

# Assessing the accuracy of interviewed recall for rare, highly seasonal events: the case of wildlife consumption in Madagascar

C. D. Golden<sup>1</sup>, R. W. Wrangham<sup>2</sup> & J. S. Brashares<sup>3</sup>

<sup>1</sup> Center for the Environment, Harvard University, Cambridge, MA, USA

<sup>2</sup> Department of Human Evolutionary Biology, Harvard University, Cambridge, MA, USA

<sup>3</sup> Department of Environmental Science, Policy, and Management, University of California, Berkeley, Berkeley, CA, USA

## Keywords

oral recall; food frequency; illegal activities; wildlife; hunting; accuracy assessment; bias; bushmeat.

## Correspondence

Christopher Golden, Center for the Environment, Harvard University, 24 Oxford Street, 3rd Floor, Cambridge, MA 02138, USA

Email: cgolden@fas.harvard.edu

Editor: Nathalie Pettorelli

Associate Editor: Ioan Fazey

Received 10 October 2012; accepted 28 March 2013

doi:10.1111/acv.12047

## Abstract

Researchers and practitioners from a range of fields including conservation biology, sociology, public health and economics rely on information gained from interviews to quantify the frequency and scale of activities or events of interest. These 'recall' data often form the basis of wildlife sustainability assessments and, ultimately, policy decisions and management actions, but they are highly vulnerable to bias, particularly when the behavior of interest has strong temporal variation. Here, we investigate bias in recalls of wildlife consumption in rural Madagascar by comparing oral recalls collected monthly and annually from male heads of household with daily diet diaries maintained by female heads of household. Daily diet calendars collected from 28 households were assumed to be the measure of true consumption and were used to validate the recalled information. While we found little interhousehold variation in accuracy of responses, we found a tendency for recalls to overreport rates of wildlife consumption. Estimating the annual frequency of rare and/or seasonal events was quantified more accurately by recalls of the prior year than by extrapolation of recalls of the prior month. We conclude that monthly variation in consumption rate leads to predictable errors in estimation of the annual consumption rate. Local consumption of wildlife has large temporal variability, reflecting human preference or the underlying life cycles of animals being consumed. Accurate assessment of consumption rates therefore requires determining an appropriate recall period by taking into account the temporal variability and frequency of the events in question.

## Introduction

The use of data derived from a subject's recall of specific events from memory is a central component of many studies in conservation science, sociology, economics and public health (Coughlin, 1990; Eisenhower, Mathiowetz & Morganstein, 1991; Gavin & Anderson, 2005; Brigham *et al.*, 2008; Ellsberg *et al.*, 2008; Jones *et al.*, 2008). Activities, behaviors and events that are difficult for researchers to directly observe often require the use of recall methodologies to estimate frequencies, durations and/or periodicity of occurrence. Policymakers depend on the reliability of these data to inform their decisions regarding household economics, human health and environmental sustainability.

In many cases, researchers estimate event frequencies by extrapolating from recall information based on a small sample of prior days or weeks (Eisenhower *et al.*, 1991; Lemmens & Knibbe, 1993). Such an approach assumes (1) that a subject can accurately recall the events of interest, (2) that the events quantified during a brief recall period remain

roughly constant in frequency throughout the year and (3) that the timing of the recall period will capture or detect events of research interest (Lemmens & Knibbe, 1993). However, these assumptions are often violated such that estimates based on extrapolation from short recall periods are vulnerable to systematic and response errors (Lemmens & Knibbe, 1993; Johansson *et al.*, 2002). This problem can be further compounded by investigators designing surveys to be conducted during periods where events of particular interest will occur relatively frequently. Assessing the accuracy and precision of these recalled estimates by using mixed-method data collection is needed to determine the validity of researcher's reliance on this form of inference (Rosenberg *et al.*, 1983).

