

Effects of vessels on behavior of individual southern resident killer whales (*Orcinus* sp.)

David E. Bain^{1*}, Rob Williams^{2,3}, Jodi C. Smith⁴, David Lusseau^{5,6}

¹Friday Harbor Laboratories, University of Washington, Friday Harbor, WA, USA

²Sea Mammal Research Unit, Gatty Marine Laboratory, St Andrews University, St Andrews, UK

³ Present address: Pearse Island, BC, Canada

⁴Coastal-Marine Research Group Ecology, Zoology and Environmental Sciences Institute of Natural Resources Massey University New Zealand

⁵University of Aberdeen, Department of Zoology, Lighthouse Field Station, Cromarty, UK

⁶Present address: Department of Biology, Dalhousie University, Halifax, NS, Canada

*dbain@grrescue.org

ABSTRACT: Southern resident killer whales numbered only 84 individuals in 2004. Disturbance by vessels may be a factor in the population's endangered status. To determine the importance of this factor, we compared behavior in the presence and absence of vessels in 2003-2005 at two different sites along San Juan Island, Washington State, USA. Theodolite tracks were summarized in terms of swimming path directness and deviation indices, travel speed, and rates of respiration and surface active display behaviors. Vessel number and distance were used in a generalized additive modelling framework as candidate explanatory variables for differences in whale behavior, along with natural factors such as sex, age, pod membership, time of day, time of year, geographic location, current and tide height. Path directness varied with number of vessels and distance to vessels. Increased distance travelled in the presence of vessels could result in increased energy expenditure relative to whales that could rest while waiting for affected whales to catch up. The likelihood and rate of surface active behavior varied with number of vessels. Number and proximity of vessels were also related to variability in respiratory intervals, path deviation index and swimming speed. The high proportion of time that southern resident killer whales spend in proximity to vessels raises the possibility that the short-term behavioral changes reported here may lead to biologically significant consequences.

INTRODUCTION

The Eastern North Pacific “Southern Resident” Stock of killer whales declined to fewer than 80 individuals in 2001, resulting in their listing as “Depleted” under the Marine Mammal Protection Act, and “Endangered” under the U.S. and Washington State Endangered Species Acts and Canada’s Species at Risk Act. The causes of this decline are uncertain, but many scientists consider a combination of reduction in prey resources, toxic chemicals, disturbance from vessel traffic, and other factors to have contributed (Krahn *et al.* 2002 and 2004, Federal Register 2004 and 2005, Killer Whale Recovery Team in press, Wiles 2004).

Vessel traffic may have contributed to the decline through a variety of mechanisms. Collisions between vessels and killer whales occur occasionally in residents and other killer whales and result in injury or death (Ford *et al.* 2000). One collision was observed in Southern Residents in 2005 that resulted in injury (K.C. Balcomb pers. comm.). Chemicals such as unburned fuel and exhaust may contribute to toxin load. The presence of noise from vessels may contribute to stress (Romano *et al.* 2004). Noise from vessel traffic may mask echolocation signals (Bain and Dahlheim 1994), thereby reducing foraging efficiency. Behavioral responses may result in increased energy expenditure, or disrupt feeding activity, which may reduce energy acquisition (Bain 2002, Williams *et al.* 2006, Bain *et al.* submitted). Energetic mechanisms for impact are of particular concern, because southern resident killer whales may be food limited (Ford *et al.* 2005).

Disturbance of wild animals could be implicated as a factor reducing the quality of life, foraging efficiency, fitness, or reproductive success of individual animals. Anthropogenic disturbance has been linked to changes in foraging behavior (*e.g.*, Galicia and Baldassarre 1997), reproductive success (*e.g.*, Safina and Burger 1983), and mating system and social structure (*e.g.*, Lacy and Martins 2003). These in turn, either singly or synergistically, could influence population dynamics.

Effects of vessel traffic vary within and between species, and include changes in respiration patterns, surface active behaviors, swimming velocity, vocal behavior, activity state, inter-individual spacing, wake riding, approach and avoidance, and displacement from habitat (Bain *et al.* 2006). Kruse (1991) and Williams *et al.* (2002ab) demonstrated short-term behavioral changes to vessel traffic in the adjacent, “northern resident” killer whales. Kruse (1991) found that northern residents increased swimming speed as vessel number increased. Northern residents swim in less predictable paths in the presence of vessels (Williams *et al.* 2002ab). Northern and southern residents are less likely to forage in the presence vessels (Williams *et al.* 2006, Bain *et al.* submitted). Adimey (1995) found percussive behavior of northern residents was inhibited in the presence of vessels, though Williams *et al.* (2002ab) found no significant differences. However, for southern resident killer whales, even subtle behavioral responses to boats have not been reported in the primary literature. This is a critical area of study because the San Juan and Gulf Islands are a region with high vessel traffic.

Because these whales are in the presence of vessels, including those not focused on whale watching, during much of the day, the potential for cumulative effects makes it important to investigate whether the behavior of southern resident killer whales is altered in the presence of vessels (Williams *et al.* 2006, Killer Whale Recovery Team in press,

Federal Register 2007). This study addresses relationships between vessel activity and Southern Resident killer whale behavior.

MATERIALS AND METHODS

Field methods. From 28 July to 30 September 2003, 1 May to 31 August 2004, and 15 May to 31 July 2005, a land-based team of observers monitored behavior of whales and activity of boats from two study sites (Figure 1). One site (the “North Site”) was located at 48° 30.561’ N, 123° 8.494’ W at an altitude of approximately 99m above mean lower low water. The other (South) site was located at Mt. Finlayson, near the southeast tip of San Juan Island. The South site was located at Mt. Finlayson (48° 27.421’ N, 122° 59.401’ W) at a height of 72m and the view of the eastern portion of Juan de Fuca Strait was unobstructed. The theodolite height was determined using the method of Williams et al. (2002ab) and Bailey & Lusseau (2004). Together, these sites were chosen to maximize sample size and to allow the behavioral observations to include the entire repertoire of the population. For the three seasons combined, data were obtained on 128 days over approximately nine months in the field.

