MARINE MAMMAL SCIENCE, 17(4):718–728 (October 2001)

DIRECT ESTIMATION OF WITHIN-GROUP HETEROGENEITY IN PHOTO-IDENTIFICATION OF SPERM WHALES

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

ABSTRACT

Heterogeneity in photo-identification rates among individuals is a potentially serious problem in many studies of cetacean biology, especially the analysis of populations. However, this heterogeneity is usually difficult to identify or measure. Two instances in which closed groups of female and immature sperm whales (*Physeter macrocephalus*) were tracked and identified using fluke photographs over periods of days off the Galápagos Islands allowed direct examination of heterogeneity in identification rates. A group of nine animals followed in 1999 provided almost no evidence for heterogeneity (permutation test for heterogeneity, $P = 0.48$), with an estimated coefficient of variation in identification rates of 0.03 (95% CI from 1,000 bootstrap replications: 0.00–0.10). In contrast, the identification rates of a group of 22 animals followed in 1995 seemed to show potentially important differences ($P = 0.058$, CV = 0.20, 95% CI = 0.07–0.28). These differences were not related to the internal social structure of the group or to differences in numbers of markings on the flukes, but smaller whales had lower identification rates. Thus, young sperm whales may be underrepresented in photo-identification studies, but adults within groups seem to have similar identification rates. Situations in which animals are photo-identified from closed populations of known size are particularly useful for examining heterogeneity. They should be vigorously exploited by those who use photo-identification to examine population or behavioral biology.

Key words: photo-identification, sperm whale, *Physeter macrocephalus,* heterogeneity.

The photographic identification of individuals has become a major tool for the study of cetaceans, especially for research into populations (*e.g.,* International Whaling Commission 1990), migrations (*e.g.,* Stone *et al.* 1990), life history (*e.g.,* Clapham and Mayo 1990), behavior (*e.g.,* Mann 2000), and social

structure (*e.g.,* Whitehead and Dufault 1999). However, the use of this technique, especially for population analysis, requires certain assumptions which should be checked carefully (Hammond 1986).

For many quantitative analyses of photo-identification data, an important assumption is that animals have the same probability of being identified. Heterogeneity among individuals in identification rates is a major concern when analyzing cetacean populations (Hammond 1986). Individuals may differ in the likelihood that they are identified because of variation in behavior (*e.g.,* shy/curious animals, or differences in the manner in which they show the marked parts of their bodies), ecology (*e.g.,* habitat use patterns), or morphology (some being more easily identified than others). These differences bias population estimates obtained by many mark-recapture estimates (Hammond 1986) and may affect other uses of the data, such as the analysis of social structure (Whitehead and Dufault 1999).

Various techniques have been developed to analyze and compensate for heterogeneity (*e.g.,* Leslie 1958, Cormack 1985). However, most tests are indirect and have little power (Carothers 1973*a*). Ideally, heterogeneity should be examined using photo-identification data from a closed population of known size. Unfortunately, such situations are very rare in studies of wild animals (Carothers 1973*a*), including cetaceans.

However, in studies of female and immature sperm whales (*Physeter macrocephalus*) off the Galápagos Islands, we have twice been in such a position. This was the result of the very low density of sperm whales off the Galápagos during the late 1990s (Whitehead *et al.* 1997), which meant that groups very rarely interacted with other groups. On two occasions (in 1995 and 1999) we were able to follow a discrete, well-defined, and closed group of female and immature sperm whales for several days. In both cases each member of the group was identified on almost every day that we were following them, and no other sperm whale was identified (except for large and distinctive mature males who visited the groups briefly on a few occasions). These opportunities allowed the distributions of the number of photographic identifications of the members of each group to be examined and tested against the null hypothesis that all members of the group had an equal probability of being identified. The degree of heterogeneity in identification rates among individuals within each group could also be estimated by decomposing the variation in numbers of identifications among individuals into the portion resulting from sampling and that due to variation in identification rates.

