Manuscript Number:

Title: Efficacy of modified observational sampling regimes for use in zoological parks

Article Type: Research Paper

Section/Category:

Keywords: observational samping; scan sampling; instantaneous sampling; activity budget

Corresponding Author: Dr. Susan W. Margulis, PhD

Corresponding Author's Institution: Lincoln Park Zoo

First Author: Susan W. Margulis, PhD

Order of Authors: Susan W. Margulis, PhD; Eric Westhus, M.S.

Manuscript Region of Origin:

Abstract:

Efficacy of Modified Observational Sampling Regimes for Use in Zoological Parks

Susan W. Margulis [1,2,3] and Eric Westhus[1]

[1] Chicago Zoological Society, 3300 Golf Road, Brookfield IL 60513 USA

[2] University of Chicago, Committee on Evolutionary Biology, Chicago, IL 60637 USA

[3] Lincoln Park Zoo, 2001 N. Clark St., Chicago IL 60614 USA

Direct all correspondence to:
Susan Margulis
Curator of Primates
Lincoln Park Zoo
2001 N. Clark St.
Chicago, IL 60614 USA
Phone: (312) 742-2345
Fax: (312) 744-4739
Email: smargulis@lpzoo.org

Abstract

Observational sampling methods provide clearly-defined guidelines for collection and analysis of behavioral data. In some situations, use of formal sampling regimes may be impractical or impossible. A case in point is data collection conducted by animal care staff at zoological parks and aquaria. Often, time is sufficiently limited that data collection is perceived as a task that cannot be accomplished given the normal constraints of the day. Here, we explore the efficacy and validity of using more variable and abridged sampling regimes in an effort to identify the appropriateness of such observation schemes for systematic monitoring of behavior. We describe the results of studies on three species (polar bears, Atlantic bottlenose dolphins, and brown bears), conducted over a period of several years at the Brookfield Zoo. Data collection schemes varied both within and across groups in order to provide a basis of comparison. In all cases, differences in overall activity budgets either did not differ or differed only slightly based on sampling regimes that ranged from continuous sampling, to sporadic point observations throughout the day. The sporadic sampling regimes provided a fairly accurate activity budget for relatively common behaviors, but not for rare behaviors (ie, behaviors that each comprise <10% of the time budget). We discuss the strengths and weaknesses of such sporadic sampling methods, and suggest that, in many instances such limited data collection may yet yield a fairly accurate picture of animal activity and should not be overlooked as a viable management tool.

Key words: observational sampling; scan sampling; instantaneous sampling; activity budget

Introduction

Observational sampling methods are designed to allow a researcher to collect representative, unbiased information about animal behavior and activity patterns. Various methodologies exist for collecting such information. The two most basic and widely used methods are instantaneous, or scan sampling, and continuous sampling (Altmann 1974; Martin and Bateson 1993). In the former, behavioral observations are made at pre-set intervals for some predetermined period of time (for example, every minute for 1 hour, every 10 seconds for 5 minutes). The latter method requires the observer to watch an animal for a predetermined period of time, and record the onset times of all behaviors. The two methods allow the observer to collect different types of behavioral information, and address different research questions. Each method has advantages and disadvantages. For example, instantaneous sampling can often be an efficient way to collect activity budget information, but it is not appropriate for observing detailed interactions. Continuous sampling allows for more detailed data collection, but often requires lengthier dedicated blocks of time and the data may be more problematic to analyze.

The method chosen and the time frame over which observations are made are designed to allow for representative data collection. However at times, the availability of observers, husbandry practices, and scheduling issues may influence the data collection regime. In some cases, this issue can be dealt with statistically, for example by including time of day as a covariate. In other situations, such sampling issues can be addressed via interpreting the data in light of these discrepancies (i.e., describing a picture of behavior limited to a specific time frame rather than as a broad generalization). In some circumstances however, it may be difficult to impose even this level of methodological or analytical control. For example, for staff at zoological institutions, it may often be impractical to evenly distribute observations throughout the day, to conduct lengthy observation sessions, or to adhere closely to specific data collection regimes. Staff may have only a limited amount of time to devote to behavioral data collection, and may have difficulty adhering to a rigid data collection protocol.

Although methodological comparisons have been done in the past, much of the work that has been done in this area does not specifically address issues associated with tracking behavioral changes using data that are collected in small quantities, on an opportunistic basis rather than on a rigorously set schedule, and outside the scope of a specific experimental design. A number of authors have compared the efficacy of different observation methods; for example, continuous versus instantaneous sampling or one-zero versus instantaneous sampling (Rhine and Ender, 1983; Suen and Ary, 1984). Kraemer et al. (1977) describe a method of estimating the costs of various sampling procedures and designing a cost effective analysis by plotting those costs against an error estimate associated with each sampling procedure. Bakeman et al. (1992) compare log-linear analysis to ANOVA for scan data collected during scheduled sessions under a rigorous experimental design. Damerose and Hopkins (2002) compare scan and focal sampling methods for a very specific application and do not explore the change in reliability of scan sampling with changing session number or interval length. Hepworth and Hamilton (2001) provide a method of handling extremely large groups of subjects, the exact opposite problem from what is usually encountered in zoo settings. Saudargas and Drummer (1996) discuss single subject, repeated measures experimental design on a conceptual level, but do not get into the nuts and bolts of sampling regimes.

Here, we contribute to the assessment of methodological application by using three distinct datasets as examples to allow us to explore the extent to which variation is sampling regime may be incorporated into a study without significantly biasing results. In the first example, we had the opportunity to collect basic activity data on two sub-adult polar bears (*Ursus maritimus*) following several different data collection methodologies and schedules. We compare the results of these overlapping studies, each using a different sampling regime, with respect to the time budgets of the two bears. The second example utilizes data collected on brown bears (*Ursus arctos*) first as a keeper-initiated study using infrequent observations over a period of years, and later as part of a behavioral monitoring project. Our analysis of our brown bear data described below draws heavily on the procedures described by Kraemer et al (1977). We first compare the long-term data with the monitoring data, and then sub-sample from the monitoring data set with gradually increasing intervals between point samples. Finally, we describe data collected during the first month of life of an Atlantic bottlenose dolphin calf (*Tursiops truncatus*). As with the brown bear example, we take frequent, intensive data collection and work backwards to determine the extent to which sample interval can be extended and yet still provide an accurate representation of behavior.

Our goal is to define the limits of scientifically accurate, unbiased behavioral data collection in an effort to describe a more flexible, yet rigorous method that is appropriate for use in zoological parks.

Methods

POLAR BEARS

Subject and data collections: The subjects of this study were two sub-adult polar bears. The female, Tiguak, was born at Brookfield Zoo in November, 1999 and hand-reared. The male, Eddy, was born at the Aquarium du Quebec in December 1999 and mother-reared until one year of age. The two bears were introduced to each other in March, 2001.