Theodolite tracking of focal individuals and boats. The theodolite tracking team consisted of three individuals who moved opportunistically between the two study sites to maximize sample size. The team recorded boat and whale positions and activity using a Pentax ETH-10D theodolite interfaced to a PC-compatible computer running Theoprog (Williams et al. 2002ab, Williams and Ashe In press), a Bushnell 40x spotting scope, binoculars, and a mini-DV camera (DeNardo et al. 2001).

As whales entered the field of view from a study site, a focal individual was selected, identified based on Ford *et al.* (2000) and catalogs updated annually by the Center for Whale Research, and tracked for at least 800 seconds. After a tracking session was completed, a new focal individual was selected, if possible. Individuals were selected haphazardly, but were drawn as evenly as practicable from all pods, age, and sex classes. We attempted to choose individuals that would not be confused with other individuals nearby, and that were sufficiently close to shore to be accurately identified (typically within 3 km). The theodolite was used to record position of the focal individual at as many surfacings as possible, and the spotting scope and computer operators, who had a wider field of view, watched for surfacings missed by the theodolite operator, to ensure an accurate record of respiration rate and surface active behavior. While the focal whale appeared to be down on a long dive, the theodolite operator recorded vessel positions. In some cases, a second theodolite tracked only vessels.

In addition to recording positions of boats and whales, Theoprog was used to record activity states, behavioral events (*e.g.*, respirations and surface active behaviors such as breaches) and other notes (Williams *et al.* 2002ab). Boat and whale data were summarized for each track, such that each track was represented only once in the analyses. Independent variables included those related to: Time (Year, Day of Year and Time of Day); Location (Site); Focal Animal (Age, Sex); and Vessel Traffic (Point of Closest Approach, Overall Boat Count, Number of boats within 100, 400 and 1000m of the focal whale, and Number of boats observed within the observers’ field of view during the track [the maximum of the total number of vessels identified by the theodolite

operator, instantaneous counts by the theodolite and computer operators, and counts by the scan sampler within 1000m]). Calculation of these candidate explanatory variables is described in greater detail in Williams et al. (2002ab) and Williams and Ashe (In press).

The five dependent (*i.e.*, whale response) variables were calculated using methods described previously (Williams *et al.* 2002ab, Williams and Ashe In press), and included:

1. Inter-breath interval [**RESP**]: A mean time between breaths was calculated (in seconds) for each track. Only tracks lasting more than 800 seconds were included in the analysis to ensure the data reliably reflected the ongoing breathing pattern (Bain 1986, Kriete 1995).
2. Swimming Speed [**SPEED**]: The average swimming speed of the whale was obtained by dividing the total distance travelled by the duration of the tracking session and reported in km/h.

Two measures of path predictability were calculated: a *directness index* and a *deviation index*.

3. Directness Index [**DI**]: The directness index is generated by dividing the distance between end-points of a path (*i.e.*, crow's flight distance) by the cumulative surface distance covered during all dives and multiplying by 100. The directness index is the ratio of the diameter of a path to its perimeter, and ranges from zero (a circular path) to 100 (a straight line).
4. Deviation Index [**DEV**]: For each surfacing in a track, we calculated the angle between the path taken by a dive and the straight-line path predicted by the dive before it. The deviation index is the mean of the absolute value of each of these discrepancies, in degrees (potentially ranging from 0 to 180), during the entire track.
5. Surface-active Behavior [**SAB**]: We recorded each time that surface-active events such as spy-hopping, tail-slapping or breaching occurred, and present this as an average rate of events expected per hour.

Analysis of theodolite data from focal individuals. We modelled heterogeneity in whale behavior using generalized additive models, GAMs (Venables and Ripley 2002), in package **mgcv** (multiple generalized cross-validation) for program R (Wood 2001). The **mgcv** approach uses thin-plate regression splines (Wood 2003) for the smooth terms of each explanatory variable, but each spline carries a penalty for excessive flexibility (Wood 2000). Flexibility is determined by the number of 'knots' (approximately one higher than the estimated degrees of freedom, edf) for each model term, between which the functional, or smoothed, relationship was modelled. The amount of flexibility given to any model term was determined in a maximum likelihood framework by minimising the GCV score of the whole model (*i.e.*, given the other terms in the model). Models were penalized for being over-parameterized, and the degree of

smoothing was automated for each model term simultaneously. This avoided the problem common to many step-wise procedures, whereby the order in which terms are presented to the model influences the apparent significance of subsequent terms.

The default smoothing value used for splines was the default value set by package *mgcv*, 10 knots in each spline, corresponding to 9 degrees of freedom (Wood, 2001). Histograms of the response variables were used to determine the appropriate family distribution and link function. The following summarizes our model specification procedure adopted for each of the five response variables, y , during this study, using the framework proposed by Wood (2001):

1. A fully saturated model was fitted to the data: $\{y \sim Year + JDay + Time + s(Tide) + s(Current) + Site + Pod + s(Age) + Sex + s(PCA) + s(BOATS) + SUM100 + s(SUM400) + s(SUM1000) + Current$ with the default degree of smoothing (10 knots, 9 df).
2. Model fit was assessed using the `summary.gam` and `plot.gam` functions in *mgcv*, which showed coefficients, GCV score, explanatory power (deviance explained and R-squared score) and fit (residual plots).
3. For each linear term, the parameter coefficient (slope) was examined to see whether it was near 0 and the significance term to see whether it was near 1. If so, the term was removed to see whether the GCV score decreased and the explanatory power of the model increased. If so, the term was dropped from the model. If no marked improvement was detected by removing the term, then it remained in the model.
4. For each smooth model term, the estimated number of degrees of freedom was examined to see if it was near 1. The 95% confidence intervals for that term were examined to see whether they included zero across the range of observations. If so, the term was dropped temporarily, to see whether the GCV score dropped and the explanatory power of the model increased.
5. A term was dropped from the final model if it satisfied all three of the conditions in step 4 (*i.e.*, $edf \approx 1$; 95% CI's include zero across range of x ; and dropping the term decreased the GCV score and increased the values for R-squared and deviance explained). If the first criterion was met ($edf \approx 1$), but not the other two, then the smooth term was replaced by a linear term.

