

## VARIATION IN THE VISUALLY OBSERVABLE BEHAVIOR OF GROUPS OF GALÁPAGOS SPERM WHALES<sup>1</sup>

HAL WHITEHEAD

Department of Biology,  
Dalhousie University,  
Halifax, Nova Scotia B3H 4J1, Canada  
E-mail: hwhitehe@is.dal.ca

### ABSTRACT

The behavior of groups of female and immature sperm whales (*Physeter macrocephalus*) was measured on 117 d within an 11-yr period off the Galápagos Islands, Ecuador. On each day, up to 18 measures of visually observable behavior were calculated. These concerned speeds, headings, movement patterns, diving synchrony, foraging formations, time spent socializing, and aerial behavior. The measured behavior of the sperm whales was considerably more variable when they were socializing than when foraging. None of the measures showed much correlation with sea-surface temperature, and only measures of consistency of movement were significantly correlated with defecation rate, an indicator of feeding success. However, month-long time periods accounted for over 50% of the variance in eight of eighteen measures, and, in the cases of surface speed and dive synchrony, the effects were statistically significant. In contrast, there was no autocorrelation with lag of one day in the residuals of any of the measures. Thus, behavior may be tracking substantial temporal variation in the whales' environment over scales of about several months. Groups of whales had significantly different travel patterns, but there was little other evidence for group-specific behavior, perhaps because tests of group-specific effects were not of adequate statistical power. Variation in sperm whale behavior, especially over time scales of a few months or longer and spatial scales of a few hundred kilometers or larger, should be considered when estimating densities from sighting surveys.

Key words: sperm whale, *Physeter macrocephalus*, behavior, Galápagos Islands.

<sup>1</sup> Although Ken Norris primarily worked in and around the Pacific Ocean, his interests were almost as wide geographically as they were scientifically. I met Ken aboard that wonderful barquentine *Regina Maris* sailing along the northeast coast of Newfoundland in 1976. During a short cruise his famous interest in just about everything and extraordinary insight into the natural world gave us a new and lively perspective into the waters and animals that we were trying to study. Another rich marine system which caught Ken's attention was the Galápagos, and one of the animals with which he was most intrigued was the sperm whale. Partially in response to his enthusiasms and ideas, we have been studying the Galápagos sperm whales since 1985. Here is an analysis of what we have seen.

Studies of variation in the behavior of populations are important for a range of purposes, including management and conservation. For instance, visual surveys of cetaceans may benefit from corrections for the time that animals are under water, the conspicuousness of their behavior, and their mode of travel (International Whaling Commission 1997). For such corrections to be applied appropriately, it is important to know how behavior varies between individuals, groups, time periods, and places.

Variation in behavior may also give information about adaptiveness and evolution. For instance, if a behavioral pattern varies with an environmental factor, then this may suggest its function [*e.g.*, Baird and Dill (1995) for variability in killer whale, *Orcinus orca*, group sizes]. Behavior which shows little correlation with the environment and is highly variable may be under weak selective pressure, whereas that which is highly correlated with changes in the environment may either be directly caused by environmental factors or be strongly selected to respond to these factors. Variation between sympatric groups can also be illustrative, being reflective of demographic conditions within groups, or, possibly, if the groups are fairly stable, of genetic or cultural drift.

When examining data from observational studies of behavior, it is common to focus on patterns which either seemed apparent in the field (such as dramatic changes in behavior with time, *e.g.*, Weinrich *et al.* 1992), or might be expected on theoretical grounds (such as changes in movement patterns with prey availability, *e.g.*, Würsig *et al.* 1991). However, more systematic, "exploratory," studies of behavioral variation are also important in that they can show general trends, for instance that a particular environmental variable has little relationship with any measurable behavior, or uncover unexpected relationships.

In this paper, I carry out a systematic survey of variation in the visually observable behavior of female and immature sperm whales (*Physeter macrocephalus*) off the Galápagos Islands, which we have been studying since 1985. Most studies of sperm whale behavior concern females [but see Mullins *et al.* (1988), Watkins *et al.* (1993) for studies of male behavior] and are conducted at the level of the group, as it is difficult to distinguish individuals in real time at sea. Groups consist of 20 or so females and immature whales (Best 1979, Gordon 1987, Whitehead *et al.* 1991).

From tracking of sperm whale groups off the Galápagos, we have found that movement patterns over time periods of hours to days are closely related to feeding success, which in turn is correlated with sea temperature (Smith and Whitehead 1993, Whitehead 1996*b*). When feeding is good—usually in cool temperatures—groups zig-zag over small areas, whereas if feeding is poor—characteristic of warm "El Niño" type conditions—groups move in fairly straight lines spending little time in any area. This indicates that sperm whales use movement strategies for optimizing feeding success in an environment which shows great variability over large temporal and spatial scales (Whitehead 1996*b*). The sizes of aggregations of sperm whale groups also seem to vary somewhat between study years off the Galápagos and between

Table 1. Galápagos studies, their date range (day/month), and number of days with sufficient data so that at least one behavioral measure could be calculated.

