of the effects (SST, wind, etc.) are more easily measurable while with large mysticetes, body condition of the animals may be related to environmental conditions, months ago and thousands of km away.

There are many cases where animals that were assumed to have limited movements were found to move over large areas as indicated by satellite tags, e.g. North Atlantic right whales on their feeding grounds (Mate *et al*., 1997), harbour porpoises (*Phocoena phocoena*) in the Gulf of Maine (Read and Gaskin, 1985) or photo-identification studies with long range comparisons (e.g. southern 'resident' killer whales that have been photographed off California). Hence, determining that a population really is limited in spatial and temporal scales is essential in determining impacts.

Measurements of prey and feeding success are important, and if financial or other resources were not limiting the study design, they would certainly be included. There are straightforward methods to do this, but it takes a large effort. However, results can be integrated with swim tracks and movement data to investigate how distribution relates to prey; cost/benefit based on caloric input vs effort can be better measured. Knowing both quantity and quality of prey is important. An animal may get more energy catching fewer prey of the same species at different times of year or in their life cycle (e.g. herring fat content in spring vs. fall, female fish carrying eggs, North Atlantic right whales showing strong preferences for stage 5 *Calanus finmarchicus*). When looking at feeding success of individuals, it is important to develop measures that average energetic input over longer periods (e.g. 30 days) as opposed to one day.

In terrestrial animals (e.g. baboons), the availability of 'weaning foods' is a prime predictor of calf survival, and these may be very different from adult prey (Altmann, 1974). In humpback whales feeding in the Gulf of Maine, years of low sand lance availability have been years of low calf survival (Rosenbaum *et al*., 2002), and young juveniles in the past seven years have been seen consistently feeding during fall on some form of plankton that adults do not generally utilise. Little is known of this phenomenon in cetaceans otherwise. However, juvenile animals do not have the physiological diving capability of adult animals, and may have other limitations as well.

The needs of different sexes also need to be taken into account in an impact assessment or energetic model. Data may often be easier to obtain on males, but females are the ones producing the offspring and on whom variability in life history characteristics may be more easily measured. Female energetic needs will also vary depending on the reproductive status of the individual female (pregnant, lactating or resting).

Ultimately, the goal is to measure food availability and quality for the animals as an explanatory variable in an impact study. Primary productivity can be used as a proxy for this, recognising that the link may or may not be strong. There are parameters that are easy to measure in the field (SST, CTD cast or other ways to measure water column temperature, etc).

Since Shark Bay was one of the areas where there is a strong database on a localised population, it was used as a 'test case' for what variables might be collected in an ideal impact study. These include water depth at each sighting (possibly indicative of both prey and predator abundance), indices of sea grass presence/benthic cover (also informative proxies for prey availability), water temperature (seasonal changes may be related to tiger shark abundance), and shark (predator) abundance itself. In harbour seals, local movements may be based on shifts in prey (herring) habitats. Water temperature affects both prey and predator (both preferences and thermal tolerances of fish) and should be collected. In shallow waters like those found in Shark Bay, measuring primary productivity probably won't measure prey as well as in pelagic areas. However, it may still affect the abundance of fish species, which may spend only a portion of their life cycle in the sea grass beds (in other words, external factors could affect their abundance prior to their visiting the sea grass beds).

The Workshop discussed how much knowledge about an animal's life was needed to make management decisions. It was noted that: (a) it was important to determine if there actually was a real problem from whalewatching that required management, and how it ranked when compared with other problems relating to life history parameters in importance (e.g. bycatch); and (b) to set up a testable hypothesis to determine if a management action, once taken, was successful in meeting its goals. Further, without solid data on which to base a proposed management measure, decisions are hard to propose and enact, and for managers/politicians to defend to their constituencies.

In general, control/experimental areas with data like Shark Bay are rare. However, the idea of looking at whether management measures succeed in their goals of protecting populations of cetaceans sets up an experiment with a testable hypothesis.

4. MODELLING: DATA REQUIREMENTS FOR MODELLING APPROACHES

Robbins presented SC/LAWE08/5 via teleconference and described the use of multi-state mark-recapture modelling to investigate humpback whale movement and fecundity in relation to whalewatching in the Gulf of Maine. Multi-state mark-recapture models estimate transitions between states as well as apparent survival and detection probabilities. A model selection process was used to evaluate hypotheses and parameters were estimated from individual sighting history data using maximum likelihood techniques. Habitat choices (whalewatching areas or not) can be thought of as states that individuals can move between, with possible survival implications. In the Gulf of Maine, this approach is being used to study two commercial whalewatching areas that are adjacent to ecologically similar coastal sites without significant whalewatching effort. The direction and magnitude of movement within paired areas can be used to investigate potential avoidance of whalewatching activities, while area-specific survival estimates provide insight into a potential cost of those decisions. Multi-state models also provide a robust framework for investigating fecundity effects, as reproduction can be modelled as a transition between calving and non-calving states. Exposure information, as well as other individual and ecological covariate data can be incorporated in a linear modelling framework to better understand observed patterns. In the Gulf of Maine, humpback whales that have low apparent exposure at one whalewatching site can nevertheless have significant exposure elsewhere. Robbins therefore emphasised the importance of a synoptic view in the development of biological models and covariate data. The results of this study will be presented to the Scientific Committee at IWC/60.

In discussion, it was noted that in the two whalewatching areas in the study (Stellwagen Bank and the Bay of Fundy) prey sources are different. This is actually a benefit to the hypothesis being tested by the model because if the effects of whalewatching were to be real, whales would do similar things in the two areas despite different ecological regimes and prey bases.

Sherwin presented SC/LAWE08/6. 'Conservation forecasting' aims to predict growth rate, extinction probability, etc, considering the effect of natural variation, management and threat scenarios, on biological factors, physical factors and intrinsic factors such as genetic variation and behaviour. Tourism may impact cetacean behaviour sufficiently to alter population forecasts. For the focal species and interacting species and processes, it is important to establish the relevant spatial and temporal scales of effects. For Shark Bay *Tursiops*, the spatial scale was identified by genetic studies, which resulted in modelling the population as two subpopulations whose long-term interchange rate was estimated from genetic data; a method that averaged dispersal over multiple generations was chosen as appropriate for the long-term (300yr) forecasts. Data from a variety of studies over two decades were used to estimate mortality, reproduction, and other input data for the program VORTEX. The population size was forecast to be stable, which agrees with long-term observations. Potential threats to this population include fishery-related mortality, a virus epidemic, pollutant spills (e.g. oil, sewage), boat traffic, and underwater obstructions. Data from other species showed that only the most extreme fisheries-mortality (7% increase) resulted in substantial population decline, and that viral infection causing 3% additional mortality was unlikely to seriously deplete the population.

Simulations showed that the population was most sensitive to reproduction, with mortality at certain age-classes also being important. In related populations and species, mortality is the most important for population forecasts. Possible reasons for this contrast are: (i) some replicates showed severe stochastic depressions of population size from which the population did not recover easily; (ii) mortality rates were very low for post-weaning juveniles, for which rates are not known in most other studies; (iii) some systematic difference between this population and others, such as the extensive seagrass and limiting salinity; and (iii) no data were available on density dependence, so were not able to model this except as a ceiling carrying capacity, but if reproduction is elevated at low densities, this could reduce its overall importance for forecasts.

The sensitivity to reproduction becomes very important when considering tourism impacts, because SC/LAWE08/2 shows that tourism activity can depress reproduction by as much as 50%. If such an effect occurred throughout the bay, the model shows that the population is likely to be extinct in about 50 years. At present, tourism is limited by the Western Australia government. Additionally, there are differences in data censoring, etc, between the study presented in SC/LAWE08/6 and the studies of Bejder and Lusseau, which might have a significant effect on forecasts, and these questions will be investigated. The current limited nature of tourism in Shark Bay makes an ideal situation for management experiments, if regulation gives forethought to experimental design and baseline data.

Tourism might have genetic effects on cetacean populations, but only under quite restricted conditions. There could be loss of genetic variation if the population size is severely depressed, but in this case demographic concerns would be paramount. Tourism might increase erosion of genetic variation without noticeable reduction of population size, for example if only one sex is affected. Major artificial subdivision caused by tourism could also affect genetic variation. If any of these situations are suspected, management experiments could include appropriate genetic data to gauge these effects.