Methods

Using 13-m auxiliary research vessels, we tracked groups of sperm whales both visually and by listening for their sounds with a directional hydrophone (Whitehead and Gordon 1986). In 1995 a group of 22 sperm whales was followed from 28 May to 3 June, apart from one nighttime break of 8 h (see Christal and Whitehead 2000 for details), and in 1999 a group of 9 whales was followed on 10–20 March, 28–31 March, 6 April, and 9–12 April. When

the whales clustered at the surface for social periods, approximately once per day, the counts of whales present were identical to the number of animals identified (22 and 9 whales, respectively).

During daylight the tracking vessel approached, from behind, whales that were breathing at the surface between deep dives. The nearest cluster of whales was generally chosen for approach. The ventral surfaces of the sperm whales' flukes were photographed as they were raised at the start of long and deep dives (Arnbom 1987). Photographs were processed using the methods described by Arnbom (1987) and Dufault and Whitehead (1995). Each photograph was given a quality value (Q) from 1 (very poor photograph) to 5 (excellent photograph), depending on its characteristics (such as focus and the size of the image of the fluke), but not the markings on the fluke (Arnbom 1987). Analyses were carried out with just the best photographs ($Q \geq 4$), as well as with moderate-quality photographs included ($Q \geq 3$). Fourteen members of the 1995 group being followed were measured using the photographic method of Gordon (1990).

To test for heterogeneity in identification rates between individuals, I needed independent data. Therefore, the photo-identification record was divided into 1-h, 2-h, 3-h, and 4-h periods, and the data sets consist of records of whether whale i was photo-identified in period j. I examined whether there was autocorrelation in the data sets, that is, whether the probability that an animal was identified in one period was independent of whether it was identified in the previous period. The data sets were permuted 100,000 times using the method of Manly (1995), so that, in each permutation, the number of animals identified during each period and the number of periods in which each animal was identified were both unchanged. The number of repeat identifications of the same animal in consecutive periods (*i.e.*, with a lag of 1 period) was calculated for the real data and each permuted data set, and these were compared to derive a significance level (the proportion of random permutations with more repeat identifications than in the real data). A large number of permutations are required for Manly's method, as they are not independent. Repetition showed that, with these data sets, 100,000 simulations gave quite stable P-values. To test for heterogeneity I used the shortest period length which did not show autocorrelation.

The null hypothesis that all members of each group had the same probability of being identified in any period was then tested using the likelihood ratio G-test statistic on the number of periods in which each animal was identified (Sokal and Rohlf 1981). However, as the data come from a series of binomial (identified or not identified within a period), not Poisson, processes, the standard Chi-squared distribution of the statistic under the null hypothesis is invalid. So, instead, the identifications within all periods were randomly permuted among members of the group (keeping the number of animals identified in each period constant, but, in this case, not the number of periods in which each animal was identified), and the actual G statistic was compared with the distribution of the statistic from 10,000 such permutations.

Carothers (1973*b*) showed that, given any sampling scheme, the effects of

heterogeneity on mark-recapture population estimates are well explained by the coefficient of variation in identification rates. Therefore, it is appropriate to quantify heterogeneity in identification using this coefficient of variation.

The total coefficient of variation in numbers of identifications among individuals, CV(I) is given by:

$$\mathrm{CV}(I)^2 = \mathrm{CV}(S)^2 + \mathrm{CV}(H)^2 \quad (1)$$

where CV(S) is the coefficient of variation due to sampling, and CV(H) the coefficient of variation due to heterogeneity in identification rates. But, $\mathrm{CV}(I)^2 = \mathrm{Var}(n_i)/\mathrm{mean}(n_i)^2$ where n_i is the number of periods in which animal i was identified. Furthermore, as the samples are assumed to be from independent binomial processes:

$$\mathrm{CV}(S)^2 = \left[\sum m_j/N \cdot (1 - m_j/N)\right] \Big/ \left[\sum m_j/N\right]^2$$

where m_j animals were identified in period j, and N is the group size.