We initiated a study to monitor behavioral changes in the two bears following the introduction. Coincidentally, several other ongoing studies enabled us to collect comparable data. Ultimately, data were collected following three distinct observation schedules and protocols: (1) the primary data collection regime, designed to assess changes in behavior following introduction (the "observer" regime), (2) a second data collection schedule, capitalizing on another ongoing study in which polar bears were videotaped every other weekend (the "taper" regime), (3) a keeper-based data collection plan, in which point samples were collected multiple times per day weekly at varying times throughout the day (the "keeper" regime).

Methodology One was designed to evaluate the changes in behavior in the two bears following their introduction. We used scan sampling to record the behavioral state of the two bears, as well as their proximity, every minute for 30 minutes (see table I). In addition, all-occurrence sampling was used to record instances of interactions (Altmann 1974). Data were collected during 1 of 3 time periods: 1000-1200 h, 1200-1400 h, or 1400-1600 h. Data were collected for 11 weeks prior to the introduction of the two bears and 11 weeks following their introduction, for a total of 62 hrs of observation. Here, we include only data collected following the introduction of the two bears (11 weeks, a total of 32 hours), March-May, 2001, since all other data collection regimes were initiated following the introduction. Data were collected by trained volunteers who had reached a level of proficiency in the behavioral methodology. Observer reliability observations were conducted and all observers reached a level of at least 85% agreement. Data collection method Two occurred concurrently with method 1, but utilized another pool of volunteer observers. These observers were videotaping the behavior of two other polar bears (an adult male and an adult female) on alternate weekends as part of a multi-institutional study on polar bear behavior (Shepherdson and Carlstead, 2001). The protocol for this study involved videotaping for 30 seconds every 10 minutes from 1000-1600 hrs. In an effort to make these lengthy observation periods more interesting for the observers, we asked them to record the behavior of Eddy and Tiguak, using the same protocol as for the primary study (method One), every 10 minutes (i.e., after each 30-second taping period). We did not, however, ask them to record all-occurrence events, only activity states and proximity. This data set then, comprised 11 observation periods during which a point sample was collected every 10 minutes for 6 hours, from April through August 2001. The third methodology utilized observations made by keeper staff. As part of an ongoing program to encourage the collection of baseline behavioral data, keepers were asked to record the activity state of the 2 bears periodically throughout the day once a week. Because the keeper staff routinely walks past the polar bears during the course of their normal routine, we wanted to evaluate how accurate data collected at less-regular intervals would be for evaluating basic time budgets. We asked keepers to try and collect data approximately every hour during their workday (approximately 0800-1700). The average inter-scan interval was 67±11.6 minutes. The same ethogram of activity states was used, although we did not ask keepers to record proximity between the 2 bears. A total of 22 keeper observations were used for analysis. These data spanned the period April through October, 2001.

Data analysis: The complete data set then, consisted of activity states for the 2 polar bears during the period March through October, 2001. Time budgets for the 2 bears were calculated using the 3 different data sets. Because data were not normally distributed, nonparametric analyses were used to compare the 3 time budgets. We calculated the percent time spent in each behavior per observation session, and then performed Kruskal-Wallis tests to compare the different collection regimes for each behavior, and for all “active” behaviors combined. Active behaviors included all behaviors except rest and out of view.. An α of 0.0035 was set for each comparison to hold the experimental wide error rate at an α of 0.05. Mann-Whitney U comparisons were used for post-hoc analysis of the significant Kruskal-Wallis comparisons.

BROWN BEARS

Subject and data collections: We conducted intensive observational sampling on 2 male brown bears during the summer of 2003. Between May 27 and June 19, 71 observations sessions were conducted by a single observer. During each 15-minute session, point samples were collected every 30 seconds, at which time the activity of each of the two bears (see table 2) were recorded using a Palm Pilot IIIXe. Due to some occasional discrepancies in session length, some observations included 28 (rather than 30) scans, therefore the raw dataset was limited to first 28 scans per sessions. Data were then synchronized into a behavioral monitoring database. Sessions were labeled chronologically 1 through 71, and scans within each session were labeled 1 through 28. The percent scans recorded for each behavior was calculated for each of the 71 sessions using all 28 scans.

Data analysis: To assess the variability in time budget calculations based on sampling interval, the 28 scans (at 30 second time intervals) from each session were sub-sampled for 1, 5, 10, and 15 minute time intervals. The 1-minute interval made use of every other scan. Two iterations were possible for this time interval: one iteration using every other scan starting with scan #1, the other starting with scan #2. The 5-minute interval resulted in eight possible iterations, using every 10th scan, and starting on each of the first eight scans, respectively. The 10-minute time interval also resulted in eight possible iterations using every 20th scan, starting on each of the first eight scans. The 15-minute time interval used just one scan from each session, this resulted in 28 possible iterations for each session.

Our first analysis of these data used ANOVA to examine the differences among iterations within each time interval to determine if there were significant differences based on which iteration was used, or if a single iteration could be used to represent that time interval. Next, we used ANOVA to test for differences in time budget among the time intervals.

Our third series of analyses utilized a method similar to that of Kraemer et al. (1977) to plot error against the “cost” of sampling. The equation used to calculate error followed Kraemer et al (19787), $\sqrt{}(S^2+\overline{B}^2)/ y$, where $\overline{B}^2=\overline{x}— y$. However, we assumed an even increase in cost with each additional session and scan rather than using their more complex method of estimating cost increases. We used this approach to evaluate how the quantity of data collected influences the variability of the resulting time budget calculations. The first set of error vs. cost analyses explored the changes in error associated with increasing the number of sessions from 5 to 60 and utilized all 28 scans per session. The dataset was sub-sampled for increasing increments using the following protocol: we randomly chose 5 of the 71 sessions and used only those five sessions to calculate the average percent-scans for each behavior (x_1) in the time budget. We repeated this procedure for 20 iterations, with none of the twenty iterations repeated, to obtain x_1 through x_{20}. We then calculated the average ($\overline{x}$) and variance (S^2) of the average percent-scans over all 20 iterations (all x_i’s) for each behavior. Those values were used to calculate the error estimate described above. This process was repeated using twenty iterations each for sub-samples of 10, 20, 30, 40, 50 and 60 sessions, and the resulting error was plotted against the increasing session numbers.

The second set of error vs. cost analyses added a third variable to the analyses. Whereas the first set of Kraemer error calculations utilized all 28 scans, we wanted to assess the impact of lengthening the inter-scan interval. We again calculated the error but with increasing scan interval lengths of 1, 5, 10 and 15 minutes. This third dimension was explored by repeating the same 5 through 60 session sub-sampling procedures outlined above for each of the scan interval categories used in the second set of ANOVA analyses (1, 5, 10, and 15 minute intervals).

Our final analysis compared the data collected during the intensive summer observation period with data collected over an extended period of time by the keepers as part of an ongoing investigation of seasonal changes in behavior (Mazrimas-Ott, 2004). Like the polar bear data previously described, keepers had been collecting behavioral data in the same manner as was used for the polar bear observations. These observations were begun in 1998. Keepers collected point-observations intermittently throughout the day (on average, 11 observations per day, range 3-17) one to three days per week. From 1998-2002, 350 observation-days were recorded. We therefore compare the results of the intensive observations conducted during the summer of 2003 (and described in more detail above) with the less regular data collected over an extended period of time, from 1998-2002, using a Kruskal-Wallis test.