Year	Dates	Days with data
1985	23/2–20/4	33
1987	3/1–28/6	57
1989	13/4–21/5	15
1991	30/3–14/4	2
1995	23/4–3/6	10

places, probably a response to varying prey availability (Whitehead and Kahn 1992).

An unusual pattern of behavioral variation among Galápagos sperm whales is the distinctive, and seemingly stable, dialects in coda vocalizations recorded from different groups (Weilgart and Whitehead 1997). This is despite the fact that a group usually consists of a temporary assemblage, for a few days, of one or more stable matrilineal units (Whitehead *et al.* 1991, Richard *et al.* 1996). There are several feasible explanations for groups, which are not themselves stable in membership, appearing to have stable dialects. The vocalization patterns of a group may largely result from the actions of one, numerically-dominant, stable unit; or the units within groups may often be members of larger social entities which have shared distinctive vocalizations, such as the “acoustic clans” found in killer whales (Ford 1991).

Such results prompted this systematic investigation of variation in visually observable behavior in sperm whales. The goals of the study were to assess the overall variability of visually observable behavior, to determine whether behavior in general is related to sea temperature, feeding success, time period, or group identity, and to identify any potentially interesting relations between individual behavioral measures and these factors.

#### METHODS

*Field methods*—Data came from studies, each lasting 1–6 mo, in 1985, 1987, 1989, 1991, and 1995 off the Galápagos Islands, Ecuador (0°S, 91°W) (Table 1). Research was carried out during 2–3-wk trips to sea aboard a 10-m sloop (1985–1989) or 13-m cutter (1991, 1995), with crews of 5–7.

During these studies we tracked groups of sperm whales, acoustically and visually, for periods of days at a time (Whitehead and Gordon 1986). In daylight hours the whales were approached from behind so that we could photograph their tail flukes as the whales “fluked-up.” These photographs were used to identify individual sperm whales (Arnbom 1987). As each identification photograph was taken, we noted whether a defecation was, or was not, observed from the whale as it dived, or whether we were unable to observe whether the whale defecated.

In 1985 and 1987 during daylight (0600–1800 local time) we recorded the range, bearing, heading, and behavior of all visible clusters of whales

during scans each 5 min. A *cluster* is a set of whales within 100 m of one another coordinating their movements (Whitehead and Weilgart 1991). Most nearest-neighbor distances within clusters were 2–15 m. In subsequent years less complete behavioral records were collected, so that many behavioral measures are available only for 1985 and 1987 (Table 2).

Positions and sea-surface temperatures were recorded each 3 h (using a Tracor Transtar SATNAV satellite navigator in 1985–1991; Trimble Transpak GPS in 1995).

*Defining groups*—I needed to define which group was present on which day. This is not straightforward for sperm whales, because groups usually cannot be distinguished in real time and may change membership over periods of days (Gordon 1991). The photographic identifications of individuals allowed determination of which group was present on which day. I used a slight modification of methods used in previous analyses (Weilgart and Whitehead 1997, Whitehead *et al.* 1998). If  $n_A$  whales were identified from good quality [Arnbohm's (1987)  $Q \geq 3$ ] photographs on day  $A$ , and  $n_B$  on day  $B$ , with  $m_{AB}$  common to the two days, then the similarity in the complements of the two days was expressed by:

$$S = m_{AB}/\text{Minimum} \{n_A, n_B\}$$

If  $A$  is the day with fewer individuals identified, then this is the proportion of animals seen on day  $A$  which were also identified on day  $B$ . As about half the whales in the group being followed were identified each day (Whitehead *et al.* 1992), if the same individuals were being followed on both days,  $S = \sim 0.5$ . I used average linkage cluster analysis (Krebs 1989, p. 316) to assign days to groups, linking sets of days with  $S > S_c$  where  $S_c$  is the "cut-off similarity." A "maximal distinguishing similarity,"  $S_m$  was used to minimize spurious clustering of days between which membership changed substantially: sets of days linked to other sets with  $S_c > S > S_m$  were omitted from the analysis. Groups were delineated in four different ways, each defined by a combination of values of  $S_c$  and  $S_m$ :

- |     | $S_c$ | $S_m$ |  |
|-----|-------|-------|--|
| (1) | 0.4   | 0.4   | ( <i>i.e.</i> , strict grouping criterion, and no days omitted)  |
| (2) | 0.4   | 0.2   | ( <i>i.e.</i> , strict grouping criterion, and days omitted to avoid potentially spurious clustering)      |
| (3) | 0.25  | 0.25  | ( <i>i.e.</i> , less strict grouping criterion, and no days omitted)                                       |
| (4) | 0.25  | 0.1   | ( <i>i.e.</i> , less strict grouping criterion, and days omitted to avoid potentially spurious clustering) |

Methods (1) and (2) may split sets of days that should be grouped, whereas (3) and (4) may cluster days which should not be grouped. Methods (2) and (4) will reduce misclassifications at the expense of smaller sample sizes. It is not clear which of these methods is optimal for detecting group-specific effects, so all were used.