In summary, although the authors of SC/LAWE08/6 cannot forecast exact probabilities of extinction or growth rate, management/threat scenarios can be compared and important parts of life cycle for management and monitoring (e.g. reproduction) identified, and a dataset exists that can be extrapolated for forecasting exercises in other populations, but probably not other species.

In discussion, it was noted that the models discussed can be broken into areas within habitats, to examine areas affected by tourism activities, taking into account immigration and emigration. Density-dependence was not included in the model, except as a maximum in carrying capacity. Because the population has never shown major fluctuations it has not been possible to estimate the effect of density on vital rates.

In other PVAs of bottlenose dolphins, the risk of extinction is most sensitive to variation in adult mortality. One way the importance of reproduction could be overstated is if the survival rates used in the model were very low, and all of the variance comes from reproduction. Mortality rates in the VORTEX model presented for ages 1, 2, and 3 did appear very low, and were derived from a survival analysis over multiple three-year periods. There may be an effect from not being able to identify calves by dorsal fin, but that would actually result in a bias towards higher mortality.

5. CASE STUDIES AND STUDY SITE FEASIBILITY: PRACTICAL ASPECTS TO CONSIDER

Hevia presented SC/LAWE08/11. Argentina has a 35-year history of commercial whalewatching, which started at Península Valdés targeting Southern right whales (*Eubalaena australis*). Nowadays commercial whalewatching has grown and is being developed in five places in Patagonia: Las Grutas (Río Negro Province) where the activity is incipient, Península Valdés and Playa Unión (Chubut Province) and Puerto Deseado and Puerto San Julián (Santa Cruz Province). Península Valdés is the main one regarding the number of tourists. The species primarily targeted by whalewatching are the Southern right whale, Commerson's dolphin (*Cephalorhynchus commersonii*), dusky dolphin (*Lagenorhynchus obscurus*) and the killer whale.

Whalewatching is boat-based at all sites, but in Chubut Province land-based whalewatching also occurs (an example of this is killer whale watching in Punta Norte, Península Valdés). There are regulations for whalewatching for Southern right whales, but so far there is a lack of regulation for dolphin watching. Authorities have the will to work with scientists (from Universities and NGOs), along with whalewatching operators, on the development and updating of regulations.

There are scientific studies of the species targeted by whalewatching at all sites, the degree of which varies between different sites and species. Not all sites have studies on the potential effects of whalewatching and for those that do have, the studies analyse short-term effects (examples of these are Commerson's dolphins in Chubut province, (Coscarella *et al*., 2003)) and the same species in Santa Cruz province (Failla *et al*., 2004). In general, there is a need to update existing studies, analyse other variables and develop long-term studies.

In discussion, it was stated that the highest density of whalewatching in Argentina is from Peninsula Valdés. Six companies operate from Golfo Nuevo from May through December, with a peak from September through November. There is also a nearby control area where whalewatching is prohibited (Golfo San Jose). The diversity of habitats, species, varied history of the industry (from well established to very new), and cooperation and interest of the government suggest that this area should be considered for inclusion in a global study. It is unlikely that whalewatching would cease in any of these areas (especially in the well developed industry of Peninsula Valdés), although the future of whalewatching in Las Grutas is hard to predict (right whales have only increased their presence recently). Local researchers come from a variety of local universities and NGOs.

One important criteria in selecting sites is having local universities and NGOs that can participate and invest in a long-term project. It was pointed out that in this site NGO photo-ID data goes back to 1971 for southern right whales. The data from studies in Argentina currently lie with the researchers or institutions that collected them, and their availability to a larger study like this is unknown.

Groch presented SC/LAWE08/12. The Right Whale Environmental Protection Area (EPA) was created on 14 September 2000, in accordance with a specific recommendation of the 1998 IWC SC's special meeting on right whales, aimed at ensuring the protection of the main concentration area of a wintering ground for southern right whales off Brazil. Whalewatching activities have been conducted in this region since 1999 and boats have operated in agreement to the national legislation. An increasing interest in this activity has been recently observed in the EPA which could put at risk the protection of these whales if not properly managed.

Preliminary results from monitoring the activity suggest that during most of the whalewatching cruises, right whales appeared to have ignored the presence of the boats and did not interrupt their behaviour, so that no clear evidence of immediate disturbance to this population was detected.

Despite available information suggesting that the boat-based whalewatching activities are not disrupting the behaviour of right whales in the EPA, available data comes from preliminary information obtained from a small industry which operates relatively few whalewatching cruises, and therefore long term effects are unknown.

Taking into consideration the results and recommendations of the intersessional Workshop on Whalewatching held in Cape Town in March 2004, the authors participated in the development of a proposal to establish area closures in the EPA. Besides the scientific desirability of having control areas implemented where no boat-based whalewatching activity occurs, other management issues were taken into account to propose the specific closure areas, including:

- shore morphology of proposed closure areas;
- topography of adjacent shore observation points; and
- boat-based tourism use of proposed closure areas.

Hence, the selection of proposed sites was defined to include six areas. Closure would apply in these areas to all boat-based whalewatching activities and also to the use of any motorised watercraft for recreation. With the exception of one beach, all proposed closure areas are regularly monitored from shore during the whale season through land-based research techniques. In addition, eight other areas where boat-based whalewatching occurs and would continue to occur are also monitored in the same fashion, thus allowing for comparisons to be made regarding whale behaviour in relation to boats.

The closure areas were adopted by a Normative Instruction in June 2006. It is hoped that the adoption and enforcement of these closure areas, and their future incorporation in the Protected Area Management Plan (currently in its initial stages of development), will allow for further improvement in gathering knowledge about the short- and long-term effects of boat-based tourism on calving southern right whales in Brazil, and in the design and implementation of adequate management measures to ensure both the species' survival and the sustainability of the whalewatching industry.

In discussion, it was noted that control bays were approximately equal in size to some whalewatch bays, but whether the habitats are of 'equal' quality is unknown. There is no current set period for which the 'no whalewatching' areas will be instituted, and it is an 'instruction' as opposed to a law. It is important that the sites chosen in a study have some government assurance that they will be set-aside for a lengthy period. It is also unknown whether the whales mix evenly throughout the area or whether there is individual preference. In the latter case, a greater opportunity to examine 'high-stress' vs. 'low-stress' whales would occur.

There was discussion about how widespread the problem of 'eco-tours' not being bound by whalewatching regulations was (as appeared to be the case in New Zealand and in some other places). It was noted that in Brazil, national regulations apply to all boats.

Carlson, participating by phone, presented an Excel spreadsheet created to assist in the assessment of site potential using a suite of categories. The exercise is a work in progress. It was suggested that the spreadsheet be used to facilitate discussion on the practical aspects of site selection and that this aspect of the work could be formula based. All agreed practical aspects (e.g. geographical accessibility, political climate, whalewatch history, species status, habitat use, research status, and external variables) would be listed and weighted as to importance for the success of the project and its import to the population/area. Once developed, such a formula could expedite the final process of site selection.

6. STUDY DESIGN: PRIORITATION OF VARIABLES TO SAMPLE

Wasser presented SC/LAWE08/7. Non-invasive physiologic and genetic measures from scat are now being used to characterise the breadth of human impacts on wildlife. A key strength of this approach stems from the fact that individual disturbances rarely occur in isolation. When one disturbance occurs, many others tend to follow. It thus becomes critical to attempt to partition their relative impacts in order to better guide mitigation efforts. The ability to acquire multiple physiological and genetic measures from the same scat sample makes this non-invasive approach well suited for dealing with such problems.

The Center for Conservation Biology, University of Washington, focuses on non-invasive physiological and genetic measures in scat because a large number of physiological and genetic products are shed in scat and it is perhaps the most accessible wildlife product found in nature. A variety of physiological and genetic measures from faeces have been well validated for the North Atlantic right whale and killer whales. Faecal endocrine measures can be used to assess reproductive condition and failure, as well as physiological and nutritional stress. Diet, immunoglobulins, parasite load, and circulating toxins can also be measured in these scat samples. Collectively, these measures can provide an overall health profile of the animal from a single faecal sample. DNA from these same samples can also confirm the species, gender and individual identity of the animal, if needed. These methods were further strengthened by developing novel methods to collect these scat samples in free-ranging whales. Detection dogs were trained to locate whale scat floating in the water, as a means to enhance samples sizes. These dogs can easily increase sampling success by 4-5 fold over alternative methods, strengthening the power of analyses aimed at assessing relative impacts of multiple disturbances (Rolland *et al*., 2006).