In this study, we use a case of hunting and wildlife consumption in north-eastern Madagascar to assess the accuracy of annual recalls for rare and seasonal events. Prior evidence has shown that hunting in this region is highly seasonal, and for some species, rare (Golden, 2009; Golden *et al.*, 2011). We compared prior month with prior year

recalls of wildlife use against daily diet calendars, assuming the calendars to be measures of true consumption because they were not subject to any recall delay. We posed the following hypotheses: (1) recalls will predict the frequency of consumption of rarely consumed taxa more accurately than frequently consumed taxa (Sudman & Schwarz, 1989; Eisenhower *et al.*, 1991; Reis & Judd, 2000); (2) recalls will estimate the consumption of aseasonally consumed taxa more accurately than seasonally consumed taxa (Lemmens & Knibbe, 1993); (3) prior month recalls conducted in a low-hunting season will provide more accurate consumption information than recalls conducted in a high-hunting season; (4) the average of the two prior month recalls (one from high-hunting and one from low-hunting season) will provide more accurate data than individual prior month recalls; and (5) annual recalls will more accurately predict the true long-term rates of consumption than prior month recalls.

Our hypothesis that long-term annual recalls would perform better than prior month recalls might seem surprising given that, in general, more recent events are more easily recollected (Reis & Judd, 2000). However, linear extrapolation requires a constant rate of occurrence through the period of extrapolation (Wentland & Smith, 1993). In the case of many seasonal events, this assumption of rate constancy is wrong. Further, because researchers can often anticipate which seasons are most productive, they sometimes design the timing of their research effort so as to increase the probability of detecting rare events (Wentland & Smith, 1993). This procedure risks producing an inflated estimate of the frequency of such events. These types of systematic bias are critical to understand in the field of sustainability science so that we do not gravely underestimate or overestimate harvest pressure on wildlife.

## Methods

### Sampling and data collection

From February 2008 through February 2009, we enlisted the participation of 28 households in one community in the Makira Natural Park of north-eastern Madagascar. These households were selected by systematic random sampling (see Golden *et al.*, 2011), and all had a male and female head of household. The female head of each household completed daily diet calendars listing the number of individuals of each type of wildlife consumed within their household. Twice during the research period (May 2008 and January 2009), a diet recall was conducted in which the male head of each household was asked to estimate the number of animals consumed of 15 wildlife taxa (14 species, plus a single category for all insectivorous bats) in the household over the past month. These were the most frequently consumed taxa in the area. The recall conducted in May occurred during the high-hunting season while the recall in January occurred during the low-hunting season. At the end of the study, we

conducted an annual diet recall by asking each male head of household to enumerate the numbers of animals of each taxa consumed during the past year (see Golden, 2009 for details).

The daily diet calendars were maintained as part of a larger public health study, and the recall data were collected during unplanned visits to the households. A research assistant visited each household weekly throughout the year to ensure that the calendars were properly maintained. We assumed that the two most important potential sources of information recall bias would be memory failure and intentional misrepresentation (Bradburn, Rips & Shevell, 1987). Because our team had been working in this community for 5 years prior to this study and had gained the trust of community members, we believe that deliberate misleading was absent or trivial. Accordingly, we assume that recall bias came primarily from misremembering.

### Cultural context

Rice agriculture was the primary form of labor activity for both men and women. Certain tasks were performed jointly such as weeding, and others were highly segregated by sex. Hunting and cooking largely fell into the second category of labor where activities were segregated by sex. It was the man's domain to hunt and the woman's domain to prepare and cook food. This was the reason why the diet calendars were maintained by the female head of household, and the oral recall was performed by the male head of household. There was very little variation in the role of men and women in producing and securing food. Thus, households were reasonable replicates of one another in this sense.

Accurately assessing rates of consumption is important in determining the sustainability of harvest and future trajectory of wildlife populations. In this case, wildlife consumption can be viewed as nearly all of harvest. In this region, there is no market for luxury species at markets or restaurants, although it is very occasionally served in urban households. Further, although certain animals are hunted as pests, they are also consumed after death. Thus, the consumption records in this study should offer a clear indication of harvest pressure.