**Table 2. Behavioral measures: study years for which data are available, sample size (*n* in d), mean, coefficient of variation (CV), and normality after transformation (as indicated by Lilliefors test). 'S' indicates measure calculated just during hours of socializing, and 'F' just during hours of foraging.**

Variable	Units	Years	<i>n</i>	$\bar{x}$	CV	Range	Transformation	Normal? Lilliefors ( <i>P</i> )
Surface speed	kn	85,87	85	1.99	0.21	1.1-3.5	Log	0.51
Surface speed (S)	kn	85,87	47	1.83	0.30	1.0-3.5	Log	0.58
Surface speed (F)	kn	85,87	83	2.06	0.20	1.1-3.0	Log	0.36
Aerial	(breaches + lob-tails)/whale	85,87	81	0.085	1.59	0.00-0.93	Log (no. + 0.5/whales)	0.35
Aerial (S)	(breaches + lob-tails)/whale	85,87	41	0.115	1.63	0.00-0.98	Log (no. + 0.5/whales)	0.08
Aerial (F)	(breaches + lob-tails)/whale	85,87	73	0.062	1.40	0.00-0.62	Log (no. + 0.5/whales)	0.64
Mean cluster size (F)	animals	85,87	85	1.60	0.19	1-2.6	Log	0.15
Heading consistency		85,87	79	0.815	0.12	0.55-0.99	Arcsine-square-root	0.62
Heading consistency (S)		85,87	43	0.748	0.20	0.37-0.96	Arcsine-square-root	0.66
Heading consistency (F)		85,87	72	0.846	0.11	0.57-1.00	Arcsine-square-root	0.58
Diving synchrony	min	85,87	77	2.9	1.41	-3.2-13.1	None	0.05
Intercluster distance	m	85,87	74	149	0.21	87-228	None	0.29
Foraging formation		85,87	56	0.33	0.24	0.18-0.62	None	0.80
Proportion time socializing		85,87,95	75	0.24	1.04	0.0-1.0	Not possible	0.00
Time of day socializing	h of day	85,87,95	50	13.8	0.22	6-19	Square	0.11
Straight-line distance	km	85,87,89,91,95	78	26.8	0.58	1.6-60.5	None	0.12
Distance through water	km	85,87,89,91,95	78	39.1	0.29	15.1-64.5	None	0.18
Consistency of movement		85,87,89,91,95	78	0.61	0.44	0.07-1.00	Not possible	0.02

*Environmental and temporal measures*—Each of the following measures was calculated for every day for which appropriate data existed:

“Sea surface temperature”: measured in °C at 0600 local time.

“Defecation rate”: the proportion of fluke-ups at which we were potentially able to observe a defecation and did observe a defecation is called the defecation rate; it is taken to be an indication of the feeding success of the whales in the previous 24 h or so (Whitehead *et al.* 1989, Whitehead 1996b).

“Year-month”: this is a categorical measure with each month being given a separate value (January 1985 = 1, February 1985 = 2 . . . January 1986 = 13, *etc.*)

*Behavioral measures*—Over time scales of hours, much of the variation in sperm whale behavioral patterns can be explained by whether the whales are socializing or foraging (Whitehead and Weilgart 1991). “Socializing” was defined as taking place in hours when, as determined by Whitehead and Weilgart (1991):  $0.278(\text{mean cluster size}) - 2.03(\text{rate of fluking-up per whale at the surface in 5 min}) > 0.5$ . If this condition was not met the whales were considered to be “foraging” (Whitehead and Weilgart 1991). Thus, socializing occurred when whales were in large clusters and not fluking-up, and foraging when whales were single or in clusters of 2–3 and fluking-up. As socializing and foraging are usually quite distinct (Whitehead and Weilgart 1991), in 1995 we informally used these criteria to assign daylight hours to either foraging or socializing in real time while at sea.

Each behavioral measure was calculated for every day for which appropriate data existed. The behavioral measures are illustrated in Figure 1.

“Surface speed”: mean scalar speed recorded from clusters at the surface as estimated by comparison with the speed of the vessel as indicated by knotmeter (in kn = 1.85 km/h). This measure was also calculated separately for periods when animals were socializing and foraging.

“Aerial behavior”: number of breaches (leaps from the water) plus lobtails (thrashes of flukes on water surface) sighted during the day, standardized by dividing by the total of the number of whales counted at the surface at each 5-min interval (only given when the total of the 5-min counts of whales at the surface during the day exceeded 100, as variance in aerial behavior increased with smaller sample sizes). This measure was also calculated separately for periods when animals were socializing and foraging.

“Mean cluster size while foraging”: mean cluster size for clusters seen within 500 m of boat. This variable was calculated only when animals were foraging, because cluster sizes when socializing are likely dependent on group size.

“Heading consistency”: an indicator of the consistency of heading of clusters of whales within an hour (mean vector recorded velocity with all speeds standardized at 1.0, and 5 or more speeds recorded in the hour),

Measure	Low value	High value
Speed		
Aerial		
Cluster size		
Heading consistency		
Diving synchrony		
Inter-cluster distance		

Measure	Low value	High value
Foraging formation		
Proprn time socializing	+-----S--+ 6:00 18:00	+-----SSSSS+ 6:00 18:00
Time socializing	+SS-----+ 6:00 18:00	+-----SS-+ 6:00 18:00
Straight-line distance		
Distance through water		
Movement consistency		

Figure 1. Illustrations of behavioral measures. "S" = hour spent socializing; "-" = hour spent foraging.

averaged over all daylight hours (see Turchin 1998, p. 251; Whitehead and Weilgart 1991). This measure was also calculated separately for periods when animals were socializing and foraging.