The use of these tools on studies conducted on southern resident killer whales in Puget Sound, Washington, was highlighted, where attempts are being made to partition impacts of whalewatching, declining prey and toxin loads, each postulated to have contributed to their 20% decline between 1996 and 2001. Declining prey is hypothesised to have a major effect on killer whale nutrition. Poor nutritional status may also cause animals to mobilise fat, releasing stored toxins back into the circulation. Whalewatching boats may compound these problems by decreasing foraging success, further impacting these direct and indirect nutrition impacts when prey availability is low. Boats may also add their own stress-related impacts on the study animal. Preliminary physiological data from scat demonstrated both stress impacts of boat traffic that may vary by type of boat, as well as nutritional impacts associated with pod-specific prey availability. Collectively, these preliminary data illustrate the incredible potential of these combined methods to address the problem of whalewatching in the complex environment in which it occurs.

In discussion, it was noted that prey in scat is determined both through identifying components in faecal contents and through genetic identification. When examining pollutants in scat, it was suggested that some pollutants could come directly from prey. However, that amount is probably negligible. In many cases, validation and control studies are done using numerous techniques.

The presentation of data that showed increased glucocorticoids in killer whales on weekends elicited discussion. When mean levels were broken down by day of the week, a substantial decline occurred on Thursday, followed by a progressive rise from Friday to Sunday, suggesting that whales took several days to 'calm down' after the weekend. Others questioned whether the Thursday data could have reflected an anomalous day of the week. This is a preliminary study, which also showed that glucocorticoid hormone increased in response to whalewatch boats as such, not just the absolute number of boats on the water.

In southern resident killer whales, 'L-pod' also shows lowest levels of T3, while 'J-pod' (the supposed dominant pod of the three in the area) has higher levels; higher T3 indicates a better food supply. J pod's T3 levels are similar to those of provisioned animals at Sea World. The decline in the southern resident community has also been disproportionately seen in the L pod. J pod spends most of its time around the San Juan Islands, while L pod has been seen as far south as California. The stress response is an adaptive mechanism. However, long term chronic stress is well documented to cause a variety of health problems. Using this in complement with other types of measures can allow real insight into health, nutritional status, reproductive status, and other factors. It is a good complement to population data from accompanying field studies, and can link individual behaviour to its underlying physiological consequences. It is also important because this is an integrating, rather than a point, measure. Faeces can actually provide a better measure of chronic stress than blood. Hormones are secreted in blood in pulses approximately every 50 minutes. Hormones in faeces reflect the last 24 hours of cumulative pulses. As sample size increases, other factors that could affect stress levels will be evened out, and in experimental design we can control for individual factors that we can compare with this integrative measure.

Measures of hormones are expressed per gram of faeces dry weight. Extensive validation studies have shown that using dry weight controls for dietary impacts on hormone excretion. The faeces are treated before assessment to remove hard parts, which would eliminate chitinous shells dumped by whale species eating euphasiids or copepods. Stress could be measured from hormones in exhalations, but the stress of approaching the animal so closely to collect blow contents, combined with the rapidity of the approach itself (as little as three minutes) will affect the level of stress hormones in the blow. In highly habituated populations, collecting hormones from exhalations may be feasible. Some additional hormones (e.g. LH) may also be uniquely present in exhalations. Some scats float more than others, and there was a question as to whether this might be diet dependant. Since with killer whales animals are presumed to all be feeding on the same thing at any given time, this should not be a major factor. However, there are situations where such variability should be examined. The Workshop **agreed** that the technique is very promising and **recommended** it be incorporated into impact studies.

Noren presented SC/LAWE08/8. A key component in the study of the ecology of animals is the study of their physiology, and in particular bioenergetics. The two key components of bioenergetics is energy acquisition in the form of prey consumption and energy expenditure due to maintenance metabolism, growth, reproduction, and daily behaviour. Disturbance from vessels has the potential to impact both energetic expenditure as well energy acquisition in cetaceans. Energetic expenditure can be increased in response to vessels if cetaceans respond to vessels via evasive tactics (e.g. increasing swimming speeds) and/or performing percussive behaviours (e.g. tail slaps, breaches, etc.). Furthermore, energy acquisition can be decreased if cetaceans respond to vessels by halting foraging behaviour. Because the behavioural responses of whales to vessels are associated with potential energetic impacts, it is important to incorporate bioenergetic studies in any investigation of vessel impacts on large whales.

Recommendations from the paper included the following.

(1) Record surface swimming speeds and respiration rates of whales (include juveniles, adult females and adult males) during periods with and without vessel presence.
(2) Record behaviour states (forage, travel, rest, etc.) of whales during periods with and without vessel presence.
(3) If possible, record surface swimming speeds, respiration rates, and behaviour states of whales that are also tagged with time-depth recorders during periods with and without vessel presence.
(4) Determine feeding rates/prey consumption with and without vessel presence. This should be done during expected periods of low feeding rates (e.g. breeding season) as well as during expected periods of high feeding rates.
(5) Measure body morphometric (length to width ratio) changes over the feeding and non-feeding periods across individuals representing all age- and sex-classes to assess changes in fat content via aerial photography. Use these data to assess changes in body condition and calculate rough estimates for energy acquisition and depletion.

There was discussion as to why female killer whales with calves had higher respiration rates at any given speed than those without calves. There are several possibilities for this, including extra costs in echelon swimming with their calves (Noren, 2008), lactation costs, or the possibility that this apparent effect may be an artefact of the tendency for mothers to surface synchronously with their dependent calves. An apparent increase in respiration rate does not necessarily indicate increased oxygen consumption because tidal volume (volume of air per breath) may not be constant. One key finding to emerge from the study is that as killer whales increased swimming speed from 1.6 to 2.8 m s^{-1}, the corresponding increase in energetic cost was relatively small. Consequently, if killer whales were to respond to boats by increasing their swimming speed slightly, one should not assume that this behavioural response alone would carry large energetic costs. Killer whales may be able to compensate for small energetic costs of boat avoidance by increasing prey intake. However, a previous study has actually demonstrated the opposite effect, namely that one behavioural response of killer whales to repeated disturbance is a reduction in time spent feeding (Williams *et al*., 2006). Animal survival hinges on maintaining a balance between energetic expenditure and acquisition, and close attention should be paid to potential reductions in prey consumption. It was noted that, in fin whales, energy for transport has been hypothesised to be substantially aided by ocean swells, suggesting that the environment may have a strong influence on how travel speeds translate to energetic needs (Bose and Lien, 1989).

Energy budgets are needed over prolonged periods and for different situations; for instance, gray whales, which need to intake their annual energy budget in only a few months, may have greater costs to lost opportunities than in other species or in other areas. D-tag data may be very helpful in refining and furthering this work.

Madsen presented SC/LAWE08/9. Cetaceans using sound for acquiring and transmitting information underwater are susceptible to noise in ways that may have consequences for overall fitness through masking, behavioural disruption or even direct physical damage. The basic, but untested assumption is that noise from whalewatching vessels is a dominant source for the observed and assumed negative effects of whalewatching. Noise from vessels is made up of (a) splashing noise from the hull moving in water, (b) engine noise and release of exhaust gases underwater, (c) propeller noise via cavitation and (d) transient noise from gear shifts. Recent studies of smaller vessels with 4-stroke outboard engines show that gear-shifts may generate transient sounds with source levels close to what may raise concern for temporary threshold shifts in cetacean auditory systems. But, the main concerns pertaining to the noise impact of whalewatching vessels relate to masking and behavioural disruption. Behavioural effects present the most challenging topic in terms of establishing whether sound exposures elicit significant behavioural responses with consequences for animal fitness. However, masking of communication and echolocation signals can readily be evaluated via the range reduction factor elicited by a vessel induced increase in ambient noise in the frequency range of the used signal. Recent research shows that speeds below what generates cavitation on propellers will dramatically reduce the masking effects of a vessel near cetaceans. Currently, there are very few data on the causal links between noise exposure and the observed short-term and long-term effects of whalewatching vessels around cetaceans. Given that noise is the primary vehicle for eliciting negative effects of whalewatching boats around cetaceans, there is every reason to believe that the problem can be reduced significantly by reducing the noise outputs from the vessels. For example, if the noise source level of a whalewatching vessel is reduced by 20dB, the volume of water around that vessel in which cetaceans potentially may suffer from adverse effects, will be reduced by a factor of 100 to 1,000 times. Therefore, many of the concerns regarding the negative effects of whalewatching related to noise can be reduced dramatically, simply by reducing vessel noise and keeping a distance.