Although Makira Natural Park was gazetted as a protected area in 2005, the laws restricting hunting remain relatively unmonitored and unenforced (Golden, 2009). Our study community, while surrounded by forest, lies 5.5 km from the boundary of the protected area. The community enrolled into this study is thus not part of the buffer area of the park, where active community engagement is taking place under the aegis of the Wildlife Conservation Society and the Ministry of Water and Forests.

### Data analysis

In all comparisons of recall types, we considered daily diet calendar data to represent 'true' wildlife consumption events. We used two metrics to assess the accuracy of oral recall methods as compared with values obtained from the

**Table 1** MSE estimates associated with each recall method for each taxa

Taxa	Annual recall MSE	Prior month recall (high hunting) MSE	Prior month recall (low hunting) MSE	Combined prior month recalls MSE
Seasonally consumed				
<i>Avahi laniger</i> <sup>a</sup>	0.18	27.88	0.15	7.36
<i>Cheirogaleus</i> sp. <sup>a</sup>	0.89	768.15	7.00	198.72
<i>Daubentonia madagascariensis</i> <sup>a</sup>	0	0.08	0.11	0.08
<i>Eulemur albifrons</i> <sup>a,b</sup>	2.86	434.46	28.44	92.92
<i>Haplemur griseus</i> <sup>a</sup>	0.07	121.04	10.81	45.72
<i>Indri indri</i> <sup>a</sup>	0	0.08	0.12	0.08
<i>Microcebus</i> sp. <sup>a</sup>	0.04	22.15	5.33	7.20
<i>Rousettus madagascariensis</i> <sup>b</sup>	5.82	2218.50	109.67	608.04
<i>Setifer setosus</i> <sup>b</sup>	2.71	211.19	193.44	147.96
<i>Tenrec ecaudatus</i> <sup>b</sup>	5.36	6104.69	71.81	1592.52
Aseasonally consumed				
<i>Cryptoprocta ferox</i> <sup>a</sup>	0	11.15	0.07	2.96
<i>Galidia elegans</i>	0	3.12	0.33	0.36
<i>Microchiroptera</i> spp.	0.04	5.54	0	1.44
<i>Potamochoerus larvatus</i> <sup>b</sup>	0.11	5.15	0.56	1.52
<i>Viverricula indica</i>	0.04	33.88	0.67	9.32
Unweighted mean MSE	1.21	664.47	28.57	181.08

Mean squared errors (MSEs) for each recall method where the records of the daily diet calendars serve as the measure of true consumption. Low numbers indicate greater accuracy. For 13 of 15 taxa, the annual recall method serves as the most accurate reflection of true consumption.

<sup>a</sup>Denotes taxa that are illegal to harvest throughout the year.

<sup>b</sup>Denotes taxa that are frequently consumed (i.e. commonly consumed five or more times per year by a given household).

daily diet calendars: (1) ratios of the geometric means of recall and calendar data across taxa per household (Johansson *et al.*, 2002) and (2) mean squared errors (MSEs) to quantify the difference between the diet calendar report and oral recall. The first method follows Johansson *et al.* (2002) and transforms the Poisson distributed count data of observed diet calendar consumption events and estimated oral recall of consumption events using the equation  $[\ln(x+1)]$ . We then calculated the geometric mean of transformed consumption data across each household for each taxa for both the recall and calendar data. Finally, we calculated a ratio of the geometric means of the recall data and the calendar data. The closer this ratio is to 1, the more accurately the recall data from a given household matched the assumed true rate of annual consumption.

For the second analysis, we used MSE to calculate the difference between the calendar count (i.e. the true count) and the recall count (i.e. the estimator of the truth). The MSE is a quantitative comparison of two values that determines the degree of similarity between the values, usually assuming that one of the values is accurate and the other subject to error (Wang & Bovik, 2009). We used this approach to compare the average recall count with each calendar count, the average calendar count with each recall count, and both the average calendar and recall count for each taxa and household.