Substituting these into equation 1 produces an estimate for the coefficient of variation in identification rates:

$$\mathrm{CV}(H) = \sqrt{\left[\mathrm{Var}(n_i)/\mathrm{mean}(n_i)^2 - \sum m_j/N \cdot (1 - m_j/N) \Big/ \left(\sum m_j/N\right)^2\right]} \quad (2)$$

If CV(S) > CV(I), there is less variation in identification numbers than expected from sampling alone and equation 2 gives an imaginary estimate of CV(H), which is best interpreted as a homogeneous data set (CV(H) = 0). Simulation showed that equation 2 (with CV(H) = 0 if CV(S) > CV(I)) leads to approximately unbiased estimators of CV(H) when the true value is in the range 0–0.8 (Fig. 1).

Confidence intervals for estimates of CV(H) were obtained using both bootstrap (producing 1,000 bootstrap replicates by resampling the identification records of members of the group with replacement) and jackknife (calculating pseudovalues of CV(H) by omitting individuals in turn) methods (see Efron and Gong 1983).

Results

The data sets from the 1999 group, and for the 1995 group with $Q \geq 4$, showed no autocorrelation over any time period from 1 to 4 h (Table 1). In contrast, there was statistically significant autocorrelation in the 1995 $Q \geq 3$ data set for periods of 1–2 h (Table 1). Therefore, for the heterogeneity analysis, I used 1-h periods, except for the 1995 $Q \geq 3$ data, which were divided into 3-h periods. The differences between the two groups in the results of these autocorrelation analyses are likely related to their sizes. Individuals within a group often spread out over a rank 1 km or so long, with individuals maintaining their approximate position within the rank over an hour or two (see Christal and Whitehead 2001). If, as seems probable, the research vessel in 1995 generally stayed to one side, or in the center, of the rank over these kinds of periods, an autocorrelation in the identifications would result. With

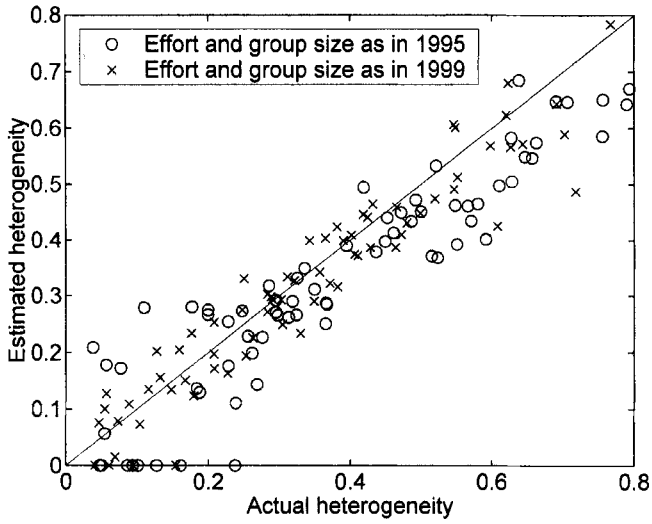

Figure 1. Results of simulation of identification process with randomly chosen photo-identification probabilities among animals, but with group sizes and effort (number of identifications in each period) as for 1995 and 1999 groups ($Q \geq 3$). For each simulated data set, estimated coefficient of variation in identification rates ("heterogeneity") from equation 2 is plotted against true coefficient of variation in simulated identification rates.

the smaller group (of 9 rather than 22 animals) followed in 1999, the rank seems to have been sufficiently short for this effect to have been insignificant. The differences in the autocorrelation analysis for the $Q \geq 3$ and $Q \geq 4$ in 1995 data sets are probably related to the considerable reduction in sample size resulting from the more stringent quality criterion (Table 2).

The group followed in 1999 were photo-identified in a mean of 42.1 1-h periods each (for all identification qualities with $Q \geq 3$) during the 18 days (Table 2). This allowed for powerful tests for heterogeneity in identification

Table 1. Results of permutation tests for autocorrelation when dividing the identification records into periods of different length, for 1995 and 1999 groups and two minimal identification qualities (Q). The P-values are the proportion of 100,000 random permutations with a greater number of repeat identifications of the same individual in consecutive periods than in the real data. P-values for two separate permutation tests for each data set are shown.