ATLANTIC BOTTLE NOSE DOLPHINS

Subject and data collections: On October 19, 2001 a dolphin calf was born at Brookfield Zoo. The standard observation protocol has always been to monitor the calf continuously (or nearly so) for the first month. Two observers are used simultaneously for most this one month period. One observer remains stationed near an underwater viewing window, with stopwatch and clipboard. Every minute, this observer records whether or not the calf is in a slipstream position. Slipstreaming is a behavior in which a calf “rides” in its mother’s slipstream, thereby reducing energy expenditure needed to maintain continuous movement and remain in proximity to its mother. This is therefore a critical behavior to track during the early days of a calf’s life. A second observer with stopwatch and radio moves throughout the underwater viewing area in order to observe all nursing bouts. The second observer times these and relays this information to the first observer, who records it on the data sheet. With this information, nursing rate (minutes/hour and bouts/hour) can be calculated.

Data analysis: In an effort to reduce the intensity of observational sampling applied to dolphin calves, we used the full data-set, consisting of over 14,000 nursing bouts and over 43,500 point samples on slipstreaming, and sub-sampled from it at varying intervals to determine at what point the validity of the data would begin to degrade. The data were broken up into 6-hour blocks, or shifts (4 shifts per day), based on the standard protocols in place for assessing dolphin calf well-being. Data were available for 4 shifts per day for 28 days, 3 shifts per day for the first day and final 2 days, for a total of 121 shifts, from October 19-November 18, 2001. Dolphin calf data were analyzed using ANOVA on square-root transformed data when data were not normally distributed. We assessed differences in nursing rate (mintues/hour) and percent time spent in slipstream. Independent variables included Day (1-31), Shift (1=midnight-0600; 2=0600-1200; 3=1200-1800, 4=1800-midnight), and sampling scheme (1, 2, 3, or 6 hours of data per shift).

Results

POLAR BEAR

There were significant differences among the data collection regimes for some, but not all, of the recorded behaviors (figure 1). When the two bears were analyzed separately, we found significant differences in the percentage of time spent by Eddy in locomotor behavior (Kruskal-Wallis test, χ^2 =14.731, df=2, P<0.0006) and out of view (χ^2 =32.52, df=2, P<0.0001), and in time the two bears spent engaged in social interaction (χ^2 = 15.62, df=2, P<0.0004). Post-hoc analyses indicate that, while there were differences among the collection regimes, these differences were not consistent. The Observer regime was significantly different from the Taper and Keeper regimes for swimming, locomotion, rest, and foraging, while the Keeper regime differed from the Observer and Taper regimes for social interaction.

BROWN BEAR

Figure 2 shows some representative examples of the differences among iterations within a scan interval. We found no significant differences among the various iterations within any time interval category, therefore for the remaining calculations all comparisons were made using a single iteration only (the iteration beginning with scan #1). The next set of analyses evaluates the impact of increasing number of observation sessions on the variability in the resulting time budget analyses. As number of sessions increases, error (not surprisingly) decreases. For common behaviors (those that encompass at least 10% of the time budget), the error is fairly low even after a small number of observations, and the error does not decrease notably as number of observation sessions increases. Rarer behaviors however, have higher levels of error initially and respond more noticeably to increases in observation number. Representative examples of the error results based on different sampling intervals are presented in figure 3. Adding a third dimension to these analyses (figures 4 and 5) illustrates more clearly the relationship between scan interval, number of observations, and error, for rare and common behaviors.

Comparison of time budget data collected via point sampling throughout the day (by keepers) and intensive point sampling by a single observer revealed that, for behaviors that are relatively common (i.e, occupying >10% of the time budget), there were no significant differences in behavior (figure 6). This was the case for rest and swimming (see table 3).

Behaviors that occurred less frequently in one or both of the data collection regimes were significantly different.

DOLPHIN

While day and shift had significant effects on nursing rate (ANOVA, $F_{30,485} = 28.23$, $P<0.0001$ for day; $F_{33,483}=7.81$, $P<0.001$ for shift), sampling scheme did not ($F_{3,483}=0.69$, $P<0.56$) (figure 7). Similarly, for percent time spent in slipstream position, day and shift had significant effects ($F_{30,483}=28.01$, $P<0.0001$ for day; $F_{3,483}=12.85$, $P<0.0001$ for shift), and sampling scheme did not ($F_{3,483}=0.09$, $P<0.96$) (figure 8).

Discussion

In an ideal observation world, an observer would monitor animals continuously, 24 hours a day, thereby creating a true and accurate picture of their behavior, interactions, and time budgets. Obviously, this is neither practical nor feasible. A less ideal, but also quite effective means of accurately monitoring behavior is to sample behavior at regular intervals for significant chunks of time and use this subset of behavior to estimate true behavior patterns. Indeed, some form of standard, observational sampling is used in virtually all behavioral research. The specific method chosen will invariably depend on the observation conditions, the research question, and the available resources (eg., Kramer et al 1977). Some variation of formal observational sampling is the best option with respect to accuracy of data and efficiency of data collection. However, in situations in which research or monitoring of behavior encompasses only a small part of an observer’s daily time budget, more creative and flexible methods of observation need to be developed. This may necessitate some loss of accuracy, but in certain situations this trade off is warranted. Such is often the case in zoological parks and aquaria, in which regular data collection is often done by staff or volunteers whose time is limited.

Here, we have demonstrated that, although such limited and varied sampling regimes cannot replace more formal observational methods, they can however provide a fairly good measure of gross activity levels. Further, deviations from presumed “accurate” behavior may be explainable and therefore controllable, either statistically (ie, by excluding rare behaviors from analysis) or by interpretation (keeper observers are likely better at spotting animals and animals may be more likely to respond with activity to keepers).

The examples described here represent three very different scenarios for behavioral data collection. In the case of the polar bears, a fortuitous situation resulted in data collection using three different observation regimes. Differences were detectable, but surprisingly minor. Given the great disparity in data collection regimes—ranging from formal, 30 minute observations multiple times a week to sporadic point samples throughout the day—the degree of similarity in time budget calculations is quite astounding. Furthermore, the nature of the management of the bears may explain some of the difference. For example, a keeper leaving the inside service area to come outside for an observation will likely be detected by the bears, who may respond by increasing locomotor activity or returning to the exhibit from an off-exhibit den. In contrast to the polar bear data, the dolphin data was collected at a much finer level of detail, with data collected continuously for a month. We found no significant differences when the observation scheme was reduced by as much as 80% (from 24 hours a day to 4 hours a day). This finding would suggest that data collection intensity can be drastically reduced without compromising data integrity.