RESULTS

Theodolite tracking of focal individuals

We collected 42 tracks in 2003, 77 tracks in 2004, and 67 in 2005 that were of sufficient quality to use in the analysis (Table 1). Roughly 50% of the individuals in the population were sampled at least once during the three seasons. Whales were tracked for an average of 25.2 minutes over 2.6 km. Only two tracks in 2003 met the criteria for good boats-absent tracks.

Results of GAM-based analyses of focal animal behavior

Respiration analysis. The model that fitted the respiration data best included only three vessel traffic variables, but no whale related variables (age, sex and pod) or temporal variables (time of day and month of year, Table 2). The model described the variation in mean respiratory interval modestly, in that it was able to account for 13% of the deviance (note that the R-squared estimate is a less informative metric than the deviance explained for models based on anything other than normal error distribution). Two of the traffic variables (BOATS and SUM100) entered the model as linear terms, while the last variable (SUM400) entered the model as a smooth term.

As number of boats increased, the inter-breath interval showed a small but significant tendency to decrease. The relationship between inter-breath interval and number of boats within 100m, though, showed the opposite slope.

The smooth term describing covariation of SUM400 and inter-breath interval indicates that dive times tended to be shorter when no boats were present within 400m of the focal whale, and increased as number of boats increased to approximately 5. When more boats were present within this range, mean inter-breath interval declined, but this relationship became non-significant (*i.e.*, the confidence intervals comfortably spanned zero). Figure 2 shows the smooth spline relating mean time between breaths to the maximum number of boats counted within 400m of the whale, the linear terms relating mean time between breaths to two boat count variables, and the residuals of the fitted GAM.

Swimming speed. The selected model included the maximum number of boats scanned within 100m and 400m of the focal animal, as well as the site from which the data were collected (Table 3). Model fit was improved by dropping the intercept term. The model described variation in swimming speed quite well, in that it was able to account for 92.9% of the deviance. The linear relationship between boat count within 100m and swimming speed was negative (*i.e.*, whale swimming speed tended to decrease as number of boats within 100m of the whale increased). The non-linear relationship between swimming speed and MAX400 mirrored this relationship, with the relationship being fairly flat until the number of boats within 400m reached approximately 6 boats, at which point swimming speed increased dramatically. Figure 3 shows the smooth spline relating swimming speed to the maximum number of boats scanned within 400m of the whale, the linear terms, and the residuals of the fitted GAM.

Deviation index. The model that fitted the path deviation index data best included two boat count variables (BOATS and PCA), two whale-related variables (AGE and POD), and three ancillary variables (SITE, CURRENT and TIDE; Table 4). Model fit was improved by dropping the intercept term. The model demonstrated good power to describe variation in deviation index, accounting for 83.5% of the deviance explained, with an adjusted R-squared value of 0.188. Term-wise parameter estimates indicate that whales adopted smoother paths (*i.e.*, lower deviation index) at the South site than at the North site. The maximum number of boats recorded by the theodolite operator (BOATS) entered the model as a linear term with negative slope, indicating that whales exhibited relatively smooth paths when few boats were observed close to the whale and more

erratic paths when many boats were present (Table 4). The strong confounding effects of TIDE and CURRENT suggest that there may be something of biological importance, perhaps foraging activity, reflected in these data, and warrants further attention.

The smooth term relating deviation index to PCA (Point of Closest Approach) in the selected model is shown in Figure 4. The spline shows weak evidence that swimming paths showed a non-linear relationship with point of closest approach, however the confidence intervals span zero across a wide range of X. Secondly, the rugplot reveals that the observations were not uniformly spread across X. In other words, few observations were made when boats approached no closer than 500-1000m of the whale. There seems to be a confounding effect of age.

Directness index. The model that fitted the path directness index data best included two vessel traffic variables (PCA and BOATS), as well as YEAR, POD, and AGE (Table 5). Model fit was improved by dropping the intercept term. The model demonstrated some power to describe variation in directness index, accounting for 53.3% of the deviance (which is a better metric for quasi-family models than the adjusted R-squared value of 0.0326). The variables YEAR, POD and PCA entered the model as linear terms. Figure 5 shows the linear and non-linear terms that entered the model and the residuals of the fitted GAM. While there is no dramatic evidence of a pattern in the residuals, there is some suggestion of asymmetry about zero. Recall that directness index was bounded between 0 and 1, but that direct paths (*i.e.*, those near 1) were much more common than those near 0. We found that the best model specification was the quasi-likelihood framework approach with a log link function. However, model fit and convergence may have been constrained by the asymmetry of this response variable.

Surface active behavior. Surface active behavior tended to occur in bouts widely separated in time. As a result, many tracks had no surface active behavior. Those that did have any at all, tended to have at least a few events and could have many. We normalized the rate of SAB to number of events per hour, but found the model was unable to fit high rates of SAB, perhaps due to disproportionately large corrections in short tracks (*i.e.*, if the interval between bouts is large compared to the sampling period, the correction for sample period would bias the data). Then we tried to treat SAB as either present (1) or absent (0) during a track (SAB.1.0). We found this value was positively correlated with track duration, as expected, but the GAM analysis could correct for this when considering other parameters. As a result, we analysed both SAB and SAB.1.0, in hopes that asking two variations on a common question with opposite bias would elucidate underlying trends.

The results for SAB 1.0 are shown in Table 6 and Figure 6. The results for SAB are shown in Table 7 and Figure 7. The model that fitted the SAB.1.0 data best included only one vessel traffic variable (SUM400) (Table 6). The analysis of SAB.1.0 suggests SAB was most likely to occur when the number of boats within 400m of the whale was small (1-3 boats; Figure 6). The model that fitted the SAB data best included only one vessel traffic variable (SUM100), but also POD and AGE (Table 7). Model fit was improved by dropping the intercept term. The analysis of SAB suggests that young animals were highly active, but this rate slowed as animals reach sexual and physical maturity, and by senescence, SAB was rare, although the trend was insignificant at most ages (Figure 7, UPPER LEFT). Rates of surface active behavior were higher in members

of K pod than the other two pods (Figure 7, UPPER RIGHT). The effect of boats (SUM100, Figure 7, LOWER LEFT) was linear; SAB was highest when boats were absent, and lowest when number of boats within 100m of the whale approached 4. The consistent trend for both SAB and SAB.1.0 to be maximized when the number of boats was small and they were in close proximity suggests that bias from sample duration was overcome in this aspect of the analysis.