We repeated the same type of analysis for three other cases: the extrapolated prior month recalls from May 2008, those performed at the end of January 2009 and a combined recall extrapolation where the counts from May and January were added together and then extrapolated to a

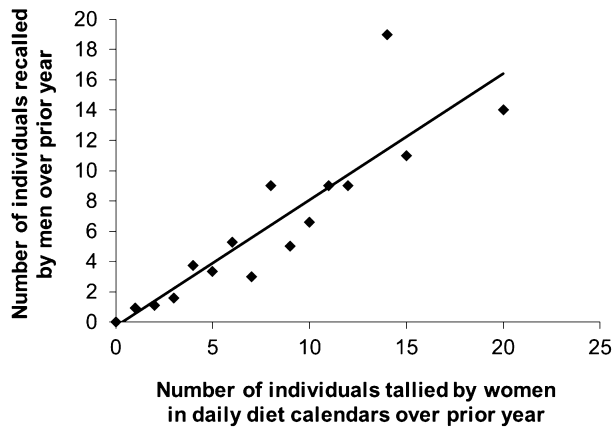
year. We then compared these extrapolated short-term recall rates with the annual recall rate to determine which recall method had better predictive accuracy of annual rates of consumption – as determined by their respective MSE value. Households in this study had conducted recalls 0–3 times prior to this research.

We used a generalized linear mixed model (GLMM) with an unstructured covariance matrix to determine the association between the MSE value (from the annual recall count and the reported dietary calendar count) and a series of covariates using the household as a random effect:

$$Y_{ijk} = \beta_0 + \beta_{0i} + \beta_1 x_{ij} + \beta_2 x_{ij} + \beta_3 x_{ij} + \beta_4 x_{ij} + \beta_5 x_{ij} + e_{ijk}$$

$Y_{ijk}$  is the MSE value in the  $i^{\text{th}}$  household for the  $j^{\text{th}}$  sample, where a lower MSE denotes greater accuracy of recall;  $\beta_0$  is a constant;  $\beta_{0i}$  is the random effect for the household;  $\beta_1$  is the reported frequency of consumption from the diet calendars;  $\beta_2$  is a binary categorical variable expressing whether hunting practices for a given taxa are seasonal or aseasonal;  $\beta_3$  is a binary categorical variable ('elementary' or 'post-elementary') characterizing the educational attainment of the male head of household who conducted the recall;  $\beta_4$  is a continuous variable listing the number of years in which the head of household has conducted an annual recall;  $\beta_5$  is the body mass of each given species (log-transformed because of positive skew); and  $e_{ijk}$  is the error term. Illegal species included only those species that are illegal to harvest throughout the year. These included an endangered carnivore species, the fosa (*Cryptoprocta ferox*) and all lemur species (Table 1 to see species list). Aseasonally hunted

species included all carnivorous species, insectivorous bat species and the bush pig (*Potamochoerus larvatus*); seasonally hunted species included all lemurs, frugivorous bat species and members of the family Tenrecidae (C. D. G., unpublished data). Seasonal refers to species that are consumed during a particular hunting season, which typically last 3–8 months per year depending on the species. This seasonality can arise from two reasons: estivation by a species or hunter behavior and preferences. We use the term



**Figure 1** Accuracy of long-term recalls in predicting annual consumption. Recall estimates from the male heads of household of the average number of individuals of a given taxa consumed in the prior year were highly correlated with the number of individuals tallied in daily diet calendars maintained by female heads of household ( $r = 0.85$ ,  $n = 102$ ,  $P < 0.0005$ ).

‘rarely consumed’ to mean consumed less than five times in 1 year by one household.