Perhaps the most informative and most complex data comparison involves the data collected on the two brown bears. We first confirmed that the internal variation of iterations within a sampling regime was insignificant and thus the main source of variation could be attributed to differences among the different sampling regimes. As scan interval increases, the variation of the resulting time budget results increases, but the extent of the increases are generally small. This is particularly the case for common behaviors (ie, those that make up at least 10% of the overall time budget). This should not come as surprise, since the probability of observing a common behavior, even when sampling interval is infrequent, should be relatively high. In contrast, the error rate is greater for rare behaviors. Thus, a large scan interval may yield fairly accurate results for the more common behaviors in a time budget, but risks over- or under-estimating the rates of more rare behaviors. This error however, has been found to decrease as the number of observation sessions increases. These results support our contention that even infrequent observations, collected over an extended period of time, may yet yield valid and appropriate results. In fact, the error rates for even rare behaviors become virtually indistinguishable from those of more common behaviors, by 50 observation sessions.

We do not argue that sporadic point observations are appropriate for all types of observational studies or for all types of behaviors. Rather, we suggest that infrequent and brief monitoring of behavior can serve as a valuable management tool to track basic activity patterns, identify changes in behavior, and serve as a starting point for more focused observational studies (Margulis and Pruett-Jones, in press). The typical animal care staff person is a highly-trained, well-educated professional with experience and expertise in animal observation. Time is often the limiting factor in the conduct of any formal research. By affirming that even infrequent observation can provide useful insights into long-term patterns of behavior and behavioral changes, we encourage staff to collect behavioral data. Even very limited amounts of data—as little as 5-10 point samples a day—can, over time, provide invaluable information for future research, management, and husbandry.

Acknowledgments

We thank the staff at Brookfield Zoo for their assistance with the data collection projects described here. In particular, we thank curators Melinda Pruett-Jones and Ann Petric, and assistant curators Greg Dye, Mary Burke, and Joan Daniels for their cooperation and support. The keepers in Fragile Kingdom and Seven Seas collected most of the data described here. We are especially indebted to Christy Mazrimas-Ott, Jennifer McGee, and Rita Stacey. Numerous volunteers assisted with data collection, volunteer scheduling, and data entry. In particular we thank Meegan Anderson, Linda Reiter, Tina Weber Murack and Kevin Murphy. We would like to acknowledge Allison Walsh for assistance with data management and Gerardo Camilo for helping to formulate the initial roadmap for these analyses.

References

Altmann, J. 1974. Observational study of behavior: sampling methods. Behaviour 49:227-267

Bakeman, R., Forthman, D.L., and Perkins, L.A. 1992. Time-budget data: log-linear and analysis of variance compared. Zoo Biology, 11: 271-284.

Damerose, E. and Hopkins, W.D. 2002, Scan and focal sampling: reliability in the laterality for maternal cradling and infant nipple preferences in olive baboons, *Papio anubis*. Animal Behavior, 63: 511-518.

Hepworth, G. and Hamilton, A.J. 2001. Scan sampling and waterfowl activity budget studies: design and analysis considerations. Behavior, 138: 1391-1405.

Kraemer, H.C., Alexander, B., Clark, C., Busse, C., and Riss, D. 1977. Empirical choice of sampling procedures for optimal research design in the longitudinal study of primate behavior. Primates, 18: 825-833.

Margulis, S.W. and Pruett-Jones, M. In Press. Integrating Science and Husbandry: Less is More

Martin, P. and Bateson, P. 1993. Measuring Behaviour: an Introductory Guide (second edition). Cambridge University Press, Cambridge. 222 pages.

Mazrimas-Ott, C. 2004. Investigating hormonal correlates of seasonal stereotypic swimming in a male Alaska brown bear (*Ursus arctos*). *Animal Keepers' Forum*, 31: 194-202.

Rhine, R.J. and Ender, P.B. 1983. Comparability of methods used in the sampling of primate behavior. American Journal of Primatology 5: 1-15.

Saudargas, R.A. and Drummer, L.C. 1996. Single subject (small N) research designs and zoo research. Zoo Biology, 15:173-181.

Shepherdson, D. and Carlstead, K. 2001. New approaches to the evaluation of zoo animal well-being using multiple institutions, assessment of behavior and temperament, and non-invasive physiological measures. Proceedings, American Zoo and Aquarium Association Annual Conference, St. Louis, MO, pp. 131-136.

Suen, H.K., and Ary, D.1984. Variables influencing one-zero and instantaneous time sampling outcomes. Primates, 25: 89-94.

Table 1. Ethogram for polar bear study

BEHAVIOR	**DEFINITION**
STATES	
Stereotypy	Repetitive movement from which the animal is not readily distracted
Enrichment use/object play	Use of toys provided by staff
Swim	Locomotion in water
Locomote	Terrestrial movement
Rest	Animal is inactive
Feed	Acquisition, manipulation, or consumption of food
Social play	Non-aggressive interaction with conspecific
Out of view	Animal is in off-exhibit den
ALL-OCCURRENCE BEHAVIORS	
Approach	One animal moves from a distance greater than 3 m to a distance of less than 3 m from the other animal
Displace	One bear moves away as the other approaches
Behaviors associated with aggression	Aggression is defined as escalation of play, associated with more rapid interactions and vocalizations
• Growl	Gutteral vocalization, often associated with escalation of play
• Rear	Animal stands on rear legs during escalation of play
• Swat	Animal slaps or swipes at another during escalated play
Other vocalization	Sounds produced other than growl

Table 2. Ethogram for brown bear study.

BEHAVIOR	**DEFINITION**
Stereotypic swim*	Repetitive circular swimming in which the animal traces the same route, pushes off wall in same location, and is not readily distracted
Enrichment use/object play	Use of toys provided by staff
Swim	Non-repetitive locomotion in water
Locomote	Terrestrial movement
Rest	Animal is inactive
Feed	Acquisition, manipulation, or consumption of food
Social play	Non-aggressive interaction with conspecific
Out of view	Animal is in off-exhibit den

*Stereotypic swim was recorded as part of the long-term keeper-based observations and not during the summer 2003 observations. Stereotypic swim and swim were combined for analyses that compared the intensive and long-term observations.

Table 3. Summary of results, polar bear data

behavior	Keeper scan %	Observer scan %	Kruskal-Wallis test
Rest	39.8	41.0	0.641
Swim	12.8	20.6	1.495
Stereotypy	15.5	13.4	0.889
Forage	13.2	3.1	**85.14**
Locomote	9.0	12.6	**8.523**
Enrichment use	5.8	8.3	**4.098**
Social	3.9	0.5	**21.106**
Aggressive	0.0	0.6	---

Figure legends.

Figure 1. Activity budgets for two polar bears based on three different data collection regimes. Bar represent means and standard errors.

Figure 2. Representative activity budgets for brown bears. The upper graph represents data collected for one bear at 10-minute intervals, and the lower graph represents data collected for the other bear at 5-minute intervals. For both bears, the complete data set is based on point-sampling at 30-second intervals, then was resampled at a range of intervals (from 30 seconds to 15 minutes). Multiple possible iterations could be used for such a comparison—for example, minutes 1, 11, 21 or 2, 12, 22 and so on for the 10-minute intervals. In all cases, we found no significant effect of iteration on time budget.