Age was not a factor in SAB.1.0, perhaps suggesting that the probability of engaging in SAB is equal for all age classes, but that younger animals tend to do more once they get started. Similarly, pods may be equal in their probability of initiating SAB, but differ in the number of events once it is initiated.

DISCUSSION

Despite a model specification approach that penalized over-parameterisation, all six models fitted the data better with boat variables included than when boat variables were excluded. The models lend support for concluding that boats exerted a small but significant effect on behavior of southern resident killer whales in 2003-2005, but that the relationships were complex and often non-linear.

Williams et al. (2002b) and Williams and Ashe (In press) suggested that vessel number and vessel proximity were different dimensions of vessel traffic, and that a whale's response to changes in vessel number is likely to occur independently of its response to changes in proximity, and vice versa. As such, an increase in proximity need not have the same effect as an increase in number. This study supports their conclusion. Further, there are hints in our dataset that number and proximity may interact. That is, boat counts at different distances entered the models independently. In some cases the effects of changing boat numbers at different distances were similar (e.g., boat numbers within 100m and 400m both led to increases in respiratory interval with increasing boat number in the range of 0-4 boats). In other cases, trends were different (e.g., respiratory interval increased with increasing boat number when the count only included boats within 100m, but respiratory interval decreased with increasing boat number when all boats within about 2 km of the focal whale were included in the count).

We also observed non-linear effects. Qualitatively, sometimes it appeared there was a baseline distribution of behavior when boats were absent, a trend from 1 to about 3 vessels, and the opposite trend when the number of vessels was large (>10). Intermediate numbers produced different trends than large or small numbers of vessels, though it is unclear whether intermediate numbers result in a distinct pattern of changes, or simply reflect the interactions of trends when vessel numbers are small with those when numbers are large. The result was sometimes a U-shaped pattern, and in other cases a linear pattern where the mean occurred at an intermediate number of boats. Similarly, non-linear trends with distance were observed.

The complexities described above may account for inconsistencies among studies, many of which simply compared a vessel present to a vessel absent condition. It is striking how well our results agree with those for Northern Residents (Williams *et al.* 2002b, Williams and Ashe in press) when vessel number is taken into account.

Directness Index. The decrease in directness of travel with vessel traffic has appeared consistently in studies such as this one (Williams *et al.* 2002ab). This pattern is consistent with whales making concerted efforts to evade boats. The GAM analysis confirmed both vessel number and proximity were significant factors even after taking natural factors into account.

Deviation Index. The path deviation index was higher when boats were present than when they were absent, which is consistent with experiments conducted on female Northern Residents in 1995-6 (Williams *et al.* 2002b) and male Northern Residents in 1998 (Williams *et al.* 2002a) produced significant differences in this index, although the differences were insignificant in males in 1995-6. However, the GAM analysis in this study found both vessel number and proximity influenced the deviation index.

The deviation index would be expected to be relatively high during socialising and foraging. Tide and current were natural factors correlated with deviation index. Felleman (1991) suggested that foraging strategies of whales should take into account current related movements of their salmonid prey. This relationship merits additional investigation.

Breathing Patterns. Breathing changes have been inconsistent from one study to another. The GAM analysis suggests that inter-breath interval increases with increasing vessel number when the number of vessels is small (from 1 to about 5 vessels), but decreases when the number of vessels is large. This “U-shaped” response pattern may account for the inconsistent results. There may be alternative tactics employed that vary depending on vessel number and proximity. Vessel proximity did not enter directly as a factor, although boat counts at different distances entered separately, suggesting distance has some relevance. Perhaps this index is more sensitive to distances throughout the track relative to momentary close approaches than other indices are. Additional data will be needed to confirm whether the result reported here is robust in a wider range of conditions.

Surface Active Behavior. Changes in surface active behavior have been significant in many studies, although the direction of the change varies from one study to another. Our results suggest the inconsistency may be due to differences in methodology. For example, our work, and that of Williams *et al.* (2002b) suggest that SAB is maximized when one or a small number of boats approach closely, but SAB may be inhibited by other configurations of vessels. Data collected when boats are primarily in an inhibitory configuration may find vessels reduce rates of SAB. Alternatively, studies that pool all configurations may find no effect.

The analysis is further complicated by the relationship between track duration and measured values. Analysis of rates may need to be limited to longer tracks than some of those used here. Longer tracks would also be helpful for one-zero sampling, as that would allow subdividing tracks into multiple short segments. The tendency of surface active behavior to occur in bouts, along with the fact that surface active behavior is a somewhat artificial class composed of behavior patterns with a wide-range of functions, make it difficult to address these behavior patterns with statistical rigor. Nonetheless, the

increased probability of SAB occurring in the presence of vessels appears robust, as the effect is large and present in numerous datasets.

One could speculate that threat displays consisting of surface active behaviors such as breaches, slaps, and fluke lifts (Tavolga 1966, Norris *et al.* 1994, Lusseau in press ab) increased when vessels were close but not close enough to trigger an escape response (see Hediger 1964 for a discussion of the concept of flight distance). At greater distances, surface active behavior could be reduced to avoid attracting the attention of vessel operators. Baseline rates would reflect the use of surface active behavior for purposes independent of vessels such as communication among whales, foraging, and non-communicative purposes such as self-grooming.

Swimming Speed. The trend in swimming speed with respect to vessel traffic has been inconsistent from one study to another (e.g., contrast Kruse 1991 with Williams *et al.* 2002b). The GAM analysis suggested that number of boats and location could be important. Given the potential for changes in swimming speed to carry energetic costs to whales, as well as reflecting their physical condition, the factors influencing swimming speed deserve more careful assessment.

Comparisons with other species

The results reported here exhibit similarities and differences with other species. Effects being stronger when vessels are within 100m than when they are farther away is a common finding (e.g., Nowacek *et al.* 2001, Ritter 2003). Increases in travel and surface active behavior are also commonly found. Increases in horizontal avoidance and energy expenditure have also been reported in other species (e.g., *Tursiops*, Yazdi 2005).