		Period length			
Group	Q	1 h	2 h	3 h	4 h
			P-values		
1995	≥3	0.047/0.051	0.022/0.016	0.677/0.675	0.409/0.414
1995	≥4	0.418/0.431	0.674/0.675	0.966/0.972	0.563/0.570
1999	≥3	0.571/0.594	0.857/0.978	0.909/0.830	0.398/0.395
1999	≥4	0.695/0.682	0.335/0.321	0.317/0.327	0.098/0.107

Table 2. Tests for, and estimates of, heterogeneity of identification rates within two groups of sperm whales, for the two groups and two minimal identification qualities (Q).

Group:	1995 (7 d)		1999 (18 d)	
Identification quality	$Q \geq 3$	$Q \geq 4$	$Q \geq 3$	$Q \geq 4$
Individuals	22	22	9	9
Period length used	3 h	1 h	1 h	1 h
Identifications	185	80	379	291
Permutation test for heterogeneity	$P = 0.058$	$P = 0.125$	$P = 0.479$	$P = 0.623$
Estimated heterogeneity, CV(H)	0.20	0.28	0.03	0.00
95% CI of CV(H) by:				
bootstrap	0.07–0.28	0.00–0.46	0.00–0.10	0.00–0.10
jackknife	0.09–0.30	0.07–0.49	0.00–0.09	0.00–0.01

rates. However, the results of tests, and estimates of the coefficient of variation in identification rates (CV(H)), all showed uniform identification rates for the members of this group, with little indication of heterogeneity (Table 2).

In contrast, despite less identification data, there was reasonably good evidence for heterogeneity for the larger 1995 group (Table 2). The permutation test for heterogeneity with all $Q \geq 3$ photo-identifications had a $P = 0.058$, and for just the higher quality, $Q \geq 4$, identifications, $P = 0.125$. The more precise estimate, using the $Q \geq 3$ identifications, suggests that the coefficient of variation in identification rates for this group was about CV(H) = 0.2 (bootstrap 95% confidence interval 0.07–0.28; Table 2).

One possible cause of the contrasting results for the two groups (suggested by an anonymous reviewer) is that, if fluking behavior varies over time for individuals, but all animals have the same long-term average, a shorter sample (as for the 1995 group) might pick up heterogeneity, but not a longer one (as in the 1999 group). I examined this by repeating the permutation tests for heterogeneity using just 4–6 day intervals within the 1999 data (10–15 March, 16–20 March, 28–31 March, 9–12 April). Of the eight tests for heterogeneity (four intervals with either $Q \geq 3$ or $Q \geq 4$), one was statistically significant at $P < 0.05$ ($P = 0.028$ for 10–15 March, $Q \geq 3$), while the other seven had $P > 0.15$. Therefore the differing results for the two groups do not seem to be a consequence of the difference in the sampling interval.

It was possible to look for potential sources of this apparent heterogeneity in the 1995 group. The group consisted of two long-term stable units with 5 and 17 members, respectively, each of which was identified separately in previous and/or subsequent years (Christal and Whitehead 2000). In 1995, despite the coherent movement of the two units as a single group, animals tended to affiliate with members of their own unit (Christal and Whitehead, in press), suggesting social segregation within the group. However, there was no indication of a significant difference between the identification rates of the two units (a mean of 11.6 identifications at $Q \geq 3$ for the members of the smaller unit, 12.1 identifications for members of the larger one; $t = 0.205$, df = 20, $P = 0.839$). If heterogeneity in identification rates was caused by differences in fluke morphology, then animals with more marked flukes should have higher rates of identification. However, among the 22 members of the 1995 group, there was no apparent relationship between the number of positions indicating marks on the trailing edge of a fluke (as defined in Whitehead 1990) and the number of $Q \geq 3$ identifications ($r_s = 0.072$, df = 20, $P = 0.380$, one-tailed test), although the number of such distinctive locations varied considerably (from 2–25). In contrast, the number of $Q \geq 3$ identifications of members of the 1995 group positively correlated with body length for the 14 measured members of the group ($r_s = 0.589$, df = 12, $P = 0.033$, two-tailed test; Fig. 2).