“Diving synchrony”: the maximum time interval in minutes during any hour with no whales seen at the surface minus the expected value of this given the mean number of whales visible at the surface and the presence or absence of calves, from the model of Whitehead (1996a), and then averaged during each day (only for days with at least five such hours). Large values indicate diving synchrony, small values asynchrony.

“Intercluster distance”: the median nearest-neighbor distance of clusters seen within 5-min periods (using just periods when at least four clusters were seen, and at least ten of these periods during a day), restricting to clusters sighted within 500 m of the boat (so that conditions of visibility and observer ability were less likely to affect the measure).

“Foraging formation”: the ratio of the length of the minor axis to that of the principal axis of positions of whale clusters seen within all 5-min periods when at least four clusters were observed (Whitehead 1989). Low values indicate clusters aligned in a rank, high values indicate clusters in a nearly circular formation. Only days with at least ten eligible 5-min periods are considered.

“Proportion of time spent socializing”: proportion of daylight hours spent socializing (for days when there were at least 7 h of data).

“Time of day of socializing”: mean time of day when socializing in hours from midnight (for days when there were at least 7 h of data).

“Straight-line distance in daylight”: distance (in km) between 0600 and 1800 positions.

“Distance traveled through the water in daylight”: sum of distances (in km) between positions at 0600, 0900, 1200, 1500, and 1800.

“Consistency of movement”: an indicator of the consistency of heading of the group over the day: mean vector velocity over 3-h periods with all 3-h speeds standardized at 1.0 (see Turchin 1998, p. 251).

*Analysis*—When possible, each of the above measures was calculated for each 12 h of daylight spent studying sperm whales. Days are appropriate units for this analysis because there were usually sufficient identifications within a day so that the identity of the group present could be evaluated. Other behavioral measures either required a full day of data (*e.g.*, time of day for socializing) or benefitted from averaging over reasonably long time periods. Time periods of longer than a day were not practical, as the group being followed often changed overnight.

All behavioral variables, except the proportion of time socializing and the consistency of movement, were approximately normally distributed (after transformations as applicable), as indicated by the Lilliefors test (see Table 2).

Relationships of behavioral measures with sea-surface temperature and defecation rate were expressed using Pearson correlation coefficients. The contri-

bution of time period to the variances of the behavioral measures was assessed from the proportion of variance accounted for by month-year periods (calculated as in Krebs 1989, p. 455) and by autocorrelations among residuals for temporal variation over scales of a few days. For each behavioral measure these methods were also used to estimate the proportion of variance accounted for by group identity. Interactions between factors were not tested, as they would have too little power to be meaningful.

Although this was principally an exploratory observational study, I wished to attach approximate significance levels to the measures of relationship (correlations, autocorrelations, and proportions of variance accounted for) so that their importance might be assessed. There are problems with using standard hypothesis-testing methods with this data set. First, with eighteen behavioral measures being compared with five environmental/temporal/group factors, ninety significance tests result, without considering interactions. If Type I error rates are controlled over this entire set of tests (using Bonferroni or Dunn-Sidak methods), each has so little power that it will not detect substantial biological effects. If the Type I error rates are not corrected, then, in ninety tests about five true null hypotheses will be falsely rejected at  $P < 0.05$ . Multivariate techniques (such as canonical correlation analysis and multivariate analysis-of-variance) can reduce these problems (Manly 1992, p. 43), but there are too many missing values in this data set for useful application of multivariate methods. Missing values, together with the opportunistic and thus unbalanced nature of the data set, also prevent two-way analyses of variance in which the effects of group identity and time period, and their interactions, could be investigated simultaneously.

To minimize these difficulties, while providing reasonably sound estimates of the significance of the results, significance values were examined in four stages. Stage-wise error rates were considered when assessing the significance of results, using the Dunn-Sidak correction (Sokal and Rohlf 1981, p. 242) within each stage. The stages of analysis were:

1. Behavioral measures *versus* sea temperature. The Dunn-Sidak-corrected significance levels of the Pearson correlation coefficients were calculated. As all coefficients were small and not statistically significant (Table 3) sea surface temperature was not considered in the following stages.

2. Behavioral measures *versus* defecation rate. The Dunn-Sidak-corrected significance levels of the Pearson correlation coefficients were calculated.

3. Behavioral measures *versus* time period. Differences in behavioral measures between month-long time periods were examined using one-way analyses of variance (ANOVA), or analyses of covariance (ANCOVA) for those variables which were significantly ( $P < 0.05$ ) or marginally significantly ( $P < 0.10$ ) correlated with defecation rate. The significance values of autocorrelations among residuals were examined to test for independence and temporal variation over short time scales.