Masking can take place all the way out to (and actually slightly beyond) the detection limit. Entire ocean basins around the world have been acoustically mapped, but most of the data are classified military information; there is limited information otherwise. However, there are some snapshot and representational area habitats that have been mapped for background acoustic levels where the masking potential can be assessed.

In measuring gear shifts in outboard engines, there are times where the gear shift approaches the received level of 200 dB re 1uPa (pp). This approaches levels of temporary threshold shifts, if the exposure takes place sufficiently often. Therefore sound exposure index over time for an animal should be calculated.

The Workshop discussed the problem of small sample sizes and noted that the best way to increase statistical power is to test a well-stated hypothesis in several small independent experiments. In any event, the Workshop **agreed** that well-stated hypotheses combined with a robust experimental design were critically important in the project's success. This will be particularly important where effects are subtle.

Lusseau presented SC/LAWE08/10, which introduced new methods to show how information on the temporal dynamics of behavioural state data can be used to infer the resilience of a particular population to behavioural disturbances. The temporal patterns in observed sequences of behavioural states can be described using a Markov process where transition probabilities to pass from one state to another are approximated from observed transitions. Markov chains can be used to understand the temporal dynamics of behavioural sequences and infer properties about the behavioural budget of the observed populations. This study shows that the resilience of behaviour is related to its unpredictability using simulations and empirical data collected over ten study sites representing over 30 years of data cumulatively. The more predictable behaviour is, the less resilient it is. However, such influences on behavioural resilience cannot be related to the magnitude of disturbance effects in inter-population comparisons. Such measures are meaningfully related to the influence of disturbances when comparing the same population exposed to different ecological conditions. Behavioural state sequencing is more resilient to disturbance, that is the effect of similar disturbances is less when the behaviour of the population is less predictable. This behaviour predictability can be driven by ecological conditions. For example in the northern resident killer whale population an increase in food availability is related to an increase in the duration of foraging bouts, hence constraining the dynamics of the population's behaviour. Such constraints increase the behavioural predictability of the population and in turn weaken its resilience to disturbance. This empirically-driven theoretical study offers a framework to manage exposure of animal populations to disturbance.

It was clarified that resilience in this study is measuring the time it takes a system to return to its normal state once it is disturbed. However, a system's resilience and the extent of disturbance vary on a sliding scale depending on the 'starting state' of the system (i.e. whether it is at a pristine state or if it is already disturbed). However, whether a system responds to a perturbation at all can depend on the inertia of the system; you have to exclude changes in thresholds at which inertia is overcome in order to make the claim that the system's resilience has changed.

It was questioned what happens when the Markov chain measures first-order changes, but the state of the system and its transition probabilities goes beyond the first order. This has been compensated for in controls, but not as much in experimental situations. One can also compensate for this by using second- and third-order Markov chains.

Underwood presented SC/LAWE08/13 on different aspects of experimental and sampling design. It is important to consider what elements are necessary or appropriate for a 'perfect' design, however unrealistic its implementation. That way, what is not possible because of logistical constraints can be identified and the influence on inferences (and their robustness) of losing the relevant information can then be made explicit. Some of the general issues are as follows.

(1) Types of impacts

Design of any study to assess potential impacts is dependent on identifying what sort of impacts may be occurring. The differences between 'press' effects (long-term changes in state of some variable(s)) and 'pulse' impacts (short-term changes to temporal patterns of change of variables) make it necessary to have designs that can detect and estimate magnitudes of either or both of these. There are also impacts that alter temporal variances of relevant variables, without necessarily altering mean values of the variables of interest. Designs must therefore be adaptable – i.e. no simple BACI design can efficiently detect all types of impacts.

(2) Various relevant designs

There have been developments over the last 30 or so years in designs of studies that have a chance of detecting impacts. Note that the emphasis is on design-based approaches, not model-based approaches, which were not considered. Using multiple times of sampling (to ensure temporal variance is estimated) *and* multiple locations (to ensure that spatial variance and non-impact interactions are estimated) are equally important considerations.

(3) Multivariate analyses

Many sets of data relevant to environmental disturbances consist of a suite of variables. This is also commonly true for behavioural studies. Consequently, each sample-unit (replicate) is an array of values. Techniques for analysis of such data are available (including factorial, 'analysis of variance-type' designs for simple models and correlation/regression procedures). One important step in multivariate analyses is to determine which variables are actually contributing to any variation among samples and which are not. If two variables are highly correlated, you don't need to continue to gather information on both of them. It is therefore more efficient to use analyses that can help reduce the number of variables. Non-parametric versions of these procedures are robust for all sorts of mixed data (continuous, discrete, etc.). New developments include all sorts of designs, but there are some potential problems with these. Many standard designs (particularly parametric designs) have very restrictive assumptions about the data and the relationships among variables.

(4) Power of tests

There was discussion about the importance in impact studies of not making type II errors (accepting the null hypotheses of no impact when there really is an influence). β, the likelihood of a type II error, can be as large as 0.8 (in other words, there is an 80% chance of making a type II error). This is a problem where sample sizes are small. Increasing sample size increases power. In contrast, where variances are large, power is relatively small, but the experimenter has no control over that. Variances are properties of the variable being sampled. The third thing that affects power is the effect size, i.e. how much difference is hypothesised to exist among the treatments. Nevertheless, if the level of α is changed (e.g. from α=0.05 to 0.10) power increases (i.e. the chance of making a type II error decreases). In multivariate analyses, there is no straightforward definition of an 'effect size.'

For frequency analyses, there has been a lot of work on how to determine the power of a test (its capacity to reject a null hypothesis when some pre-determined alternative is actually operating). Attention to the power of tests requires good understanding of 'effect-sizes' (the predicted differences among the treatments being compared), some estimates of the variances of the variables being tested and trade-offs between sizes of samples and probability of Type I error (or 'alpha'). Power for many types of impacts cannot be determined analytically, but can be done by appropriate modelling of a hypothesised impact.

(5) Bio-equivalence

In analyses of impacts (or other responses to disturbances), attention often needs to be paid to recovery after management or mitigation has been implemented. 'Traditionally', this was done by testing the null hypothesis that the disturbed values of a variable have changed so that there is no difference from the values of the variable in the reference or control areas. Bio-equivalence, in contrast, defines a value of the reference variable that is considered to be recovered. The hypothesis tested is then that the values of the variables in the disturbed areas exceed this defined reference. This reverses the onus of proof and avoids the problems of Type II error that plague more traditional approaches.

In discussion, it was noted that non-parametric tests are useful in some cases, but are typically less powerful than the equivalent parametric tests. Further, assumptions are often not that different between parametric and non-parametric tests, except that the former often assume that data are normally distributed (e.g. Kruskal-Wallis tests vs. ANOVAs). ANOVAs are very good for power functions or exploring the sources of variation. GLMs, however, are used as alternatives where errors are non-normally distributed or when variables are a combination of ordinal and categorical data.