Discussion

Heterogeneity of identification is perhaps the most difficult issue facing those who attempt to analyze cetacean populations using photo-identifications.

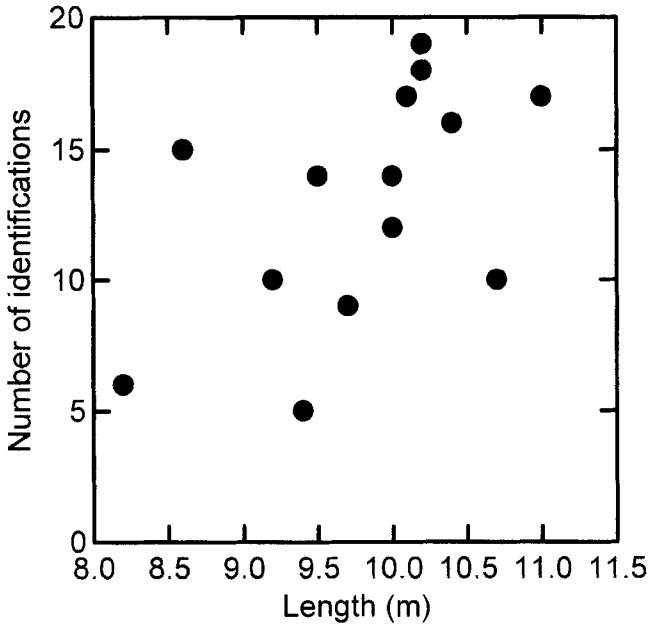

Figure 2. Number of identifications ($Q \geq 3$) of 14 members of sperm whale group followed in 1995 plotted against their photographically estimated lengths ($r_s = 0.589$, $P < 0.05$).

While some earlier studies ignored the issue, since the publication of Hammond's (1986) important review, there have been many attempts to investigate variability in identification rates either directly from photo-identification records (*e.g.,* Hammond 1990, Gowans and Whitehead 2001) or to look at its underlying causes in the animals' behavior (*e.g.,* Perkins *et al.* 1985, Rice *et al.* 1987) or morphology (*e.g.,* Carlson *et al.* 1990, Dufault and Whitehead 1995, Childerhouse and Dawson 1996, Blackmer *et al.* 2000, Friday *et al.* 2000). These attempts, while worthy and useful, are either indirectly aimed at heterogeneity or have little power to detect and deal with it. In contrast, photo-identification studies in which the population is closed and of known size can provide a powerful test for heterogeneity and examine its causes in some detail.

The two sperm whale groups examined gave contrasting results. The smaller group followed extensively in 1999 had highly homogeneous identification rates, with a coefficient of variation estimated to be nearly zero, and probably less than 10%. Unfortunately for those who use sperm whale photo-identifications in analyses of sperm whale populations, the results from the larger group were different. Despite smaller sample sizes, and thus reduced power to detect heterogeneity, the tests suggest that it was present in this group, with an estimated coefficient of variation in photo-identification rates among members of the group of about 0.2.

The analyses indicate that these differences in identification rates were not

due to underlying social structure, as the permanent units that constituted the group had very similar identification rates, or to the morphology of the fluke, as more-marked animals had similar identification rates to those with few marks. Instead, the length of the whale was a reasonably good predictor of its identification rate (Fig. 2). It seems that animals less than about 10.0 m, or about 15 yr of age (Christal and Whitehead 2000), had lower rates of identification than their elders. Personal observations of sperm whales suggest that this may be because the smaller animals sometimes dived without clearly showing their flukes, while older whales almost always raised their flukes as they dived. This pattern, of younger animals not showing their flukes as reliably, was found by Perkins *et al.* (1985) in their analysis of humpback whale (*Megaptera novaeangliae*) identification photographs from West Greenland. Members of our 1999 group, which showed no heterogeneity, were not measured, but there is no indication in the field notes from the tracking that any of the animals were particularly small. Furthermore, the principal scientists on this study report: "None of the animals struck us as being noticeably smaller than the others. . . [W]e were with the group so long and with consistently the same nine animals that by the end we were recognizing some of the animals by eye, so if any of the animals were small it would almost certainly have been noticed."[1]