4. Behavioral measures *versus* group identity. Differences in behavioral measures between groups were examined using one-way analyses of variance or analyses of covariance for those variables which were significantly ( $P < 0.05$ )

Table 3. Behavioral measures correlated with sea-surface temperature (SST) and defecation rate, and proportion of variance accounted for by month-long time periods and group identity. *P*-values corrected using Dunn-Sidak procedure in each column.

Variable	Correlation with:		Proportion of variance accounted for by:	
	SST	Defecation rate	Time period	Group
Surface speed	-0.18	0.04	0.70+	0.04
Surface speed (S)	-0.31	0.05	0.59	-0.57
Surface speed (F)	-0.20	0.03	0.80**	0.16
Aerial behavior	0.06	0.09	-0.39	-0.22
Aerial behavior (S)	0.17	-0.39	-0.64	0.19
Aerial behavior (F)	-0.15	0.27	0.47	0.31
Mean cluster size (F)	0.08	-0.02	0.69	0.22
Heading consistency	-0.20	0.04	-0.12	0.15
Heading consistency (S)	-0.15	0.29	0.08	-0.77
Heading consistency (F)	-0.14	-0.08	0.65	0.42
Diving synchrony	0.31	-0.23	0.86**	0.23
Intercluster distance	0.04	0.21	0.35	0.44
Foraging formation	0.15	0.14	-0.66	0.21
Propn time socializing	-0.07	-0.07	0.27	0.12
Time of day socializing	-0.12	0.07	0.26	0.45
Straight-line distance	0.02	-0.40*	0.55	0.55
Distance through water	-0.05	-0.29	0.58	0.70**
Consistency of movement	-0.01	-0.36+	0.16	0.41

+  $0.05 < P < 0.10$

\*  $0.01 < P < 0.05$

\*\*  $P < 0.01$

or marginally significantly ( $P < 0.10$ ) correlated with defecation rate. Auto-correlations among residuals were examined to test for independence.

## RESULTS

*Variability in behavior*—Means and coefficients of variation ( $CV = SD/\bar{x}$ ) of all behavioral measures (before transformation) are listed in Table 2. Some measures show little variability ( $CV < 0.3$ ): speed at the surface, mean cluster size when foraging, heading consistency, intercluster distance, foraging formation, and distance traveled through the water. In contrast, the rates of aerial behavior, diving synchrony, and proportion of time spent socializing varied greatly between days ( $CV > 1$ ). The measures available separately for socializing and foraging (surface speed, aerial behavior, and heading consistency) were considerably more variable during socializing (Table 2). In the cases of surface speed and heading consistency (but not aerial behavior), this difference in variability was statistically significant (*t*-test of CVs,  $P < 0.01$ ).

*Behavioral measures versus sea temperature*—Correlation coefficients between behavioral measures and sea-surface temperature all had absolute magnitudes less than 0.32 and none were significant at  $P < 0.20$  (Table 3). Thus sea-

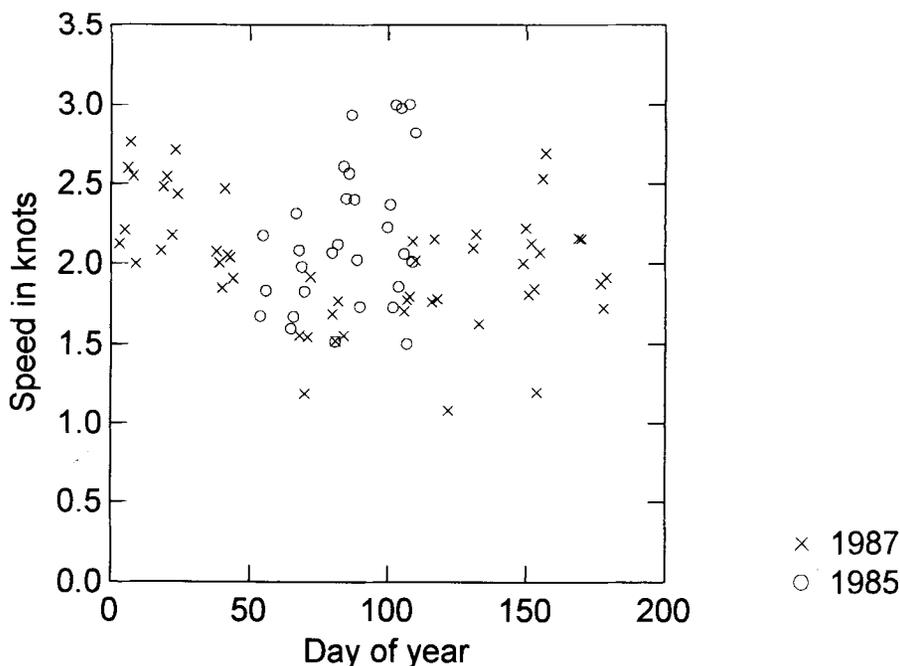


Figure 2. Mean surface speed, in minutes, with day of year ("1" = 1 January) for 1985 and 1987.

surface temperature explained less than 10% of the variance in each behavioral measure (as the squares of the correlation coefficients are less than 0.1).