Figure 3. Error (as defined by Kraemer et al. 1977) for all behaviors by increasing session number with 5 minute scan interval for one of the bears (upper figure) and 15 minute scan interval for the second bear (bottom figure). Although only two example graphs are presented here, for all behaviors, error decreases as the number of sessions increase and error increases as scan interval increases.

Figure 4. Error (Kraemer et al. 1977) by scan interval and session number for behavior Swimming for one of the two bears. Swimming was one of the more common behaviors. This graph shows that even at lower session numbers and longer scan intervals, error is relatively low. Error rapidly declines as session number increases and scan interval decreases.

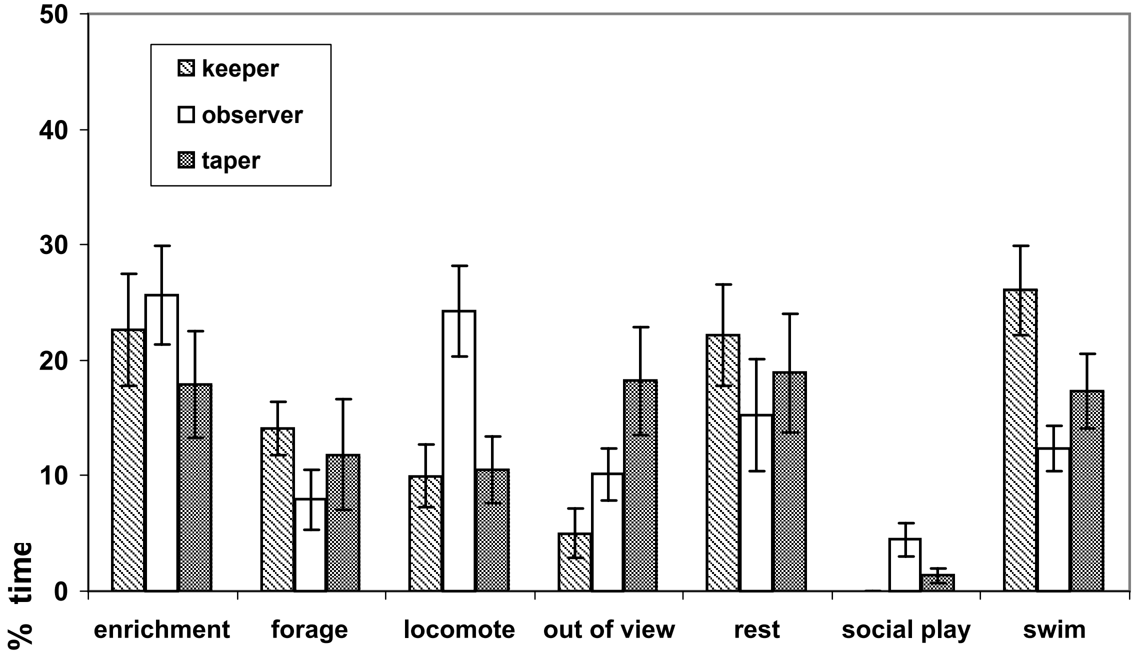
Figure 5. Error (Kraemer et al. 1977) by scan interval and session number for behavior Forage for the second bear. Foraging was an uncommon behavior, so its error is more variable at low session number and longer scan intervals than for the common behavior Swimming.

Figure 6. Comparison of activity budget for one brown bear based on intensive (hatched bars) and sporadic (open bars) observations. Intensive data collection involved recording behavior every 30 seconds for 15-minute observation blocks. Sporadic observations were based on a varying number of data points (average 11 points per day, range 3-17) collected throughout the day.

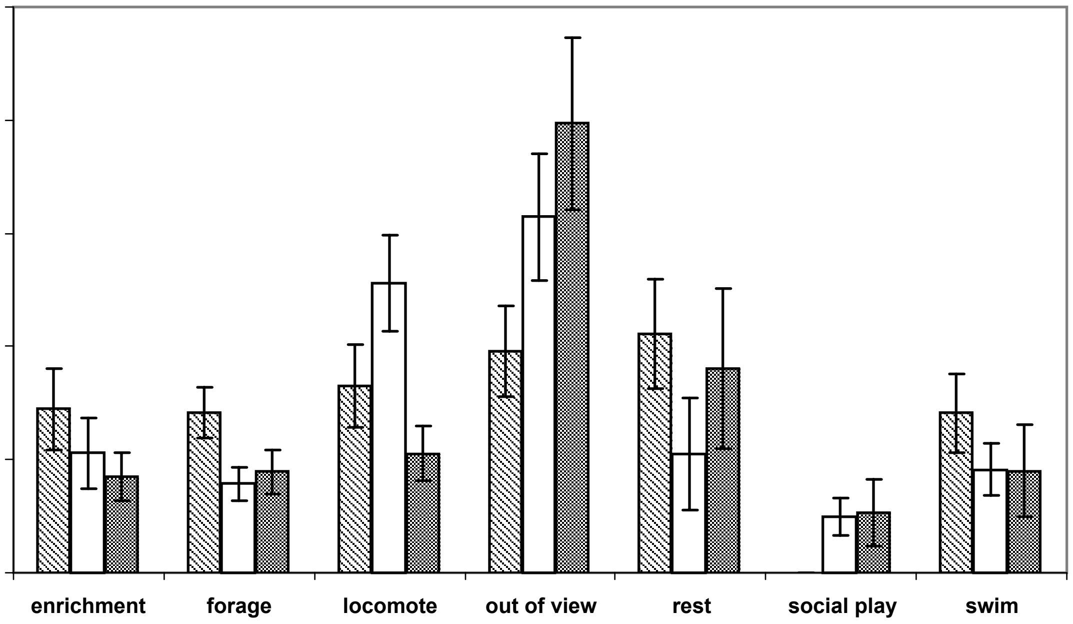
Figure 7. Minutes per hour spent nursing by a dolphin calf during the first 30 days of life. Each day is divided into 4 six-hour shifts (0000-0600, 0600-1200, 1200-1800, 1800-0000). Observations were conducted continuously. The graph represents differences in patterns obtained when utilizing 1, 2, 3, or 6 hours of data per 6 hour shift. There were no significant differences in nursing time across any of the sampling regimes.

Figure 8. Percent time spent in slipstream position by a dolphin calf during the first 30 days of life. . Each day is divided into 4 six-hour shifts (0000-0600, 0600-1200, 1200-1800, 1800-0000). Scan sampling was used to score calf as in or not in slipstream position every minute. The graph represents differences in pattern when utilizing 1, 2, 3, or 6 hours of data per shift. There were no significant differences in slipstream across any of the sampling regimes.