## Results

### Variation in accuracy based on recall length and timing

There were 420 total paired samples (15 taxa in 28 households) of the taxa under investigation. In comparing the annual recalls to the daily diet calendars, 318 of the 420 paired samples were 0,0 pairs where both the male’s recall of consumption and the female’s diet calendar from the same household recorded no consumption of that taxa. Excluding the 0,0 pairs, there was a high correlation between the number of individual animals reported to be consumed by men and tallied in women’s daily diet calendars ( $r = 0.85$ ,  $n = 102$ ,  $P < 0.0005$ , Fig. 1).

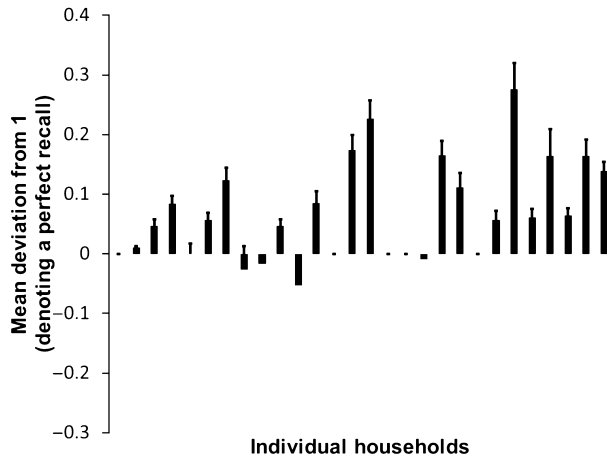
We found that the annual recall performed best overall when examining the MSEs associated with the consumption reported in the daily diet calendar as compared with the annual, prior month (from both high and low hunting) and combined month recalls (15 of 16 cases, Table 2; one-sided binomial test  $P = 0.0003$ ). Following our initial hypotheses, the prior month recall from the low-hunting season performed better than the prior month recall from the high-hunting season in predicting true rates of annual consumption (14 of 16 cases, Table 2; one-sided binomial test  $P = 0.002$ ). Moreover, the prior month recall from the low-hunting season performed better than the combined prior month recalls from both the high hunting and low-hunting season (12 of 16 cases, Table 2; one-sided binomial test

**Table 2** MSE estimates associated with each recall method for a given level of true consumption (measured by the daily diet calendar)

Recorded consumption on calendar	Annual recall MSE	Prior month recall (high hunting) MSE	Prior month recall (low hunting) MSE	Combined prior month recalls MSE
0	0.01	254.15	4.19	69.05
1	0.94	52.30	42.23	32.70
2	1.70	1053.16	92.10	311.68
3	3.20	1259.22	304.13	565.71
4	0.75	764.75	20.75	158.75
5	6.78	6855.11	19.29	2323.29
6	6.14	3283.86	115.86	887.29
7	16.00	9.00	2025.00	441.00
8	1.00	2601.00	81.00	441.00
9	16.00	49.00	–	–
10	24.20	1960.33	37.00	556.33
11	4.00	–	81.00	–
12	9.00	225.00	225.00	225.00
14	25.00	841.00	361.00	25.00
15	16.00	2401.00	121.00	361.00
20	45.00	49 729.00	205.00	10 609.00
Unweighted mean MSE	10.98	4755.86	248.97	1214.77

The first column is a list of how many animals of each given taxa were consumed in a particular household over a 1-year period according to the daily diet calendars. The daily diet calendars are assumed to be a true measure of consumption. Each subsequent column lists the mean squared error by each recall method of all the recall estimates associated with a given level of true consumption. Rare consumption is defined as being consumed less than five times per year.

MSE, mean squared error.



**Figure 2** Calculating agreement between recall-based estimates of household wildlife consumption and data from daily diet calendars. The 0 line represents a perfect fit between recall and diet calendars. Each household's ( $n=28$ ) accuracy of recall is depicted, averaged across the 15 taxa of interest. Negative deviation represents underreporting whereas positive deviation represents overreporting. In general, annual recalls were highly accurate despite a trend toward overreporting of wildlife consumption events for 15 wildlife taxa.