Williams presented a summary of results of a number of vessel-interaction studies that have been undertaken on northern resident killer whales (British Columbia, Canada) since 1995. That project has used the same land-based observation site from which behaviour of whales was measured in summer in the presence and absence of boats, both inside and adjacent to Robson Bight (Michael Bigg) Ecological Reserve in Johnstone Strait. The study has used a combination of small-scale experiments on individually recognisable focal animals, opportunistic observations (under conditions when it was unfeasible to manipulate traffic around whales), and scan sampling of activities of boats and whales. Over the years, the experimental design has compared behavioural responses of individuals during control periods and three different treatment levels. The first experiment showed that animals responded to experimental approach by one small boat following whalewatching guidelines (paralleling whales' paths at 100m) by adopting horizontal avoidance tactics in the form of less direct swimming paths (Williams *et al*., 2002b). Opportunistic observations suggested that behavioural patterns showed opposite relationships as boats got closer and as boat numbers increased. The second experiment (Williams *et al*., 2002a) considered an erratic approach by a 'leapfrogging' vessel, and found that whales displayed the same stereotypical evasive tactics with stronger effect size than those observed when the experimental boat followed whalewatching guidelines. The third set of experiments measured response of whales to approach by multiple vessels (Williams and Ashe, 2007). Results showed that the stereotypical evasive tactics previously observed in the population were only observed when few (1, 2 or 3) boats approached whales within 1,000m. Whales approached by 4-17 boats adopted more predictable swimming paths. The average of these two opposing behavioural responses (i.e. a simple presence/absence comparison) would have led to a false conclusion, with very strong statistical confidence, that no effect took place. Opportunistic observations showed a similar curve between boat number and path directness after attempting to account for confounding effects, with statistical support for an inflection point in the curve occurring when three boats were found within 1000m of the focal whale. Scan sampling of activities of all whales observed in the study area during nine summers (June-September, 1995-2003) showed that animals spent less time feeding in the presence of boats than in their absence (Williams *et al*., 2006). While the relationship between time spent feeding and consumption of prey is not straightforward, these analyses suggested that the energetic cost to whales of reduced energy acquisition had the potential to be 5-6 times greater than that of increased energetic expenditure due to avoiding boats. This assessment of relative energetic costs is important for prioritising future research efforts, because impacts of whalewatching on time-activity budgets have the potential to cause fitness or population-level consequences. These data were also used to demonstrate that the grouping behaviour of northern resident killer whales was influenced by relative abundance of Chinook salmon in Johnstone Strait (Lusseau *et al*., 2004). This ecological context suggests that inter-annual variability in prey density may buffer the whales' ability to cope with anthropogenic disturbance, and confound our ability to measure disturbance. Finally, results were shown from a systematic line transect survey in which average summertime density of resident killer whales in Johnstone Strait was estimated to be 40 times higher in Johnstone Strait than in BC coastal waters overall (Williams and Thomas, 2007). Given that the core whalewatching area for this population targets the seasonal aggregation of whales in Johnstone Strait, there is the potential for boat traffic to be contributing to general habitat degradation in a very small but critical component of the overall range of this vulnerable population.

In discussion, it was suggested that for some species, such as a deep diving toothed whale like a beaked whale, if a dive is shortened, the only option the animal has is to reduce feeding time (at the base of the dive). Hence, even though the total dive may be cut by a small percentage, the reduction in foraging time may be far greater. Williams noted that his variable for 'travel' is really also foraging time; the animals are likely doing both simultaneously. In general, how time categories are defined (e.g. when is foraging really foraging?) is both hard to do and important in study design and conclusions. Further, the energetic 'cushion' the animal has to make up from these boat avoidance consequences is unknown, and these effects really are most important when a population is food limited.

It was noted that a comparison of the ratio of time in each behavioural state rather than an absolute, would be useful, i.e. a decrease in time from 80% to 60% of time is the same ratio as 40% to 30% of time. Finally, it was noted that while the study area Williams used, and the areas where the southern residents were studied, are so heavily used that controls are impossible, there are several areas along the coast where whalewatch boats are less frequent, that are not widely studied, and may serve as a control in a large scale study.

Trites presented a series of cautions in marine mammal science. Intensive studies have occurred in the North Pacific to determine whether the decline of Steller sea lions was caused by disturbance (commercial fishing). Trying to resolve this question has cost about US $200 million and has resulted in a lack of consensus among scientists. A number of lessons can be learned from the concerted research programme that has been underway in the North Pacific for 15 years that may help guide the design of future studies seeking to understand the effects of human activities (e.g. whalewatching) on marine mammals. In particular, it is important to:

Know your study animal. Interpreting what is normal or abnormal cannot be done without knowing the timing of major life history events (weaning, births, mating, etc.), seasonal energy requirements, movement/migration patterns, and diurnal behaviour (time spent foraging, resting, socialising).

Don't accept common wisdom about your animal without data to back it up. Some of what is believed to be common knowledge about a species is little more than fanciful 'stories' that may mask key aspects of life history needed to understand how species respond to disturbance. Check your facts before quoting fiction.

Think about how marine mammals evolved. What are the natural behaviours they have evolved to survive? Knowing the range of conditions that an animal is adapted to contend with can help to understand the range of anthropogenic conditions an animal might or might not be able to handle.

Results that are consistent with expectations should be scrutinised more carefully. Be on the lookout for 'unquestioned answers'. People typically accept results that fit their expectations or preconceived ideas about how they think the world works. Such results are usually the ones that need to be scrutinised more carefully.

Be cautious of 'one-way stories'. Correlations are all too often based on short time series of one-way events (e.g. a species declined while a fishery increased). Causation is usually inferred from correlations that fit preconceived ideas. Longer time series are needed containing observations where the supposed threat is removed or reversed to ensure that the correlation stands up.

The system you see is not the real system. Ecosystems are dynamic and in constant flux. They represent a continuum of change and different steady states, making it difficult to know what is normal and where one is currently along this continuum. Carrying capacities change, yet many biologist hold a 'Garden of Eden' view of the world despite the fact that species have never all co-existed at high levels of abundance in the absence of humans. There is no such thing as a pristine ecosystem.

Prey quality and quantity. Food needs to be thought of in terms of *quantity* and *quality* of prey. Many people falsely believe that marine mammals will do well as long as there are fish in the ocean because they are opportunistic feeders. However, mammals are limited by the size of their stomachs and time needed to capture and digest food. Food consumption can vary by an order of magnitude depending upon the caloric density of prey eaten, which in turn varies seasonally between and within species.

Energy requirements of marine mammals change seasonally. The ability to handle stress also likely varies by season. Males and females may also have different seasonal responses to disturbance. One needs to account for seasonality to properly interpret data.

Predation. The risk of being eaten drives a lot of observed behaviour. Yet the risk of predation is not widely acknowledged or considered when field observations are interpreted. Substantial insight into marine mammal behaviour may be gained by estimating how vulnerable animals are to being eaten when in different habitats at different times of day.

Assessing long-term response to disturbance requires multi-factorial data to test alternative hypotheses. Factors that might affect how animals respond to disturbance include the quality of the occupied habitat, the distance, availability, and quality of other sites, the risk of predation, density of competitors, or the investment that an individual or group has made in a site. Responses may also be specific to an individual, or may occur at a small group or population level. The types of data needed to assess long-term responses and test alternative hypotheses may include: diet (quantity and quality), oceanographic conditions, predation, birth rates, survival rates, age structure, activity budgets, habitat quality, body condition, etc.

Successfully resolving multi-factorial problems requires increasing multidisciplinary inputs (oceanographers, fisheries biologists, biostatisticians, archaeologists, etc.). Marine mammals are part of an ecosystem and are not independent of it. Understanding an ecosystem and the marine mammals that live within it requires interdisciplinary expertise.

Detecting disturbance effects likely requires significantly increasing sample sizes. Individual variability is larger than most people recognise. A small number of observations will likely result in concluding there is no effect of disturbance unless there is a large effect. Multivariate analysis can help work with smaller more realistic samples, as can repeated-measures study designs.

Ultimately physiological effects need to be linked to life history consequences. The most logical way to do this is by constructing mathematical models (that articulate conceptual models about how we think species are affected by disturbance) and to later test them with field data.

The nature of mathematic models. Mathematical models help to articulate conceptual models of understanding, organise existing knowledge, and identify missing information and key parameters. They also help to anticipate the unexpected by allowing different actions to be simulated. But all models are partial truths and people need to appreciate the limitations of model predictions. Models need to be validated by constructing independent models and they should be tested with independently collected data.

Measuring short-term response to disturbance is easier than assessing recovery. Recovery is generally recognised as a return to an original state or normal condition, but is more often operationally defined as a return to 50% or more of pre-disturbance numbers or behaviours. It is important to establish the criteria for concluding that recovery has occurred.