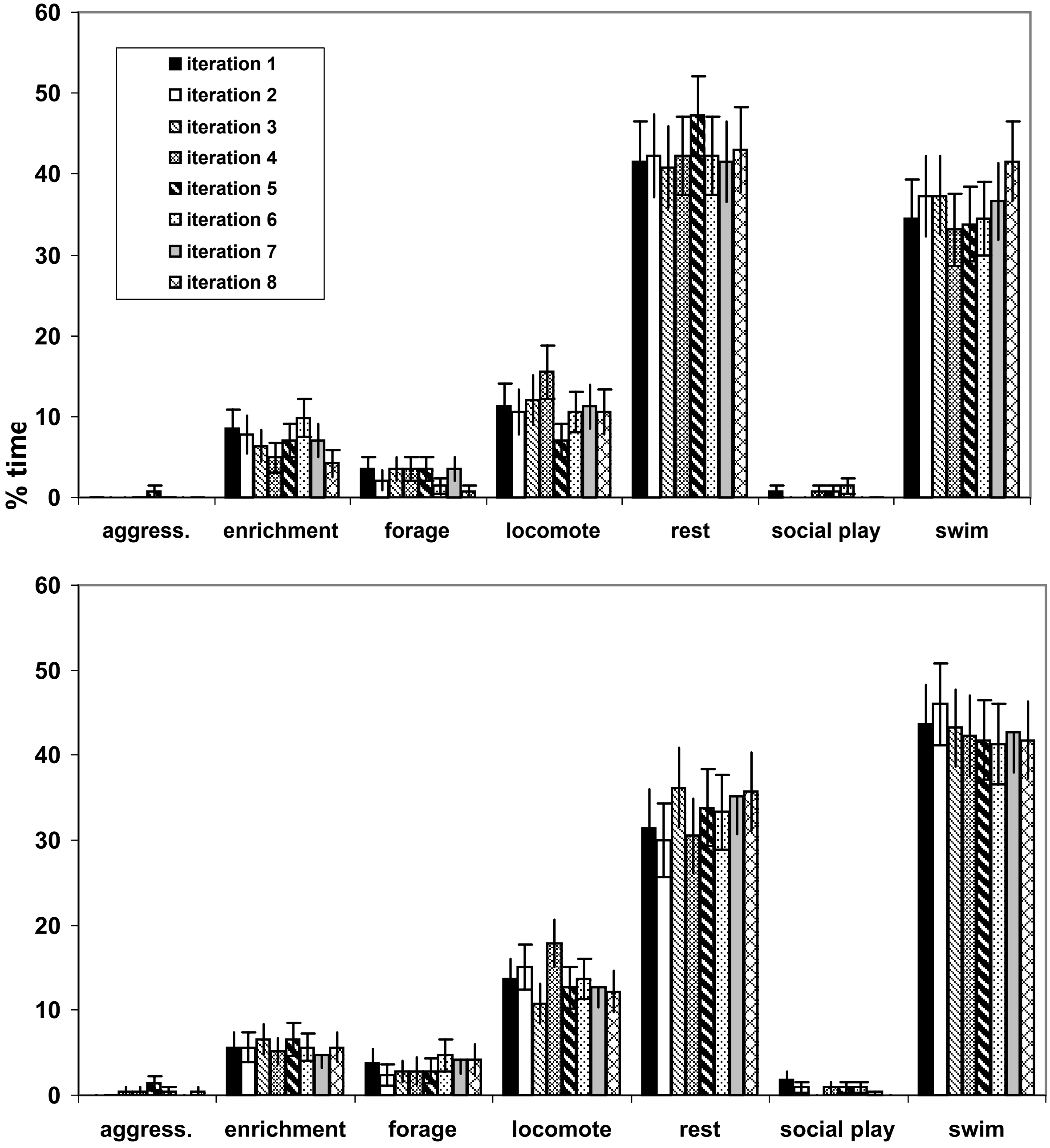
Figures 1-8

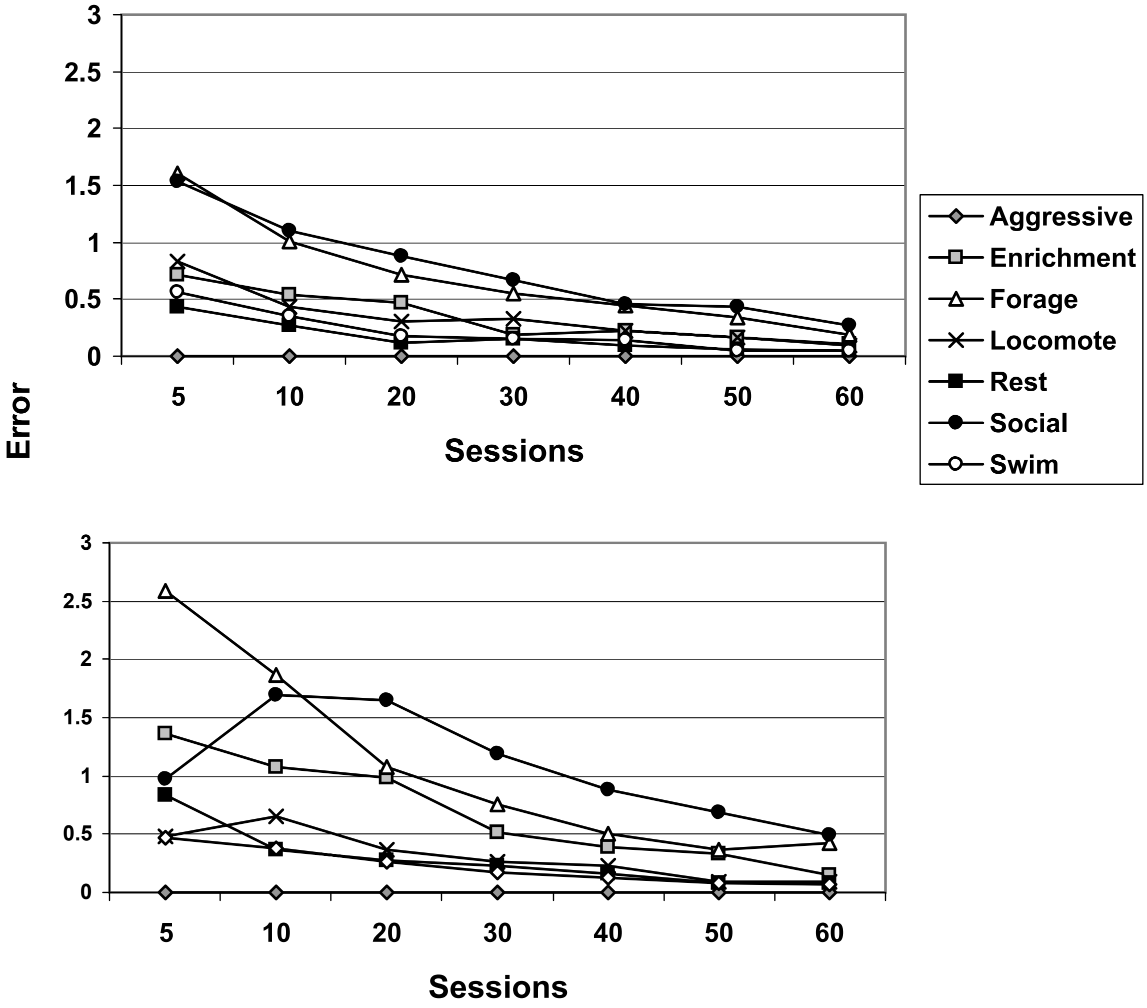


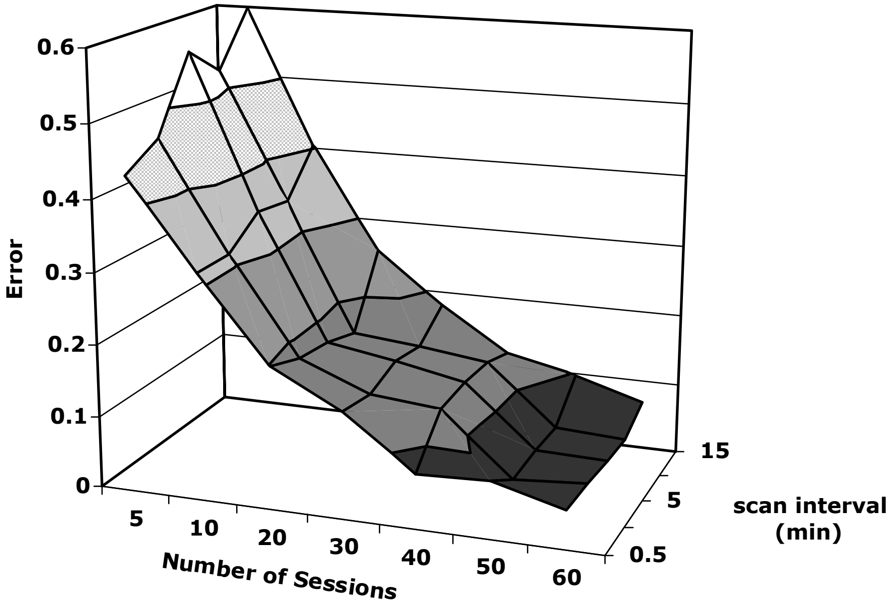
50
40
30
20
10
0
% time
▧ keeper
☐ observer
▩ taper
enrichment
forage