$P = 0.04$ ). For example, based only on the extrapolated rates of hunting from prior month recall during the high-hunting season, we calculated that the dwarf lemur (*Cheirogaleus* sp.) was consumed on average 1.7 times its true consumption rate and potentially as high as 8.6 times the true rate. More strikingly, using the same extrapolation from prior month recall, we calculated that the common tenrec (*Tenrec ecaudatus*) was consumed on average 5.9 times its true consumption rate and as high as 48 times the true rate.

### Causes of variation in recall discrepancies

The mean ratio of the log-transformed geometric means of the recall count data and the calendar count data for wildlife consumption was 1.056 (median = 1.044; standard error = 0.013), demonstrating high predictive accuracy of annual recalls on the daily diet calendars (Fig. 2). There was little variation between households in terms of accuracy of estimating annual consumption through long-term recalls (Fig. 2). Twenty-four of the 28 households slightly overreported their annual consumption rates. Dividing the taxa into those that were legally versus illegally harvested, the mean deviation from a perfect recall for illegally harvest taxa was 0.03 whereas the mean deviation for legally harvested taxa was 0.11.

In examining causes of variation across the MSEs, certain factors arose as possible sources of variation. The number of times a taxa was consumed according to the diet calendar was the strongest predictor for MSE, where more frequent consumption increased the MSE ( $\beta = 1.08$ ,  $P = 0.006$ , Table 3). Taxa that were consumed less frequently were more reliably reported by recalls than taxa that were con-

**Table 3** The association between household and species' characteristics and the accuracy of oral recall reports ( $n = 420$ )

Variable	Coefficient <sup>a</sup>
Calendar count <sup>b</sup>	1.08***
Seasonality <sup>c</sup>	0.35**
Male education <sup>d</sup>	-0.26
Recall history <sup>e</sup>	-0.77
Species body mass <sup>f</sup>	-0.14*

High species body mass increased the accuracy of recall of wildlife consumption whereas high frequency of consumption and seasonality reduced the accuracy of recall.

<sup>a</sup>Significance level for each coefficient \* $P < 0.10$ ; \*\* $P < 0.05$ ; \*\*\* $P < 0.005$ .

<sup>b</sup>Calendar count refers to the annual rates of consumption of each taxa according to a household's daily diet calendar.

<sup>c</sup>Seasonality is a binary categorical variable expressing whether hunting practices for a given species are aseasonal (0) or seasonal (1).

<sup>d</sup>Male education is the maximal educational attainment of the male head of household who conducted the recall (categorized as either elementary or post-elementary).

<sup>e</sup>Recall history is a continuous variable listing the number of years in which the head of household has conducted an annual recall.

<sup>f</sup>Species body mass is the weight of each given species (log-transformed because of positive skew).

sumed less frequently (100 vs. 80% of cases, Table 1). If a given taxa was only seasonally harvested (Table 1), the accuracy of recall also decreased (the absolute value of MSE increased) significantly ( $\beta = 0.35$ ,  $P = 0.025$ , Table 3). If a taxa had a higher body mass, the accuracy of recall increased ( $\beta = -0.14$ ,  $P = 0.097$ , Table 3). There were no significant statistical associations between the accuracy of recall estimators and either the male head of household's education or the number of times which he has conducted a recall for past research (Table 3). This means that rarity, more than large body size, was more memorable to an individual recalling consumption events (Table 3). Additionally, frequency of consumption was more likely to skew the accuracy of recall than the seasonality of consumption (Table 3).