It is essential to collect data that will allow you to discriminate similar outcomes from different mechanisms. Simple interpretation of disturbance effects can be easily confounded by concurrent natural seasonal changes in behaviours, or by daily variability in numbers and behaviours that can be attributed to weather, tidal cycle stage, and other factors. There may be more than one possible explanation for a given observation. One must therefore recognise the different mechanisms and be able to rule those out that are inconsequential.

Most field sites thought to be experimental controls are likely little better than pseudo-controls. Marine systems are complex. It is not possible to perfectly pair sites for comparison. Researchers will therefore probably have to rely on 'Before-After' experiments as their primary tool to investigate disturbance effects.

Scale matters when selecting experimental sites. Study areas should represent significant portions of a species or individual's range or life history event to ensure that observations are meaningful and can be generalised.

Importance of individual recognition. Some of the biggest advancements in understanding marine mammals have come from being able to recognise individual animals. Branding, tagging, or natural markings make it possible to determine birth rates, survival rates and movement patterns of animals and provide key pieces of information needed to understand whether or not animals respond to anthropogenic stimuli.

Weight of evidence. It is theoretically possible to draw a consensus from a range of studies using a preponderance of evidence approach. However, issues pertaining to marine mammals are unlikely to be evaluated objectively and will likely be influenced by people's personal belief systems. Evaluating scientific findings is often not objective, and will likely result in a lack of consensus over issues that are emotionally charged.

Resilience and unexpected changes. Species have great resilience and are likely to respond in unpredictable ways to punctuated events. Marine ecosystems are products of a physical environment which has never been constant on a long time scale, but is always in flux. Change is constant.

Additional information and papers from the University of British Columbia Marine Mammal Research Unit are available at *http://www.marinemammal.org*.

7. PROPOSAL OF LARGE SCALE STUDY: LAWE

The Workshop noted that the Scientific Committee has agreed that there is new compelling evidence that the fitness of individual odontocetes repeatedly exposed to cetacean watching vessel traffic can be compromised and that this can lead to population effects. The Workshop therefore **agreed** that the overarching aim of the proposed research should be to establish if there are population level effects in odontocetes exposed to prevailing whalewatching, and whether such effects also exist in mysticetes exposed to whalewatching.

The Workshop recognised that, although not comprehensive, cetaceans targeted by whalewatching can be roughly divided into four categories: (1) resident populations where breeding, nursing, and feeding occur in the same area; (2) cetaceans on their breeding grounds; (3) cetaceans on their feeding grounds; and (4) cetaceans on their migratory corridors.

The Workshop considered the application of the US National Academy of Sciences Population Consequences of Acoustic Disturbance (PCAD) model as a framework for assessing the potential for population-level impacts of whalewatching (Annex D). While the model presents certain challenges in practical application, the Workshop **agreed** that this model is a useful framework for considering the potential for population-level consequences on cetaceans from human activities of any kind.

The Workshop **agreed** that there is a crucial need to investigate mitigation measures. To do so the concept of bio-equivalence offered a useful way forward to provide a framework in which to develop measures of recovery and test for changes that can show such recovery. Therefore, it was agreed that the research design should account for the possibility to test the efficacy of these mitigation measures. Such measures should rest on the precautionary principle.

The Workshop recognised that demonstrating a population level effect was very difficult and will take a long time. However, the Workshop also recognised that advice may be welcome from IWC member states on measures to mitigate any impact of whalewatching activities whether or not there are population level effects. Management objectives may be precautionary and be aimed at individual effects. In this context, research to understand better the mechanisms that result in shorter term or individual effects is essential to inform the most appropriate management action.

There are already management measures in place to control whalewatching operations in a number of places around the world. The Workshop **agreed** that these situations should be used as part of a research design, where possible; that is, that the study should be done in parallel with management, as appropriate. It also agreed that any study in which mitigation measures were implemented should continue for sufficient time to investigate whether or not recovery of the response variable(s) in question occurred. Care is needed in the definition of recovery.

The Workshop discussed the value of developing a management framework for whalewatching, similar to the IWC's RMP and AWMP for managing whaling and a similar framework developed for managing bycatch (Winship *et al*., 2007). This could potentially be used to explore the population consequences of whalewatching impacts on individuals and the effects of management actions and serve as a framework for developing standardised management advice.

7.1 Objectives

(1) Determine whether the results described in Shark Bay, Doubtful Sound, and Croatia studies can be observed in other situations.
(2) Determine how exposure to whalewatching affects the ecology, behaviour and/or physiology of cetaceans.
(3) Conduct short-term studies to inform the likelihood of long-term population impacts.
(4) Initiate long-term research programmes to assess temporal dynamics of response to disturbance including habituation, tolerance, and sensitisation.
(5) Develop a modelling framework to explore potential population consequences of changes in life history parameters given observed effects and effect sizes and use an additional dataset to test model predictions. The model can also be used to identify missing information.
(6) Determine the effectiveness of mitigation measures employed to reduce the effects of whalewatching.
(7) Develop a management framework for whalewatching that accounts for uncertainties, and includes monitoring and feedback mechanisms.

7.2 Research design

In planning a study covering a series of complex scenarios, it is important to keep in mind an ideal design. From this, it is possible to understand and account for components of the design that cannot be achieved because of logistical or other constraints. These can then be examined and commented on when making inferences from incomplete experimental data.

All aspects of the designs of particular components of the research programme should be based on pilot studies, cost-benefit and power analyses. These should be considered for all levels of the intrinsic hierarchies of the projects (e.g. spatial and temporal scales, species, populations, individuals, etc.).

The use of control areas near whalewatch sites will enable estimation of changes to variables not due to whalewatching, i.e. due to other ecological and environmental variation.

One successful method to understand the impacts of whalewatching would be to stop whalewatching in experimental areas. This is unlikely to be possible, but opportunities that exist should be exploited. There are, for example, areas (Brazil, Pacific NW, and San Ignacio Lagoon) where governments have been willing to displace boats which would make such experiments possible. In general, it is important that data are not neglected just because they are harder or more expensive to collect, making it tempting to focus on the 'easy' or 'standard' data.

The Workshop **agreed** that experimental designs need to take into account the spatio-temporal scale of the animal's life, as well as the portion of it where the whalewatching takes place. The Workshop agreed on stratification into the list of resident, feeding, breeding and migratory habitats. The key measure is over what range of an animal's home range and its life, whalewatching takes place in. Some animals have special resting areas (e.g. spinner dolphins in Hawaii) but there was no consensus about how to deal with such exceptions in the study design, unless a nearby control site can be designed.

The Workshop questioned whether it is better to choose a few sites with a very detailed understanding of most important variables, to best understand the role of whalewatching, or whether it is better to use many sites, with fewer covariates, where common results across sites are more likely to be related to the disturbance of interest. Some covariates are relatively easy to measure and it is important to establish the minimum of what needs to be measured in order to gain sufficient understanding of the collective systems to examine the role of whalewatching. This is especially true given the history of finding comparable study sites or reference sites. It is critical that all participants measure these variables correctly and in the same way. This is more important than having a long list of covariates.

There was discussion about the nature of control sites. Areas that are nearby that have different levels of whalewatching activity, but have the same animals moving among them, would not act as a control in a study that is looking for population-level effects. Hence, for animals with large home ranges, the closer two sites are, the more one can control for environmental variation, but the more confounded life history parameters are by a lack of independence. For species for whom feeding, breeding, and migrating are separated spatially, designs need to account for events in all parts of the animal's annual home range. Dispersal between populations needs to be estimated and considered in design, otherwise the empirical work cannot be connected to the longer-term forecasts.

In some cases, it may be possible to have comparable sites where whalewatching is well established, moderately established, and just starting. Unfortunately, in many of the cases where whalewatching is just starting, there is often poor information on life history parameters.

Caution is needed about studying species at the edges of their range, where animals may be more susceptible to effects. However, the added susceptibility may make these important study areas, although it becomes harder to tease out the impacts of whalewatching as opposed to the synergistic stressors to which the animals are exposed.