In contrast, measures of swimming speed have varied among species and among studies within species, with some studies reporting increases (*Orcinus sp.*: Kruse 1991; *Tursiops truncatus*: Nowacek *et al.* 2001), some reporting no change (*Orcinus sp.*: this study; *Globicephala macrorhynchus*, *Stenella coeruleoalba*, *Steno bredanensis*: Ritter 2003), and some reporting both increases and decreases depending on vessel speed (*Stenella frontalis*, *Tursiops truncatus*: Ritter 2003). Some species show increased dive times as reported here (e.g., *Eschrichtius*, Sumich 1983), while other species shorten dives in the presence of boats. Some species are displaced from regions by vessels (e.g., *Tursiops*: Allen and Read 2000, Yazdi 2005, Bejder *et al.* in press), in contrast to resident killer whales who continued to use the same range in the presence of vessels.

The term habituation has been used both in the strict psychological sense, and a more general sense. Strictly, habituation is a lessening of a response to repetitive stimuli that is not due to fatigue. In discussions regarding habituation of killer whales to vessel traffic, it has also been used to refer to a lessening in response due to suites of related stimuli through a variety of mechanisms. Whalewatching does not lend itself to habituation in the strict sense. Boats are frequently changing the stimuli received by killer whales. Stimuli change as engines are turned on and off, engine speed changes, directions of travel change, distances between vessels and killer whales change, and spatial arrangements of multiple vessels change. Thus it is not surprising that after more than 30 years of exposure to whale-oriented vessel traffic, both northern and southern Residents still exhibit behavioral changes.

Previous studies (Williams et al. 2002ab, Williams and Ashe In press) suggested killer whales adopted various tactics in the presence of vessels with various operating practices (speed, noise, distance, numbers, etc.). Thus killer whales may optimize their tactics in response to vessels over time, resulting in reduced responses that would correspond to habituation in the less strict sense.

This study found evidence of small but biologically significant changes in behavior in the presence of vessels. These effects indicate the need to develop and enforce regulations for vessel operation near killer whales.

CONCLUSIONS

Horizontal avoidance (evidenced by changes in directness and deviation indices) appears consistently across studies, although the statistical significance may depend on the sample size of the study and the pattern of vessel traffic experienced by the exposure group. This may lead to an increase in energy expenditure. Surface active behavior often shows significant differences depending on vessel activity, although such results are inconsistent in their magnitude and direction. This indicates that surface active behavior can be triggered by vessels, in addition to its natural occurrence for other purposes. Surface active behavior is largely composed of threat displays, so a relationship to vessel traffic is not unexpected. Non-linear responses to changes in vessel proximity and number probably account for the inconsistent results. Since many surface active behaviors are threat displays, they may be indicative of stress, and we urge additional research on captive and free-ranging killer whales to assess potential linkages among anthropogenic activity, stress and rates of surface active behavior. Some surface active behaviors like breaching require increased energy expenditure, so should be considered when calculating cumulative effects. Average inter-breath interval (IBI) and swimming speed do not show consistent changes across studies.

The behavior of southern residents in the presence of vessels is consistent with that observed in northern residents. This increases the confidence that can be placed in cross-population extrapolations, and in using northern residents as proxies for southern residents when conditions preclude experimentation on southern residents (Williams and Ashe In press). Indeed, it is time for a meta-analysis of existing data from both populations, given the potential to increase statistical strength through the larger sample size to answer questions that small sample size precludes addressing through single studies alone. Future research should focus on prey acquisition, and potential impact through other mechanisms such as noise, stress and toxins.

Acknowledgments. We would like to thank B. Gleason of San Juan Island National Historic Park for permission to use the South Site (permits SAJH-2004-SCI-0003 and SAJH-2005-SCI-0003), and the Reiss family for permission to use the North Site. We thank E. Ashe, A. Coleman, B. Cuddihy, A. Foote, M. Foss, J. Hogan, A. Miller, F. Robertson, D. Ross, and R. Taylor for assistance with data collection. The Whale Museum's Soundwatch Program, Marine Mammal Monitoring Program (M3) and The Whale Watch Operators Association Northwest assisted with increasing the quantity of no-boat data at our North Site. We give special thanks to Dawn Noren for her help with the design of this project and review of previous versions of the manuscript. The data used in this report were funded under National Marine Fisheries Service Contracts No. AB133F03SE0959, AB133F04CN0040 and AB133F05SE3965. The

views expressed in this manuscript are those of the authors, and not necessarily those of the funding agency.

LITERATURE CITED

Adimey NM (1995) A descriptive study of the percussive behavior of orcas, *Orcinus orca*, in Johnstone Strait, British Columbia. MS. thesis, Nova Southeastern University, Fort Lauderdale, FL

Allen MC, Read AJ (2000) Habitat selection of foraging bottlenose dolphins in relation to boat density near Clearwater, Florida. *Mar Mamm Sci* 16:815-824.

Bailey H, Lusseau D. (2004) Increasing the precision of theodolite tracking: Modified technique to calculate the altitude of land-based observation sites. *Mar Mamm Sci* 20:880-885.

Bain DE (1986) Acoustic behavior of *Orcinus*: sequences, periodicity, behavioral correlates and an automated technique for call classification. In: Kirkevoold BC & Lockard JS (eds) Behavioral biology of killer whales. Alan R. Liss, New York, p 335-371.

Bain DE (2002) A model linking energetic effects of whale watching to killer whale (*Orcinus orca*) population dynamics. Orca Relief Citizens Alliance, Friday Harbor, WA http://orcarelieve.org/docs/bain_paper.pdf

Bain DE, Dahlheim ME (1994) Effects of masking noise on detection thresholds of killer whales. In: Loughlin TR (ed) Marine mammals and the *Exxon Valdez*. Academic Press, San Diego, CA, p 243-256

Bain DE, Smith JC, Williams R, Lusseau D (2006) Effects of vessels on behavior of southern resident killer whales (*Orcinus* spp.). NMFS Contract Report No. AB133F03SE0959 and AB133F04CN0040, Seattle, WA

Bejder L, Samuels A, Whitehead H, Gales N, Mann J, Connor R, Heithaus M, Watson-Capps J, Flaherty C, Krützen M (in press) Decline in relative abundance of bottlenose dolphins exposed to long-term disturbance. *Conserv Biol*

DeNardo C, Dougherty, M, Hastie, GD, Leaper, R, Wilson, B, Thompson, PM (2001) A new technique for investigating variability in spatial relationships within groups of free ranging cetaceans. *J Appl Ecol* 38: 888-895.