### Discussion

It is critical to validate the accuracy of survey methods that provide numerical responses on which management and policy actions are based. Our results show that annual recalls performed better than prior month recalls from either low- or high-hunting seasons and better than prior month recalls that combined a month from both seasons. The 1-month recall from the low-hunting season performed better than both the prior month recall during the high-hunting season and the combined 1-month recalls from both the high hunting and low-hunting season in predicting annual rates of consumption. This demonstrates the risk of overestimating true rates of consumption by conducting recalls during the high-hunting season. In this case, and in other similar cases where events of interest are rare and/or seasonal, a longer recall period could be the best approach,

or extrapolation should be limited to the season of occurrence. The annual recall is more cost-effective than two separate recall periods and produces a more accurate estimate in this case. It should be noted that the two prior month recalls may have affected the accuracy of the annual recall, and it should also be noted that many participants had performed recalls for this research in the past. However, other studies have shown that rapid assessments can only provide qualitative but not quantitative accuracy and should only be used for restricted applications (Gavin & Anderson, 2005).

Contrary to the suggestions of other studies (Casswell, Huckle & Pledger, 2002; Johansson *et al.*, 2002), 24 of 28 households overreported their consumption behaviors. With more than half of the taxa being illegal to harvest throughout the year, one might expect that consumption behaviors would be underreported rather than overreported (e.g. Rist *et al.*, 2010). Our result is similar to the finding that fishermen tended to overreport their catch (Lunn & Dearden, 2006), but dissimilar to Jones *et al.* (2008) who found that respondents overreported low levels of consumption and underreported high levels of consumption. However, in our study, the ratio of the recall to the calendar shows only a small margin of overreporting. This is likely due to one of two reasons: (1) the general lack of fear of reporting hunting activities in this area as local monitoring and enforcement are minimal or (2) the possibility that men are overreporting wildlife consumption as a point of pride (e.g. the influence of social desirability, DeMaio, 1984). It is possible that both men and women were simultaneously underreporting use of illegal taxa. Our data indicated that there was a high correlation between responses, and that legality and illegality of taxa did not affect the recall accuracy.

Our results underline the danger of extrapolating from short-term recall while assuming that behaviors of interest occur at a constant rate over time. Yet our research does not suggest that an annual recall is always better than a short-term recall period. It does suggest that we must be diligent in extrapolating rare and/or seasonal events. Rare events, or what Reis & Judd (2000) would call salient or distinctive events, are more accurately recalled than very common events. Understanding the temporal variability of behaviors of interest is critical to designing research methods that produce accurate estimates and are least likely to suffer from systematic bias. Thus, we must be sure that the recall is long enough to detect an event of interest but short enough for recall to remain vivid. Further, researchers should not feel obliged to utilize only one recall period in a given study but should adjust this recall period to be meaningful to the specific study subject. We recommend: (1) conducting a pilot study with focus groups to understand the rarity and seasonality of the events in question prior to creating surveys; (2) designing the study so that subjects are not exclusively observed during highs and/or lows; (3) creating recall periods that will allow the researcher to detect the event without risking memory inaccuracies; and (4) only extrapolating responses to the season of occurrence, if this is

relevant. These recommendations have significant implications for affecting study design to minimize bias and reduce systematic errors that could inadvertently mislead managers and policymakers.

## Acknowledgments

We would like to thank J. Gagnon-Bartsch for statistical consultation. We also acknowledge financial support by the National Geographic Society Conservation Trust #C135-08, the Margot Marsh Biodiversity Fund # 023815, the Mohamed bin Zayed Species Conservation Fund #1025935 and the National Science Foundation doctoral dissertation improvement grant # 1011714 and National Science Foundation Coupled Natural Human Systems grant #NSF-GEO1115057. We would also like to thank L. Fernald for her financial generosity in purchasing the kitchen scales for this project.