7.3 Variables

The value of certain techniques was highlighted because they could help collect data needed for several identified variables at once. For example, photo-identification studies can help to infer most population biology parameters as well as inform on habitat use of individuals. Similarly, hormonal extraction from scat sampling provides information on nutritional status (through T3 hormones), reproductive status and stress levels. Using some techniques in combinations (e.g. passive acoustics, photo-id, and tracking) can increase the level of details that can be obtained about some variables (e.g. habitat use).

Behavioural studies will focus on the following response variables: indices of socially-related calf survival, estimates of activity budgets, swim speed, and habitat use (given that those will be sampled for standardised periods). For physiological studies it will be important to sample the physiological status of individuals at the beginning and the end of annual sampling periods (nutritional and reproductive status). This could be done by assessing body condition in large whales and monitoring for reproductive hormones. Such studies should prioritise metabolic indices, particularly respiration rate. In addition, it will be valuable to obtain stress hormonal levels.

Independent variables should focus on the quantity and rate of exposure to the number and types of boats, the distance at which approaches take place and the received sound levels. Without simultaneous measures of prey availability and quality, it will be very difficult to tease out the effects of the whalewatching exposure. Covariates will have to be monitored for priority to prey availability, ambient noise and oceanographic variables.

7.4 Species/sites

Several species (bottlenose dolphins, killer whales, humpback whales and southern right whales) were quickly identified as likely study species, with some discussion of several other species (gray whales, *Eschrichtus robustus*; blue whales, *Balaenoptera musculus*; sperm whales, *Physeter macrocephalus*; pilot whales, *Globicephela* spp.; and Commerson's (*Cephalorhynchus commersonii*) and Hector's dolphins) that may be looked into further. A draft table of key variables around sites where the four primary species are subject to tourism impacts was developed and presented to the group. This will be further developed and refined for the development of the research proposal.

7.5 Funding

If the proposal is endorsed by the Scientific Committee, the Workshop anticipates that some seed funds will become available from the Commission. However, funds from external sources will be essential. Such sources include IWC member governments, science foundations in US and Europe, and the Oil and Gas Producers' Joint Industry Program. The Commission may be able to assist in making funds available from these sources.

It is critical for the success of this project that funding be coordinated among different study sites. Therefore, the Workshop **identified** advantages to having external funds administered by the IWC.

8. WORK PLAN

The Workshop **agreed** that further development of the proposal would be made by an editorial group of Bejder, Bjørge, Hammond, Lusseau, Underwood and Weinrich. The proposal will be sent to all members of the Workshop for comments prior to submission to the Scientific Committee for review and approval. Further, the Workshop **requested** that the Scientific Committee conveys the proposal to the Commission, with the aim of having the Commission's endorsement and assistance in finding funding.

9. ADOPTION OF REPORT

Bjørge acknowledged Weinrich and other members of the Workshop for their excellent work as rapporteurs. The report was adopted at 19:15 on 4 April 2008. The Workshop thanked the Chair for his hard work and guidance, and noted that he has finally now been on a whalewatching vessel!

REFERENCES

Altmann, J. 1974. Observational study of behavior: sampling methods. *Behaviour* 49: 227-67.

Baker, C.S. and Herman, L.M. 1989. Behavioural responses of summering humpback whales to vessel traffic. Experimental and opportunistic observations. Report NPS-NR-TRS-89-01 to the US National Park Service, Anchorage, AK. [Available from Glacier Bay National Park, PO Box 140, Gustavus, Alaska 99826].

Barlow, J. and Clapham, P.J. 1997. A new birth-interval approach to estimating demographic parameters of humpback whales. *Ecology* 78(2): 535-46.

Bejder, L., Samuels, A., Whitehead, H., Gales, N., Mann, J., Connor, R., Heithaus, M., Watson-Capps, J. and Flaherty, C. 2006. Decline in relative abundance of bottlenose dolphins exposed to long-term disturbance. *Conserv. Biol.* 20(6): 1,791-1,98.

Bose, N. and Lien, J. 1989. Propulsion of a fin whale (*Balaenoptera physalus*): why the fin whale is a fast swimmer. *Proc. R. Soc. Lond. Ser. B.* 237(1287): 175-200.

Coscarella, M.A., Dans, S.L., Crespo, E.A. and Pedraza, S.M. 2003. Potential impact of unregulated dolphin watching activities in Patagonia. *J. Cetacean Res. Manage.* 5(1): 77-84.

Danil, K. and Chivers, S.J. 2007. Growth and reproduction of female short-beaked common dolphins, *Delphinus delphis*, in the eastern tropical Pacific. *Canadian Journal of Zoology* 85: 108-21.

Failla, M., Iñiguez, M.A., Fernandez-Juricic, E. and Tossenberger, V. 2004. Effect of vessel traffic on Commerson's dolphin (*Cephalorynchus commersonii*) in Bahia San Julian, Patagonia, Argentina. Paper SC/56/WW7 presented to the the IWC Scientific Committee, July 2004, Sorrento, Italy (unpublished). 5pp. [Paper available from the Office of this Journal].

Gabriele, C.M., Straley, J.M., Mizroch, S.A., Baker, C.S., Craig, A.S., Herman, L.M., Glockner-Ferrari, M.J., Ferrari, M.J., Cerchio, S., Von Ziegesar, O., Darling, J., McSweeney, D., Quinn, T.D., III and Jacobsen, J.K. 2001. Estimating the mortality rate of humpback whale calves in the central North Pacific Ocean. *Can. J. Zool.* 79: 589-600.

Gaspar, R. 2003. Status of the resident bottlenose dolphin population in the Sado Estuary: Past, present and future. PhD thesis, St Andrews University, St Andrews. 213pp.

Hoyt, E. 2001. Whale watching 2001 – worldwide tourism, numbers, expenditures and expanding socioeconomic benefits. Report to IFAW, Crowborough, Sussex, UK. 157pp. [Available from *www.ifaw.org*].

Kraus, S.K. and Rolland, R.R. 2007. *The Urban Whale: North Atlantic Right Whales at the Crossroads*. Harvard University Press, Cambridge, MA. 576pp.

Lusseau, D., Williams, R., Wilson, B., Grellier, K., Barton, T., Hammond, P.S. and Thompson, P.M. 2004. Parallel influence of climate on the behaviour of Pacific killer whales and Atlantic bottlenose dolphins. *Ecology Letters* 7: 1068-76.

Mate, B.R., Nieukirk, S.L. and Kraus, S.D. 1997. Satellite-monitored movements of the northern right whale. *J. Wildl. Manage.* 61(4): 1393-405.

Mizroch, S.A., Herman, L.M., Straley, J.M., Glockner-Ferrari, D.A., Jurasz, C., Darling, J., Cerchio, S., Gabriele, C.M., Salden, D.R. and von Ziegesar, O. 2004. Estimating the adult survival rate of central north Pacific humpback whales (*Megaptera novaenagliae*). *J. Mammal.* 85(5): 963-72.

Noren, S.R. 2008. Infant carrying behaviour in dolphins: costly parental care in an aquatic environment. *Funct. Ecol.* 22: 284-88.

Ramp, C., Bérubé, M., Hagen, W. and Sears, R. 2006. Survival of adult blue whales *Balaenoptera musculus* in the Gulf of St. Lawrence, Canada. *Mar. Ecol. Prog. Ser.* 319: 287–95.

Read, A.J. and Gaskin, D.E. 1985. Radio tracking the movements and activities of harbor porpoises, *Phocoena phocoena* (L.), in the Bay of Fundy, Canada. *Fish. Bull.* 83(4): 543-52.

Rolland, R.M., Hamilton, P.K., Kraus, S.D., Davenport, B., Gillett, R.M. and Wasser, S.K. 2006. Faecal sampling using detection dogs to study reproduction and health in North Atlantic right whales (*Eubalaena glacialis*). *J. Cetacean Res. Manage.* 8(2): 121-25.

Rosenbaum, H.C., Weinrich, M.T., Stoleson, S.A., Gibbs, J.P., Baker, C.S. and DeSalle, R. 2002. The effect of differential reproductive success on population genetic structure: Correlations of life history with matrilines in humpback whales of the Gulf of Maine. *J. Hered.* 93(6): 389-99.

Stolen, M.K. and Barlow, J. 2003. A model life table for bottlenose dolphins (*Tursiops truncatus*) from the Indian River Lagoon system, Florida, USA. *Mar. Mammal Sci.* 19(4): 630-49.