This analysis had the potential to uncover effects from many, but not all, possible sources of heterogeneity in identification rates (including morphology and individual fluking behavior). In particular, attributes of different groups, such as movement patterns which seem to vary between groups (Whitehead 1999) and which might cause them to be more easily discovered, approached, or followed, could introduce heterogeneity at the level of the group, a particular concern for large-scale population analysis.

An important issue when planning and analyzing photo-identification studies is how long should be spent with a group of whales for the members of the group to have roughly equal probability of being available for identification. The autocorrelation analysis suggests that for a small group of about 10 sperm whales an hour is sufficient, but that for a larger group of about 20, some members of the group may not have been available for identification if it is followed for less than 3 h. Although all animals may be similarly available for identification during 3 h of tracking, it is usually preferable to spend much longer than this with a group of sperm whales in order to increase the overall photo-identification rate.

In conclusion, it seems from the analysis of these two groups that, in studies using sperm whale photo-identifications, the robustness of results should be checked against heterogeneity levels of about CV = 0.2. Consideration should also be given to a probable underrepresentation of younger animals. More generally, the method used here has a much greater power to detect hetero-

[1] Personal communication from L. Rendell, Department of Biology Dalhousie University, Halifax, Nova Scotia B3H 4J1, Canada and G. Merlen, Casilla 17-01-3891, Quito, Ecuador, 25 May 2000.

geneity, estimate its extent, and find its underlying causes, than the more piecemeal and indirect methods generally employed. However, it depends on a situation where there is a closed population of known size. In such circumstances, prolonged photo-identification effort may seem superfluous, but it is not. Exploitation of these rare situations is key to assessing the validity of the method in other contexts.

Acknowledgments

I am grateful to all those involved with the field work, but particularly to Jenny Christal, Godfrey Merlen, and Luke Rendell for their large roles in the collection and analysis of the photo-identification data. This study was carried out with the assistance and collaboration of the Armada of Ecuador, the Charles Darwin Station, the Galápagos National Park Service, and the Instituto Nacional de Pesca. It was funded by the Natural Sciences and Engineering Research Council of Canada. Shannon Gowans, Andrea Ottensmeyer, and two anonymous reviewers gave useful reviews of manuscripts.

Literature Cited

Arnbom, T. 1987. Individual identification of sperm whales. Report of the International Whaling Commission 37:201–204.

Blackmer, A. L., S. K. Anderson and M. T. Weinrich. 2000. Temporal variability in features used to photo-identify humpback whales (*Megaptera novaeangliae*). Marine Mammal Science 16:338–354.

Carlson, C. A., C. A. Mayo and H. Whitehead. 1990. Changes in the ventral fluke pattern of the humpback whale *(Megaptera novaeangliae),* and its effects on matching; evaluation of its significance to photo-identification research. Report of the International Whaling Commission (Special Issue) 12:105–111.

Carothers, A. D. 1973*a*. Capture-recapture methods applied to a population with known parameters. Journal of Animal Ecology 42:125–146.

Carothers, A. D. 1973*b*. The effects of unequal catchability on Jolly-Seber estimates. Biometrics 29:79–100.

Childerhouse, S. J., and S. M. Dawson. 1996. Stability of fluke marks used in individual photoidentification of sperm whales at Kaikoura, New Zealand. Marine Mammal Science 12:447–451.

Christal, J., and H. Whitehead. 2000. A week in the life of a sperm whale group. Pages 157–159 *in* J. Mann, R. C. Connor, P. L. Tyack and H. Whitehead, eds. Cetacean Societies. University of Chicago Press, Chicago, IL.

Christal, J., and H. Whitehead. 2001. Social affiliations within sperm whale (*Physeter macrocephalus*) groups. Ethology:107–340.