*Behavioral measures versus defecation rate*—Defecation rate was negatively correlated with straight-line distance ( $r = -0.40$ ,  $P < 0.05$ ), aerial behavior when socializing ( $r = -0.39$ ,  $P > 0.10$ ), and consistency of movement ( $r = -0.36$ ,  $P < 0.10$ ) (Table 3). Other behavioral measures were poorly correlated with defecation rate ( $-0.30 < r < 0.30$ ), and not statistically significant ( $P > 0.50$ ).

*Behavioral measures versus time period*—Month-long time periods explained 50% or more of the variance in eight of eighteen measures. However, the tests were not powerful and the effect was only significant or marginally significant for speed, speed when foraging, and diving synchrony. Mean foraging speed and dive synchrony are plotted against day of year in Figure 2, 3. Within years, both measures seem to vary over periods of a few months. There is no clear consistency with calendar month between 1985 and 1987 in either measure. For all behavioral measures, autocorrelations with a lag of 1 d of residuals after the ANOVA or ANCOVA on group identity were less than 0.26 and not statistically significant.

*Behavioral measures versus group identity*—Results of analyses for group-specific behavior were similar whichever of the four methods of delineating groups was used. Results from method (4) (with  $S_c = 0.25$  and  $S_m = 0.1$ ) are shown in Table 3. Of the behavioral measures, group identity explained more than

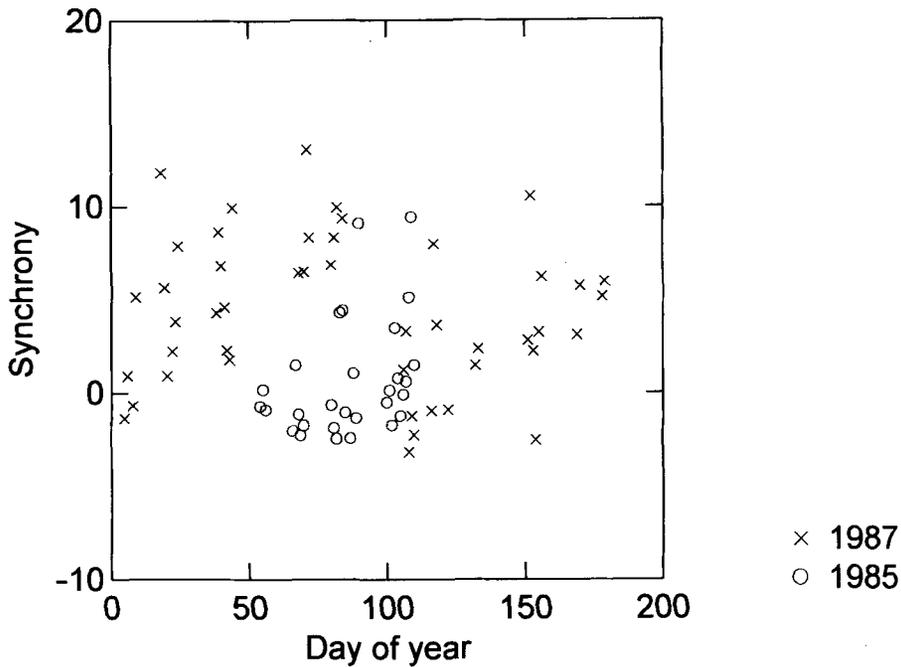


Figure 3. Mean synchrony of dives (in min), in residual minutes with no animals visible at surface, with day of year ("1" = 1 January) for 1985 and 1987.

50% of the variance in straight-line distance and distance traveled through the water in 12 h. However, the tests for group-specific effects were not powerful, and only in the case of distance traveled through the water was this significant or marginally significant between time periods. The four methods of defining groups produced significance levels for a group effect on distance traveled of  $P = 0.109, 0.066, 0.009,$  and  $0.007$  and suggested that group identity accounted for an estimated 62%–70% of the variance in this measure. The distance traveled through the water in 12 h is plotted against group identity [determined using method (4)] in Figure 4 (for groups for which this measure was available for two or more days). This Figure shows the strong differences between groups in this measure. Although only eight groups had data available for more than one year, the plot indicates that groups retained some of their idiosyncracies in this measure over periods of years. For all behavioral measures, autocorrelations with a lag of 1 d of residuals after the ANOVA or ANCOVA on group identity were less than 0.17 and not statistically significant.

#### DISCUSSION

*Variability in behavior*—The variables examined in this analysis showed considerable contrast in their variation. Behavior related to foraging (such as foraging formation and intercluster distance) was generally less variable than



This suggests that there are important aspects of the environment which are not well correlated with either sea-surface temperature or defecation rate but which affect the behavior of sperm whales. The marine environment shows considerable variation over periods of months (*e.g.*, McGowan 1990), and this could have an important influence on the behavior of sperm whales through intermediate variables such as prey type. In contrast, there was no substantial or significant residual autocorrelation of lag 1 d in any measure, indicating that environmental variation over temporal scales of a few days is not an important determinant of behavior. This may be a consequence of the low level of variability in marine environments over scales of a few days (Steele 1985).