Taylor, B.L., Chivers, S.J., Larese, J. and Perrin, W.F. 2007. Generation length and percent mature estimates for IUCN assessments of cetaceans. US NMFS Administrative Report LJ-07-01, Southwest Fisheries Science Center.

Williams, R. and Ashe, E. 2007. Killer whale evasive tactics vary with boat number. *J. Zool., London.* 272(4): 390-97.

Williams, R., Bain, D.E., Ford, J.K.B. and Trites, A.W. 2002a. Behavioural responses of male killer whales to a 'leapfrogging' vessel. *J. Cetacean Res. Manage.* 4(3): 305-10.

Williams, R., Lusseau, D. and Hammond, P.S. 2006. Estimating relative energetic costs of human disturbance to killer whales (*Orcinus orca*). *Biol. Conserv.* 133(3): 301-11.

Williams, R. and Thomas, L. 2007. Distribution and abundance of marine mammals in the coastal waters of British Columbia, Canada. *J. Cetacean Res. Manage.* 9(1): 15-28.

Williams, R.M., Trites, A.W. and Bain, D.E. 2002b. Behavioural responses of killer whales to whale-watching traffic: opportunistic observations and experimental approaches. *J. Zool. (Lond.)* 256: 255-70.

Winship, A.J., Deaville, R., Jepson, P.D., Learmouth, J.A., Northridge, S.P. and Hammond, P.S. 2007. Framework for estimating the growth rate of harbour porpoise populations in the North Sea and European Atlantic using a population model and data on abundance, life history and bycatch. 12pp. Paper SC/59/SM26 presented to the IWC Scientific Committee, May 2007, Anchorage, USA (unpublished). 12pp. [Paper available from the Office of this Journal].

Zeh, J., Poole, D., Miller, G., Koski, W., Baraff, L. and Rugh, D. 2002. Survival of bowhead whales, *Balaena mysticetus*, estimated from 1981-98 photo-identification data. *Biometrics* 58: 832-40.

Annex A

List of Participants

Lars Bejder
Murdoch University
Centre for Fish & Fisheries Research
South Street
Murdoch WA 6150
Australia
+ 61 424 508 498
l.bejder@murdoch.edu.au

Arne Bjørge
Institute of Marine Research
Gaustadalleen 21
Gatty Marine Laboratory
Oslo 349
Norway
+ 47 22958751
arne.bjorge@imr.no

Tamara Chapman
Department of Environment & Conservation
Locked Bag 104
Bentley Dc WA 6983
Australia
+ 61 8 9334 0516
tamara.chapman@dec.wa.gov.au

Dan Costa
University of California
100 Shatter Road
Santa Cruz California 95060
United States
+ 1 831 459 2786
costa@biology.ucsc.edu

Renaud De Stephanis
CIRCE (Conservation Information
& Research on Cetaceans)
Cabeza De Manzaneda 3
Pelayo
Algeciras, Cadiz 11390
Spain
+ 34 605 998195
renaud@stephanis.org

Caterina Fortuna
Central Institute for Marine Research
Via Di Casalotti 300
Roma 0166
Italy
+ 39 0661 570444
c.fortuna@icram.org

Karina Groch
Right Whale Project
Caixa Postal 201
Imbituba Sc 88780 000
Brazil
+ 55 4832 5529 22
krgroch@terra.com.br

Philip Hammond
University of St Andrews
Sea Mammal Research Unit
St Andrews, Fife KY16 8lB
United Kingdom
+ 44 1334 463222
psh2@st-andrews.ac.uk

Marta Hevia
Fundación Cethus
Potosí 2087
Olivos Buenos Aires B1636BUA
Argentina
+ 54 11 4796 3191
marta.hevia@cethus.org

David Lusseau
University of Aberdeen
Institute of Biological & Environmental Sciences
Aberdeen AB242tTZ
Scotland
d.lusseau@adbn.ac.uk

Peter Madsen
Aarhus University
Build 1131
Dept of Biological Sciences
Aarhus, Midtjyland 8200 Denmark
+ 45 89422602
peter.madsen@biology.au.dk

Dawn Noren
NOAA NMFS Northwest Fisheries Science Center
2725 Montlake Boulevard East
Seattle Washington 98112
United States
+ 1 206 302 2439
Dawn.Noren@noaa.gov

Milena Rafic
Department of the Environment, Water Heritage & the Arts
Cetacean Policy & Recovery Section
5 Farrell Place
Canberra 2600 ACT
Australia
+ 61 2 6274 2130
milena.rafic@environment.gov.au

Kate Reading
Department of Environment & Conservation
Locked Bag 104
Bentley Dc WA 6983
Australia
+ 61 8 9334 0516
kate.reading@dec.wa.gov.au

William Sherwin
University of New South Wales
Ecology & Evolution Research Centre
School of Biological Earth & Environmental Sciences
Sydney NSW 2052
Australia
+ 61 2 9385 2119
w.sherwin@unsw.edu.au

Andrew Trites
University of British Columbia
Marine Mammal Research Unit
Room 247, AERL, 2202 Main Mall
Vancouver BC V6T 1Z4
Canada
+ 1604 822 6557
a.trites@fisheries.ubc.ca

Tony Underwood
University of Sydney
Centre for Research on Ecological Impacts of Coastal Cities
Marine Ecology A11
Sydney NSW 2006
Australia
+ 61 2 9351 2590
aju@bio.usyd.edu.au

Sam Wasser
University of Washington
Box 351800, Department of Biology
Seattle, Washington 98195
United States
+ 1 206 543 1623
wassers@u.washington.edu

Mason Weinrich
Whale Center of New England
PO Box 159
Gloucester, MA 1930
United States
+1 978 281 6351
mason@whalecenter.org

Rob Williams
University of British Columbia
Marine Mammal Research Unit
Room 247, AERL, 2202 Main Mall
Vancouver BC V6t1z4
Canada
+ 1 604 822 6557
r.williams@fisheries.ubc.ca

Annex B

Agenda

1. Introductory items
2. Whalewatching: status of knowledge – strengths and weaknesses of current EIAs
3. Cetacean biology and ecology: key variables to monitor
4. Modelling: data requirements for modelling approaches
5. Study design: prioritation of variables to sample
6. Study site feasibility: practical aspects to consider
7. Proposal of large scale study: LaWE
8. Work plan
9. Adoption of report

Annex C

List of Documents

SC/LAWE08/

1. L. BEJDER and D. LUSSEAU. Valuable lessons from studies evaluating impacts of cetacean-watch tourism.
2. L. BEJDER. Is the reproductive success of female bottlenose dolphins influenced by home range characteristics and cumulative tour vessel exposure?
3. M. WEINRICH. Whale watch effects studies: What do we know?
4. P.S. HAMMOND. Information on cetacean life history and population dynamics.
5. J. ROBBINS. Potential use of multi-state mark recapture models to study whale watching impacts: an example from the Gulf of Maine.
6. J. MCDONALD, M. KRUETZEN, J. MANN, R. CONNOR, M. HEITHAUS and W.B. SHERWIN. Population viability analysis of bottlenose dolphins in Shark Bay World Heritage Area, Western Australia.
7. S. WASSER. Physiological measures of stress.
8. D. NOREN. Brief review of marine mammal bioenergetics and field methods to estimate field metabolic rates (FMRs) and potential bioenergetic impacts of vessel disturbance for large whales.
9. P. MADSEN. Cetaceans and noise from whale-watching vessels.
10. D. LUSSEAU, R. WILLIAMS, L. BEJDER, K.A. STOCKIN, D. BAIN, M.A. METHE, F. CHRISTIANSEN, E. MARTINEZ and P. BERGGREN. The resilience of animal behavior to disturbance.
11. M. HEVIA. LaWE 2008 Workshop: Argentina's Case.
12. K. GROCH and J.T. PALAZZO Jr. A proposal to improve monitoring of the effects of boat-based whale-watching in the right whale environmental protection area, Santa Catarina, Brazil.
13. A.J. UNDERWOOD. Considerations for assessing potential impacts of whale-watching.

Annex D

Population Consequences of Acoustic Disturbance (PCAD) Model