locomote
out of view
rest
social play
swim

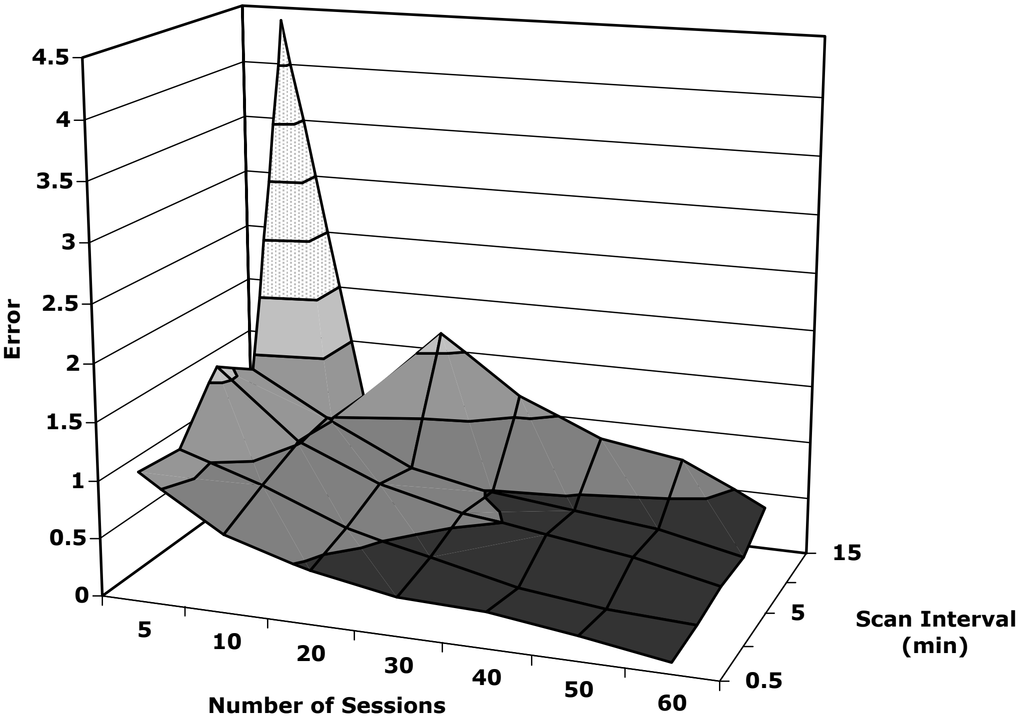


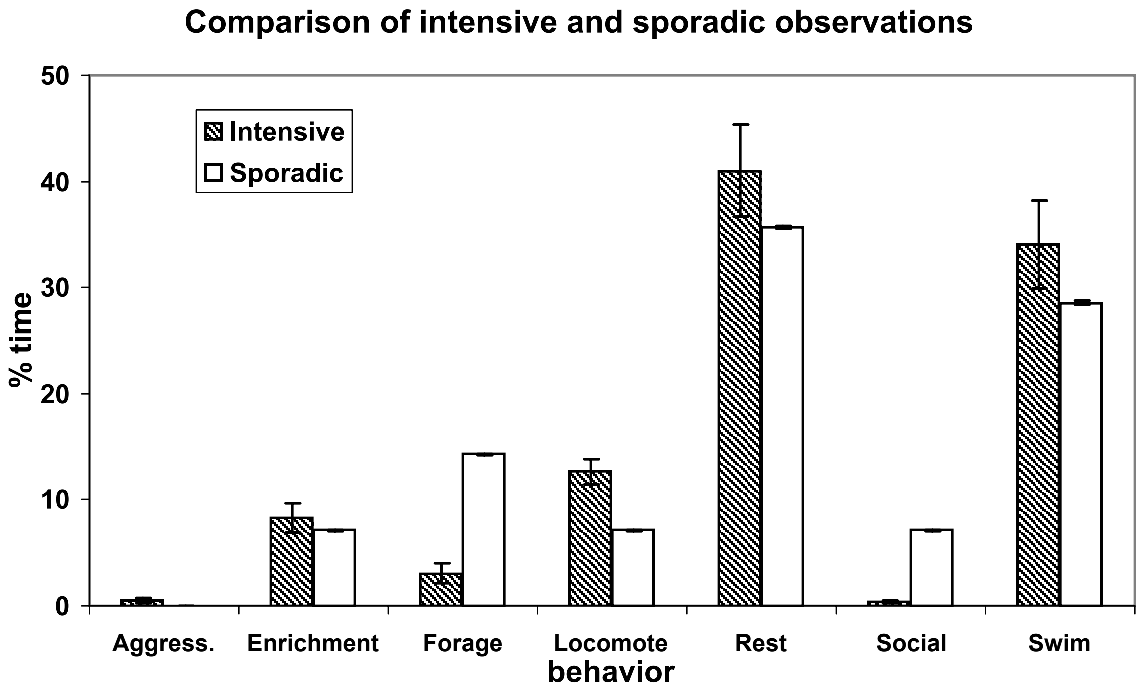
50
40
30
20
10
0
enrichment
forage
locomote
out of view
rest
social play
swim