## References

- Bradburn, N.M., Rips, L.J. & Shevell, S.K. (1987). Answering autobiographical questions: the impact of memory and inference on surveys. *Science* **236**, 157–161.
- Brigham, J., Lessov-Schlaggar, C.N., Javitz, H.S., McElroy, M., Krasnow, R. & Swan, G.E. (2008). Reliability of adult retrospective recall of lifetime tobacco use. *Nicotine Tob. Res.* **10**, 287–299.
- Casswell, S., Huckle, T. & Pledger, M. (2002). Survey data need not underestimate alcohol consumption. *Alcohol Clin. Exp. Res.* **26**, 1561–1567.
- Coughlin, S.S. (1990). Recall bias in epidemiologic studies. *J. Clin. Epidemiol.* **43**, 87–91.
- DeMaio, T.J. (1984). Social desirability and survey measurement: a review. In *Surveying Subjective Phenomena*: 257–281. Turner, C. (Ed.). New York: Russell Sage Foundation.
- Eisenhower, D., Mathiowetz, N. & Morganstein, D. (1991). Recall error: sources and bias reduction techniques. In *Measurement errors in surveys*: 127–144. Biemer, P., Groves, R., Lyberg, L., Mathiowetz, N. & Sudman, S. (Eds). Hoboken, NJ: Wiley-Interscience.
- Ellsberg, M., Jansen, H.A., Heise, L., Watts, C.H. & Garcia-Moreno, C. (2008). Intimate partner violence and women's physical and mental health in the WHO multi-country study on women's health and domestic violence: an observational study. *Lancet* **371**, 1165–1172.
- Gavin, M. & Anderson, G. (2005). Testing a rapid quantitative ethnobiological technique: first steps towards developing a critical conservation tool. *Econ. Bot.* **59**, 112–121.
- Golden, C.D. (2009). Bushmeat hunting and use in the Makira Forest north-eastern Madagascar: a conservation and livelihoods issue. *Oryx* **43**, 386–392.
- Golden, C.D., Fernald, L.C.H., Brashares, J.S., Rasolofoniaina, B.J.R. & Kremen, C. (2011). Benefits of wildlife consumption to child nutrition in a biodiversity hotspot. *Proc. Natl. Acad. Sci. USA* **108**, 19653–19656.

- Johansson, I., Hallmans, G., Wikman, A., Biessy, C., Riboli, E. & Kaaks, R. (2002). Validation and calibration of food-frequency questionnaire measurements in the Northern Sweden Health and Disease cohort. *Public Health Nutr.* **5**, 487–496.
- Jones, J.P.G., Andriamarivololona, M., Hockley, N., Gibbons, J. & Milner-Gulland, E.J. (2008). Testing the use of interviews as a tool for monitoring trends in the harvesting of wild species. *J. Appl. Ecol.* **45**, 1205–1212.
- Lemmens, P.H. & Knibbe, R.A. (1993). Seasonal variation in survey and sales estimates of alcohol consumption. *J. Stud. Alcohol Drugs* **54**, 157–163.
- Lunn, K. & Dearden, P. (2006). Monitoring small-scale marine fisheries: an example from Thailand's Ko Chang archipelago. *Fish. Res.* **77**, 60–71.
- Reis, H. & Judd, C. (2000). Event-sampling and other methods for studying everyday experience. In *Handbook of research methods in social and personality psychology*: 190–222. Reis, H. & Gable, S. (Eds). Cambridge: Cambridge University Press.
- Rist, J., Milner-Gulland, E., Cowlshaw, G. & Rowcliffe, J.M. (2010). Hunter reporting of Catch per unit effort as a monitoring tool in a bushmeat-harvesting system. *Conserv. Biol.* **24**, 489–499.
- Rosenberg, M.J., Layde, P.M., Ory, H.W., Strauss, L.T., Rooks, J.B. & Rubin, G.L. (1983). Agreement between women's histories of oral contraceptive use and physician records. *Int. J. Epidemiol.* **12**, 84–87.
- Sudman, S. & Schwarz, N. (1989). Contributions of cognitive psychology to advertising research. *J. Advert. Res.* **29**, 43–53.
- Wang, Z. & Bovik, A. (2009). Mean squared error: love it or leave it? *IEEE Signal Process. Mag.* 98–117.
- Wentland, E. & Smith, K. (1993). *Survey responses: an evaluation of their validity*. New York: Academic Press.