*Group specific variation in behavior*—Although intergroup variability seemed quite important for a few behavioral measures, especially those concerned with movement, for only one variable was the difference between groups statistically significant: distance moved through the water. This result is consistent with the suggestion that within-group, culturally transmitted information may be important in allowing sperm whales to exploit very large home ranges efficiently and thus to survive long-term and large-scale downturns in their environment, such as those caused by the El-Niño phenomenon (Whitehead 1996b).

In this study I measured behavior at the level of the group, but groups are unstable over periods of days, each consisting of one or more much more stable units (Whitehead *et al.* 1991), and any idiosyncrasies in behavior would generally be expected to be expressed at the level of the unit. The method used for delineating groups should result in days being grouped only when the animals being followed were largely common to the days, usually because one numerically dominant unit was present on both days. However, changes in the animals being observed between days will have reduced the power of this analysis to detect unit-specific behavior.

In addition, because of the number of behavioral measures examined, and the use of Dunn-Sidak corrections for multiple tests, the tests for time period and, especially, group-specific effects are not powerful. Time period or group identity had to account for about 60%–80% of the variation in a measure to produce a statistically significant result at  $P < 0.05$ . This is despite the fact that we spent 117 d at sea observing sperm whales and during ninety of them recorded the visually observable behavior of all visible whales every 5 min of daylight.

Compared with this analysis, studies of variation in acoustic behavior derived from the same field studies have been much more informative about group-level and other effects (Weilgart and Whitehead 1997). This contrast results from several factors: vocalizations can be recorded accurately at longer ranges and through a much larger part of the daily routine of the whales than visually observable behavior, and our record of sounds is probably closer to the perception of the whales themselves than is our visually derived data. Therefore, where both are practical, acoustic studies of the behavior of sperm

whales may be expected to be generally more informative than those which use visually collected data.

*Conclusion*—This analysis has indicated the general importance of some factors in determining the behavior of sperm whales: sea-surface temperature has little direct effect; feeding success as indicated by defecation rates principally relates to movement patterns; and other unidentified factors influence sperm whale behavior patterns over periods of months. These results have practical implications for censuses of sperm whales using visual surveys, as several aspects of surface behavior which might affect the rates at which they are counted on visual censuses seem to vary substantially between month-long time periods—movement patterns, cluster sizes, and diving synchrony. As variation in the marine environment with spatial scale is in some ways analogous to its variation with temporal scale (Haurly *et al.* 1978, Whitehead 1996b), we might expect a similar substantial variation in sperm whale behavior over ranges of hundreds of kilometers. These sources of variation should be considered when trying to correct visual census counts for whale visibility. Modeling might be used to indicate the potential effects of large-scale behavioral variation. If effects are found to be significant, collection of quite detailed behavioral data during censuses may allow counts to be corrected separately for each spatial or temporal stratum.

#### ACKNOWLEDGMENTS

I am very grateful to the many colleagues who helped collect data in the field, and to Tom Arnbom, Jenny Christal, Susan Dufault, and Susan Waters who analyzed the identification photographs. The research was largely funded by the Natural Sciences and Engineering Research Council of Canada, the Whale and Dolphin Conservation Society, the International Whaling Commission, Cetacean Society International, and the Green Island Foundation. Godfrey Merlen, the Galápagos National Park Service, the Armada of Ecuador, the Charles Darwin Research Station, and the Instituto Nacional de Pesca of Ecuador helped with permits and assisted in other ways. Thanks to Jenny Christal, Shannon Gowans, Peter Tyack, and two anonymous reviewers for useful comments on manuscripts.

#### LITERATURE CITED

- ARNBOM, T. 1987. Individual identification of sperm whales. Report of the International Whaling Commission 37:201–204.
- BAIRD, R. W., AND L. M. DILL. 1995. Occurrence and behavior of transient killer whales: Seasonal and pod-specific variability, foraging behavior and prey handling. *Canadian Journal of Zoology* 73:1300–1311.
- BEST, P. B. 1979. Social organization in sperm whales, *Physeter macrocephalus*. Pages 227–289 in H. E. Winn and B. L. Olla, eds. *Behavior of marine animals*. Volume 3. Plenum, New York, NY.
- FORD, J. K. B. 1991. Vocal traditions among resident killer whales (*Orcinus orca*) in coastal waters of British Columbia. *Canadian Journal of Zoology* 69:1454–1483.
- GORDON, J. C. D. 1987. Sperm whale groups and social behaviour observed off Sri Lanka. Report of the International Whaling Commission 37:205–217.
- GORDON, J. C. D. 1991. The World Wildlife Fund's Indian Ocean Sperm Whale