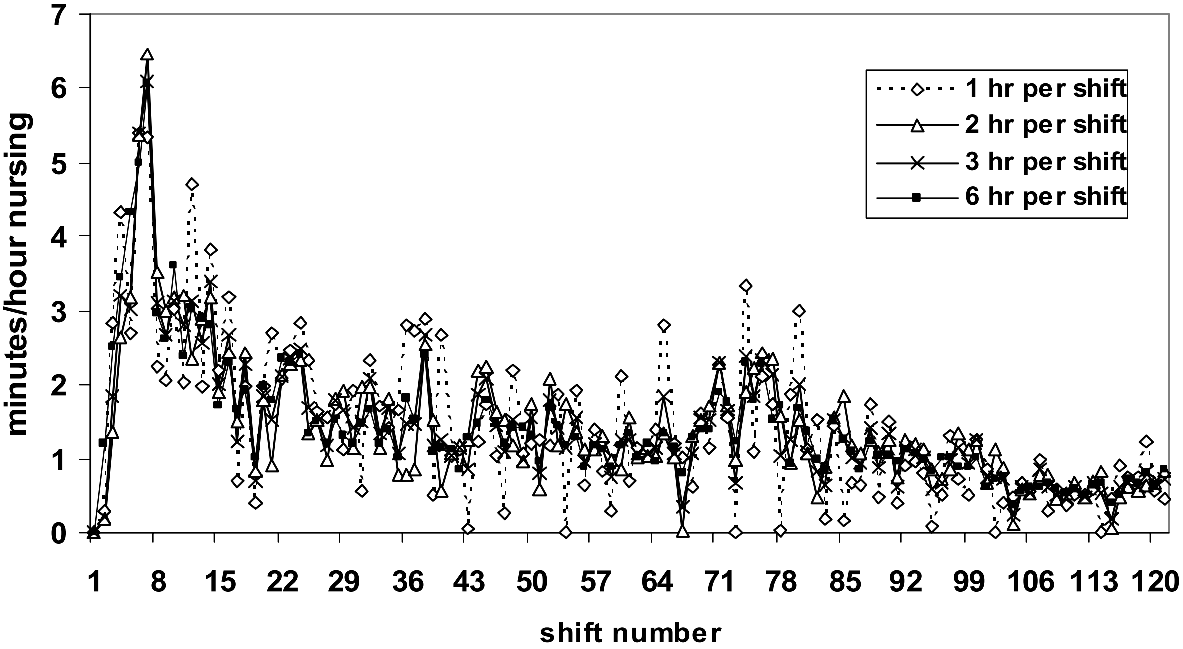

60
50
40
30
20
10
0
% time
iteration 1
iteration 2
iteration 3
iteration 4
iteration 5
iteration 6
iteration 7
iteration 8
aggress.
enrichment
forage
locomote
rest
social play
swim
60
50
40
30
20
10
0
aggress.
enrichment
forage
locomote
rest
social play
swim


3
2.5
2
1.5
1
0.5
0
5
10
20
30
40
50
60
Sessions
Error
Aggressive
Enrichment
Forage
Locomote
Rest
Social
Swim
3
2.5
2
1.5
1
0.5
0
5
10
20
30
40
50
60
Sessions

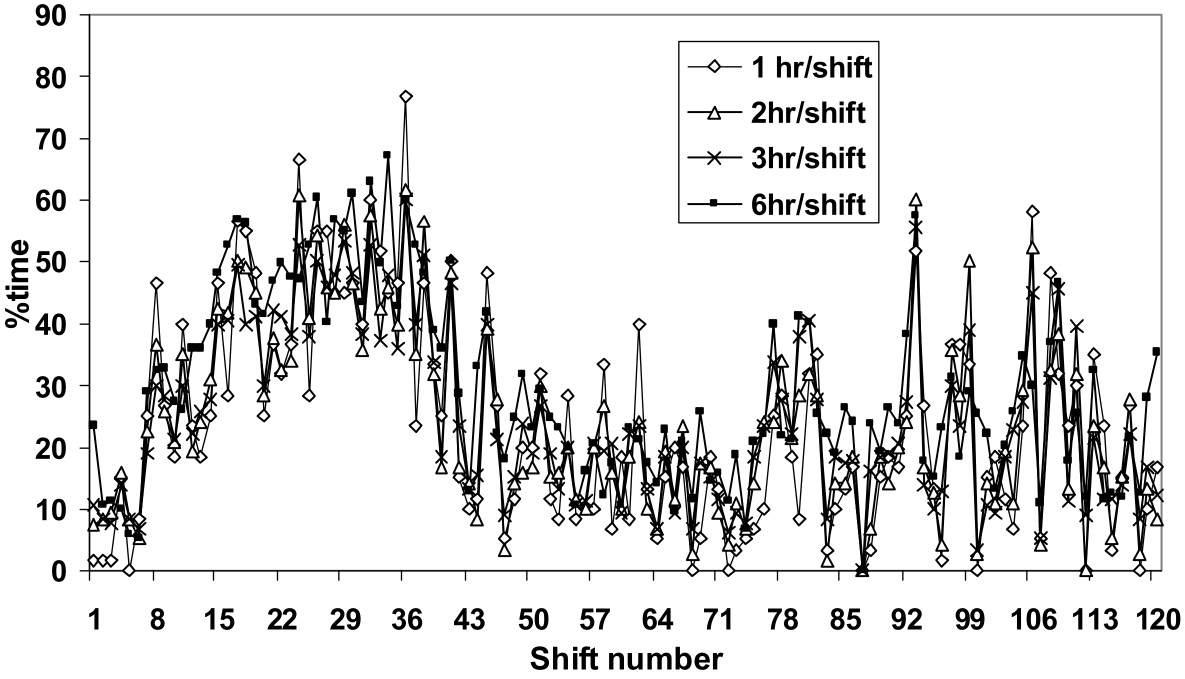

0.6
0.5
0.4
0.3
0.2
0.1
0
Error
5
10
20
30
40
50
60
Number of Sessions
15
5
0.5
scan interval
(min)


Error
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
5
10
20
30
40
50
60
Number of Sessions
15
5
0.5
Scan Interval
(min)


Comparison of intensive and sporadic observations
▨ Intensive
☐ Sporadic
% time
50
40
30
20
10
0
Aggress.
Enrichment
Forage
Locomote
Rest
Social
Swim
behavior



minutes/hour nursing
0
1
2
3
4
5
6
7
1
8
15
22
29
36
43
50
57
64
71
78
85
92
99
106
113
120
shift number
1 hr per shift
2 hr per shift
3 hr per shift
6 hr per shift



%time
90
80
70
60
50
40
30
20
10
0
1 hr/shift
2hr/shift
3hr/shift
6hr/shift
1
8
15
22
29
36
43
50
57
64
71
78
85
92
99
106
113
120
Shift number