Federal Register (2004) Endangered and Threatened Wildlife and Plants: Proposed Threatened Status for Southern Resident Killer Whales. Federal Register 69:76673-76682.

- Federal Register (2005) Endangered and Threatened Wildlife and Plants: Endangered Status for Southern Resident Killer Whales. Federal Register 70:69903-69912.
- Federal Register (2007) Protective regulations for killer whales in the Northwest Region under the Endangered Species Act and Marine Mammal Protection Act. Federal Register 72:13464-13467.
- Felleman FL, Heimlich-Boran JR, Osborne RW (1991) Feeding Ecology of the killer whale (*Orcinus orca*). In: Pryor KW, Norris KS (eds) Dolphin societies: discoveries and puzzles. University of California Press, Berkeley, CA p 113-147
- Ford JKB, Ellis GM, Balcomb KC (2000) Killer whales: the natural history and genealogy of *Orcinus orca* in British Columbia and Washington. 2nd edn. University of British Columbia Press, Vancouver, BC
- Ford JKB, Ellis GM, Olesiuk PF (2005) Linking prey and population dynamics: did food limitation cause recent declines of 'resident' killer whales (*Orcinus orca*) in British Columbia? Canadian Science Advisory Secretariat Research Document 2005/042
- Galicia E, Baldassare GA (1997) Effects of motorized tourboats on the behavior of non-breeding American flamingos in Yucatan, Mexico. Conserv Biol 11:1159-1165.
- Hediger H (1964) Wild animals in captivity. Dover, New York.
- Killer Whale Recovery Team (in press) Recovery Strategy for the Northern and Southern Resident Killer Whales (*Orcinus orca*) in Canada. Species at Risk Recovery Strategy Series, Fisheries and Oceans Canada, Ottawa, Canada
- Krahn MM, Ford MJ, Perrin WF, Wade PR, Angliss RP, Hanson MB, Taylor BL, Ylitalo GM, Dahlheim ME, Stein JE, Waples RS (2004) 2004 Status review of southern resident killer whales (*Orcinus orca*) under the Endangered Species Act. U.S. Dept. Commer., NOAA Tech. Memo. NMFSNWFS-62
- Krahn MM, Wade PR, Kalinowski ST, Dahlheim ME, Taylor BL, Hanson MB, Ylitalo GM, Angliss RP, Stein JE, Waples RS (2002) Status Review of Southern Resident killer whales (*Orcinus orca*) under the Endangered Species Act. U.S. Dept. Commer., NOAA Tech. Memo. NMFS-NWFS-54
- Kriete, B (1995) Bioenergetics in the killer whale, *Orcinus orca*. PhD dissertation, University of British Columbia, Vancouver, BC
- Kruse S (1991) The interactions between killer whales and boats in Johnstone Strait, B.C. In: Pryor, K, Norris, KS (eds.) Dolphin societies: discoveries and puzzles, University of California Press, Berkeley, CA p 149-159

- Lacy KE, Martins EP (2003) The effect of anthropogenic habitat usage on the social behavior of a vulnerable species, *Cyclura nubila*. *Anim Conserv* 6:3-9.
- Lusseau, D (in press a) Why do dolphins jump? Interpreting the behavioral repertoire of bottlenose dolphins in Doubtful Sound, New Zealand. *Behavioral Processes*.
- Lusseau, D (in press b) The short-term behavioral reactions of bottlenose dolphins to interactions with boats in Doubtful Sound, New Zealand. *Mar Mamm Sci*
- Norris KS, Würsig B, Wells RS (1994) Social behavior. In: Norris KS, Würsig B, Wells RS, Würsig M (eds) *The Hawaiian spinner dolphin*. University of California Press. Berkeley, CA, p 103-121
- Nowacek SM, Wells RS, Solow, AR (2001) Short-term effects of boat traffic on bottlenose dolphins, *Tursiops truncatus*, in Sarasota Bay, Florida. *Mar Mam Sci* 17: 673-688.
- Ritter, F (2003) Interactions of cetaceans with whale watching boats – implications for the management of whale watching tourism. M.E.E.R. e.V., Berlin, Germany
- Romano TA, Keogh MJ, Kelly C, Feng P, Berk L, Schlundt CE, Carder DA, Finneran JJ (2004) Anthropogenic sound and marine mammal health: measures of the nervous and immune systems before and after intense sound exposure. *Can J Fish Aquat Sci* 61:1124-1134.
- Safina C, Burger J (1983) Effects of human disturbance on reproductive success in the black skimmer. *Condor* 85:164-171
- Sumich, J., L. (1983): Swimming velocities, breathing patterns, and estimated costs of locomotion in migrating gray whales, *Eschrichtius robustus*. *Canadian Journal of Zoology* 61(3): 647-652.
- Tavolga MC (1966) Behavior of the bottlenose dolphin (*Tursiops truncatus*): social interactions in a captive colony. In Norris KS (ed) *Whales, dolphins and porpoises*. University of California Press. Berkeley, CA. p 718-730
- Venables WN, Ripley BD (2002) *Modern applied statistics with S*. Springer, New York
- Wiles GJ (2004) Washington State status report for the killer whale. Washington Department Fish and Wildlife, Olympia, WA
- Williams R, and Ashe E (in press) Killer whale evasive tactics vary with number of boats. *J Zool Lond*

- Williams R, Bain DE, Ford JKB, Trites AW (2002a) Behavioural responses of male killer whales to a “leapfrogging” vessel. *J Cet Res Manage* 4:305-310
- Williams R, Trites A, Bain, DE (2002b) Behavioural responses of killer whales (*Orcinus orca*) to whale-watching boats: opportunistic observations and experimental approaches. *J Zool Lond* 256:255-270
- Williams R, Lusseau D, Hammond PS (2006) Estimating relative energetic costs of human disturbance to killer whales (*Orcinus orca*). *Biol Conserv* 133:301-311
- Wood SN (2000) Modelling and smoothing parameter estimation with multiple quadratic penalties. *J R Statist Soc B*. 62:413-428
- Wood SN (2001) mgcv: GAMs and generalized ridge regression for R.” *R News* 1: 20-25.
- Wood SN (2003) Thin-plate regression splines. *J R Statist. Soc. B* 65:95-114
- Yazdi, P (2005) Einfluss der Tourismusboote auf das Verhalten und die Energetik der Großen Tümmler (*Tursiops truncatus*) vor der Insel Choros, Chile. PhD Dissertation. Christian-Albrechts-Universität, Kiel. (In German with English summary).