- Project: An example of cetacean research within the Indian Ocean Sanctuary. Pages 219–239 in S. Leatherwood and G. P. Donovan, eds. Cetaceans and cetacean research in the Indian Ocean Sanctuary. UNEP Technical Report Number 3, Nairobi, Kenya.
- HAURY, L. R., J. A. MCGOWAN AND P. H. WIEBE. 1978. Patterns and processes in the space-time scales of plankton distributions. Pages 277–327 in J. H. Steele, ed. Spatial pattern in plankton communities. Plenum, New York, NY.
- INTERNATIONAL WHALING COMMISSION. 1997. Report of the abundance estimation working group. Report of the International Whaling Commission 47:261–290.
- JAQUET, N., AND H. WHITEHEAD. 1999. Movements, distribution and feeding success of sperm whales in the Pacific Ocean, over scales of days and tens of kilometers. *Aquatic Mammals* 25:1–13.
- KREBS, C. J. 1989. Ecological methodology. Harper and Row, New York, NY.
- MANLY, B. F. J. 1992. Multivariate statistical methods. Chapman & Hall, New York, NY.
- MCGOWAN, J. A. 1990. Climate and change in oceanic ecosystems: the value of time-series data. *Trends in Ecology and Evolution* 5:293–299.
- MULLINS, J., H. WHITEHEAD AND L. S. WEILGART. 1988. Behaviour and vocalizations of two single sperm whales, *Physeter macrocephalus*, off Nova Scotia. *Canadian Journal of Fisheries and Aquatic Sciences* 45:1736–1743.
- RICHARD, K. R., M. C. DILLON, H. WHITEHEAD AND J. M. WRIGHT. 1996. Patterns of kinship in groups of free-living sperm whales (*Physeter macrocephalus*) revealed by multiple molecular genetic analyses. *Proceedings of the National Academy of Sciences of the United States of America* 93:8792–8795.
- SMITH, S. C., AND H. WHITEHEAD. 1993. Variations in the feeding success and behaviour of Galápagos sperm whales (*Physeter macrocephalus*) as they relate to oceanographic conditions. *Canadian Journal of Zoology* 71:1991–1996.
- SOKAL, R. R., AND F. J. ROHLF. 1981. *Biometry*. 2nd Edition. W.H. Freeman, San Francisco, CA.
- STEELE, J. H. 1985. A comparison of terrestrial and marine ecological systems. *Nature* 313:355–358.
- TURCHIN, P. 1998. Quantitative analysis of movement. Sinauer Associates, Sunderland, MA.
- WATKINS, W. A., M. A. DAHER, K. M. FRISTRUP, T. J. HOWALD AND G. NOTARBARTOLO DI SCIARA. 1993. Sperm whales tagged with transponders and tracked underwater by sonar. *Marine Mammal Science* 9:55–67.
- WEILGART, L., AND H. WHITEHEAD. 1997. Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whales. *Behavioural Ecology and Sociobiology* 40:277–285.
- WEINRICH, M. T., M. R. SCHILLING AND C. R. BELT. 1992. Evidence for acquisition of a novel feeding behaviour: Lobtail feeding in humpback whales, *Megaptera novaeangliae*. *Animal Behaviour* 44:1059–1072.
- WHITEHEAD, H. 1989. Formations of foraging sperm whales, *Physeter macrocephalus*, off the Galápagos Islands. *Canadian Journal of Zoology* 67:2131–2139.
- WHITEHEAD, H. 1996a. Babysitting, dive synchrony, and indications of alloparental care in sperm whales. *Behavioural Ecology and Sociobiology* 38:237–244.
- WHITEHEAD, H. 1996b. Variation in the feeding success of sperm whales: Temporal scale, spatial scale and relationship to migrations. *Journal of Animal Ecology* 65: 429–438.
- WHITEHEAD, H., AND J. GORDON. 1986. Methods of obtaining data for assessing and modelling sperm whale populations which do not depend on catches. Reports of the International Whaling Commission (Special Issue 8):149–166.
- WHITEHEAD, H., AND B. KAHN. 1992. Temporal and geographical variation in the social structure of female sperm whales. *Canadian Journal of Zoology* 70:2145–2149.

- WHITEHEAD, H., AND L. WEILGART. 1991. Patterns of visually observable behaviour and vocalizations in groups of female sperm whales. *Behaviour* 118:275-296.
- WHITEHEAD, H., V. PAPASTAVROU AND S. C. SMITH. 1989. Feeding success of sperm whales and sea-surface temperatures off the Galápagos Islands. *Marine Ecology Progress Series* 53:201-203.
- WHITEHEAD, H., S. WATERS AND T. LYRHOLM. 1991. Social organization in female sperm whales and their offspring: Constant companions and casual acquaintances. *Behavioural Ecology and Sociobiology* 29:385-389.
- WHITEHEAD, H., S. WATERS AND T. LYRHOLM. 1992. Population structure of female and immature sperm whales (*Physeter macrocephalus*) off the Galápagos Islands. *Canadian Journal of Fisheries and Aquatic Sciences* 49:78-84.
- WHITEHEAD, H., M. DILLON, S. DUFAULT, L. WEILGART AND J. WRIGHT. 1998. Non-geographically based population structure of South Pacific sperm whales: Dialects, fluke-markings and genetics. *Journal of Animal Ecology* 67:253-262.
- WÜRSIG, B., F. CIPRIANO AND M. WÜRSIG. 1991. Dolphin movement patterns. Information from radio and theodolite tracking studies. Pages 79-112 in K. Pryor and K. S. Norris, eds. *Dolphin societies. Discoveries and puzzles*. University of California Press, Berkeley, CA.

Received: 11 May 1998

Accepted: 24 March 1999