Clapham, P. J., and C. A. Mayo. 1990. Reproduction of humpback whales *(Megaptera novaeangliae)* observed in the Gulf of Maine. Report of the International Whaling Commission (Special Issue) 12:171–175.

Cormack, R. M. 1985. Examples of the use of GLIM to analyse capture-recapture studies. Pages 243–273 *in* B. J. T. Morgan and P. M. North, eds. Statistics in ornithology. Springer-Verlag, Berlin.

Dufault, S., and H. Whitehead. 1995. An assessment of changes with time in the marking patterns used for photo-identification of individual sperm whales, *Physeter macrocephalus.* Marine Mammal Science 11:335–343.

Efron, B., and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. The American Statistician 37:36–48.

Friday, N., T. D. Smith, P. T. Stevick and J. Allen. 2000. Measurement of photo-

graphic quality and individual distinctiveness for the photographic identification of humpback whales. Marine Mammal Science 16:355–374.

GORDON, J. C. D. 1990. A simple photographic technique for measuring the length of whales from boats at sea. Report of the International Whaling Commission 40:581–588.

GOWANS, S. AND H. WHITEHEAD. 2001. Photographic identification of northern bottlenose whales (*Hyperoodon ampullatus*): Sources of heterogeneity from natural marks. Marine Mammal Science 17:76–93.

HAMMOND, P. S. 1986. Estimating the size of naturally marked whale populations using capture-recapture techniques. Reports of the International Whaling Commission (Special Issue) 8:253–282.

HAMMOND, P. S. 1990. Heterogeneity in the Gulf of Maine? Estimating humpback whale population size when capture probabilities are not equal. Report of the International Whaling Commission (Special Issue) 12:135–139.

INTERNATIONAL WHALING COMMISSION. 1990. Report of the Workshop on Individual Recognition and the Estimation of Cetacean Population Parameters, La Jolla, 1–4 May 1988. Report of the International Whaling Commission (Special Issue) 12: 3–40.

LESLIE, P. H. 1958. Statistical appendix. Journal of Animal Ecology 27:84–86.

MANLY, B. F. J. 1995. A note on the analysis of species co-occurrences. Ecology 76: 1109–1115.

MANN, J. 2000. Unraveling the dynamics of social life: Long-term studies and observational methods. Pages 45–64 *in* J. Mann, R. C. Connor, P. L. Tyack and H. Whitehead, eds. Cetacean societies. Field studies of dolphins and whales. University of Chicago, Chicago, IL.

PERKINS, J. S., K. C. BALCOMB, G. NICHOLS, A. T. HALL, M. SMULTEA AND N. THUMSER. 1985. Status of the West Greenland humpback whale feeding aggregation, 1981–1983. Report of the International Whaling Commission 35:379–383.

RICE, M., C. CARLSON, K. CHU, W. DOLPHIN AND H. WHITEHEAD. 1987. Are humpback whale population estimates being biased by sexual differences in fluking behavior? Report of the International Whaling Commission 37:333–335.

SOKAL, R. R., AND F. J. ROHLF. 1981. Biometry. 2nd Edition. W. H. Freeman, San Francisco, CA.

STONE, G. S., L. FLÓREZ-GONZALEZ AND S. KATONA. 1990. Whale migration record. Nature 346:705.

WHITEHEAD, H. 1990. Computer assisted individual identification of sperm whale flukes. Report of the International Whaling Commission (Special Issue) 12: 71–77.

WHITEHEAD, H. 1999. Variation in the visually observable behavior of groups of Galápagos sperm whales. Marine Mammal Science 15:1181–1197.

WHITEHEAD, H., AND S. DUFAULT. 1999. Techniques for analyzing vertebrate social structure using identified individuals: Review and recommendations. Advances in the Study of Behavior 28:33–74.

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., J. CHRISTAL AND S. DUFAULT. 1997. Past and distant whaling and the rapid decline of sperm whales off the Galápagos Islands. Conservation Biology 11:1387–1396.

Received: 7 June 2000
Accepted: 28 December 2000