Tables.

Table 1. Sample size of theodolite tracks.

Sample size (number of tracks)		2003	2004	2005	Total
Study site					
	North	28	52	47	127
	South	14	25	20	59
Month					
	May	0	33	14	47
	June	0	13	27	40
	July	4	19	26	49
	August	16	12	0	28
	September	22	0	0	22
Track duration					
	13.3-20 min	15	27	16	58
	>20 min	27	50	51	128
Sex of focal animal					
	Female	7	39	16	62
	Male	27	38	41	106
	Unknown	8	0	10	18
Pod of focal animal					
	J	8	44	19	71
	K	4	9	11	24
	L	21	24	26	71
	Unknown	9	0	11	20
Traffic (number of unique boats in theodolite track)					
	0	2	19	4	25
	1	2	2	10	14
	2	0	4	2	6
	3	3	6	3	12
	4	1	6	2	9
	5	6	2	2	10
	6-10	17	15	13	45
	11-15	6	11	16	33
	16-20	1	3	4	8
	21-25	2	2	9	13
	26-30	2	3	1	6
	31-35	0	1	0	1
	36-40	0	1	1	2
	41-45	0	2	0	2
Minimum number of focal individuals sampled (ignoring unknowns)					
		13	34	24	45

GCV score = 1.1132

Scale est. = 1.0523

n = 153

Table 4. Summary of selected model describing heterogeneity in path deviation index as linear (top) and smooth (*i.e.*, non-linear, bottom) functions of covariates selected by **mgcv**.

Family: quasi

Link function: log

Formula:

DEV ~ SITE + POD + s(AGE) + s(PCA) + BOATS + CURRENT + TIDE - 1

Parametric coefficients:

	Estimate	std. err.	t ratio	Pr(> t)
SITENORTH	3.8345	0.1368	28.02	< 2.22e-16
SITESOUTH	3.7238	0.1459	25.52	< 2.22e-16
PODK	-0.4178	0.1456	-2.869	0.0047283
PODL	-0.017296	0.1029	-0.1681	0.86677
BOATS	-0.015458	0.005596	-2.762	0.0064791
CURRENT	0.12404	0.06914	1.794	0.074881
TIDE	-0.002979	0.0009788	-3.044	0.0027730

Approximate significance of smooth terms:

	edf	chi.sq	p-value
s(AGE)	5.202	15.052	0.014762
s(PCA)	5.523	14.487	0.022309

R-sq.(adj) = 0.188

Deviance explained = 83.5%

GCV score = 202.28

Scale est. = 180.42

n = 164

Table 5: Summary of selected model describing heterogeneity in path directness index as functions of covariates selected by **mgcv**.

Family: quasi

Link function: log

Formula:

DI ~ YEAR + POD + s(AGE) + PCA + s(BOATS) - 1

Parametric coefficients:

	Estimate	std. err.	t ratio	Pr(> t)
YEAR	-0.00014202	1.553e-05	-9.145	2.9417e-16
PODJ	-0.040724	0.03777	-1.078	0.28262
PODK	0.090947	0.04594	1.98	0.049479
PODL	-0.050198	0.0372	-1.349	0.17916

PCA	3.4896e-05	1.596e-05	2.186	0.030307
-----	------------	-----------	-------	----------

Approximate significance of smooth terms:

	edf	chi.sq	p-value
s(AGE)	1.967	4.6553	0.098005
s(BOATS)	1	5.4682	0.020626

R-sq.(adj) = 0.0544	Deviance explained = 53.3%	
GCV score = 0.060885	Scale est. = 0.058298	n = 164

Table 6. Summary of selected model describing heterogeneity in likelihood of surface active behavior as linear (top) and smooth (bottom, *i.e.*, non-linear) functions of covariates selected by **mgcv**.

Family: binomial
Link function: logit

Formula:
SAB.1.0 ~ s(SUM400)

Parametric coefficients:

	Estimate	std. err.	t ratio	Pr(> t)
(Intercept)	-0.99329	0.1941	-5.118	3.0939e-07

Approximate significance of smooth terms:

	edf	chi.sq	p-value
s(SUM400)	3.214	7.0758	0.08098

R-sq.(adj) = 0.0389	Deviance explained = 5.97%	
UBRE score = -0.0097255	Scale est. = 1	n = 153

Table 7. Summary of selected model describing heterogeneity in rates of surface active behavior as linear (top) and smooth (bottom, *i.e.*, non-linear) functions of covariates selected by **mgcv**.

Family: quasi
Link function: log

Formula:
SAB ~ POD + s(AGE) + SUM100 - 1

Parametric coefficients:

	Estimate	std. err.	t ratio	Pr(> t)
--	----------	-----------	---------	----------

PODJ	0.52211	0.394	1.325	0.18751
PODK	1.6992	0.4161	4.084	7.7327e-05
PODL	1.1554	0.308	3.751	0.00026512
SUM100	-0.64357	0.4003	-1.608	0.11034

Approximate significance of smooth terms:

	edf	chi.sq	p-value
s(AGE)	4.154	5.755	0.24128

R-sq.(adj) = 0.0582
 GCV score = 13.705

Deviance explained = 29.9%
 Scale est. = 12.889 n = 137

Figure captions

Figure 1. The study area, with the North and South theodolite sites marked with stars.

Figure 2. Relationships between smoothed component (solid line; UPPER LEFT) of the explanatory variables (x-axis) used in the fitted GAM (after accounting for the linear effects of BOATS and SUM100; UPPER RIGHT and LOWER LEFT, respectively) and the response variable, mean time between breaths. The explanatory variable in the smooth term, SUM400, represents the maximum number of boats ever observed within 400m of the whale. SUM400 was allowed up to 9 d.f. in model selection, and the degree of smoothing was automated by **mgcv**. The x-axis contains a rugplot, in which small ticks mark locations of observations. Zero on the y-axis corresponds to no effect of the covariate on the estimated response (here, inter-breath interval). Values above zero on the y-axis indicate positive correlation, *i.e.*, long inter-breath intervals. The y-axis is labeled *s*(covariate name, estimated degrees of freedom) indicating the curve is smoothed. The dashed lines represent ± 2 standard errors, or roughly 95% confidence intervals. The last plot (BOTTOM RIGHT) shows the residuals of the fitted model. The slight positive skew (asymmetry about zero on the y axis) suggests that some overdispersion in respiration rate remained unmodeled by the selected GAM, but that the model provided a reasonably good fit to the data overall.

Figure 3. TOP LEFT - the smooth spline relating swimming speed to the maximum number of boats scanned within 400m of the whale suggests that whales tended to swim slowly when a few boats were observed within 400m, and then swimming speed tended to increase as boat number increased. TOP RIGHT and BOTTOM LEFT - the linear terms that entered the model. All other things being equal, whales tended to swim faster at the South Site than the North Site. Whales generally slowed down as number of boats within 100m increased. BOTTOM RIGHT- the residuals of the fitted GAM after accounting for the linear effects of Site, and maximum number of boats within 100m, and a smooth spline of the maximum number of boats within 400m of the whale. No pattern is evident in the residuals, indicating that the model fitted the data well, and that the maximum likelihood approach was able to account for the overdispersion in the response data.

Figure 4. TOP LEFT, and TOP CENTER -- Relationship between smoothed component (solid line) of the two non-linear explanatory variables (AGE and PCA) selected in the fitted GAM (after accounting for the effects of the other terms), and the response variable, path deviation index. The explanatory variable, maximum number of boats observed within 1000m, was allowed up to 9 d.f. and the degree of smoothing (≈ 1.48 d.f.) was automated by **mgcv**. The x-axis contains a rugplot, in which small ticks mark locations of observations. Zero on the y-axis corresponds to no effect of the covariate on the estimated response (deviation index). Values above zero on the y-axis indicate positive correlation, *i.e.*, an erratic, or above-average, path deviation index. The y-axis is labeled *s*(covariate name, estimated degrees of freedom). The dashed lines represent ± 2 standard errors, or roughly 95% confidence intervals. NEXT FIVE PLOTS (top right, middle row, bottom left) – The linear terms that entered the model. BOTTOM RIGHT –

This plot shows the residuals of the fitted model. No pattern is evident, indicating that the model fitted the data well, and that the maximum likelihood approach was able to account for any overdispersion in the response data.

Figure 5. TOP-- Relationships between smoothed component (solid line) of the smoothed explanatory variables (AGE and BOATS) selected in the fitted GAM, and the response variable, path directness index. The explanatory variables, age and maximum number of boats recorded by the theodolite operator, was allowed up to 9 d.f. and the degree of smoothing was automated by **mgcv**. The x-axis contains a rugplot, in which small ticks mark locations of observations. Zero on the y-axis corresponds to no effect of the covariate on the estimated response (deviation index). Values above zero on the y-axis indicate positive correlation, *i.e.*, a more direct path. The y-axis is labeled $s(\text{covariate name, estimated degrees of freedom})$. The dashed lines represent ± 2 standard errors, or roughly 95% confidence intervals. Overall, these plots suggest that young and old animals tended to mill in the study areas more than middle-aged whales (ca. 40a). The boat traffic plot suggests that when fewer than 10 boats were in the study area with the whale, paths tended to be less direct than when more than 10 boats were in the area. NEXT THREE PLOTS – Relationships between the three linear terms (YEAR, POD and PCA) and the response variable. Note that YEAR and PCA explained negligible components of the variance, but their retention improved overall model fit. BOTTOM RIGHT -- Residuals of the fitted model. The model fitted the data reasonably well, and the maximum likelihood approach (with a constant variance term) was able to account for overdispersion in the response data. Some evidence of asymmetry about zero in the y-axis remains, suggesting that some heterogeneity remains in the fitted model. [Note that while approximately half of the values are positive, and half negative, the largest negative values are larger than the largest positive values – this reflects the boundaries of the original scale, which was bounded by 0 and 1, but values of DI tended to be nearer 1 than 0.]

Figure 6. LEFT – Relationship between smoothed component (solid line) of the explanatory variables (SUM400) selected in the fitted GAM, and the response variable, probability of surface active behavior occurring. The explanatory variables were allowed up to 9df and the degree of smoothing was automated by **mgcv**. The x-axis contains a rugplot, in which small ticks mark locations of observations. Zero on the y-axis corresponds to no effect of the covariate on the estimated response (likelihood of surface active behavior). The y-axis is labeled $s(\text{covariate name, estimated degrees of freedom})$. The dashed lines represent ± 2 standard errors, or roughly 95% confidence intervals. RIGHT – Residuals of the fitted model. The plot of residuals indicates a good ability to classify samples into those with a high versus those with a low probability of having a surface active event occur.

Figure 7. TOP – Relationship between smoothed component (solid line) of the explanatory variables (AGE, POD and SUM100) selected in the fitted GAM, and the response variable, rate of surface active behavior (average number of events per hour). The explanatory variables were allowed up to 9 d.f. and the degree of smoothing was automated by **mgcv**, but only AGE was selected as a smoothed term. POD and SUM100

entered the model as linear terms. The x-axis contains a rugplot, in which small ticks mark locations of observations. Zero on the y-axis corresponds to no effect of the covariate on the estimated response (rate of surface active behavior). Values above zero on the y-axis indicate positive correlation, *i.e.*, higher rates of surface active behavior. The y-axis is labeled s(covariate name, estimated degrees of freedom). The dashed lines represent ± 2 standard errors, or roughly 95% confidence intervals. **BOTTOM RIGHT** – Residuals of the fitted model. The plot of residuals indicates poor ability to explain high rates of SAB.











