

Reports

Can Anthropologists Distinguish Good and Poor Hunters? Implications for Hunting Hypotheses, Sharing Conventions, and Cultural Transmission

Kim Hill and Keith Kintigh

School of Human Evolution and Social Change, Arizona State University, P.O. Box 872402, Tempe, Arizona 85287-2402, U.S.A. (kim.hill@asu.edu). 31 VII 08

CA+ Online Supplements: figures, tables.

Numerous articles examine the relationship between men's hunting skill and other important biological and social traits. We analyzed more than 14,000 hunter days during 27 years of monitoring the Ache of Paraguay by using resampling methods to demonstrate that large sample sizes are generally required in order to distinguish individual men by hunting skill. A small published study on !Kung hunters shows that large-game hunters are even more difficult to distinguish by individual skill level. This is a serious problem because regressions using noisy hunting data as the independent variable systematically underestimate the association of hunting ability with other biosocial traits. The analysis suggests that some coresidents in many small-scale societies will be unable to accurately distinguish hunters by skill level, possibly favoring groupwide meat-sharing conventions and biased cultural transmission that emphasizes prestige rather than perceived hunting skill.

Anthropologists have been interested in hunting for many reasons. First, the nutritional dependence on hunting represents an important shift in ape feeding patterns possibly associated with other notable human behaviors such as food sharing, high levels of cooperation, the sexual division of labor, and important life-history shifts (e.g., Kaplan et al. 2000; Stanford and Bunn 2001). Because our hominin ancestors were probably strongly dependent on meat for the past 2 million years and because faunal remains preserve well, much of the archeological record is dominated by discussions of hunting (and scavenging). Second, hunting still contributes significant nutrients to many small-scale societies in most regions of the world (e.g., Robinson and Bennett 2000). Third, success in hunting is often assumed to be associated with status and other social perks (Wiessner 1996), as well as suc-

cess in mating competition (Hawkes and Bird 2002; Kaplan and Hill 1985), better offspring survivorship (Hill and Hurtado 1996), and higher biological fitness through a variety of mechanisms (Gurven and Von Rueden 2006; Smith 2004).

Because of these interests, data from a number of societies have been employed to investigate associations between hunting success rate of individual men and other theoretically interesting traits. These include relationships between hunting and the probability of marrying a younger or more attractive wife (Gurven and Von Rueden 2006; Marlowe 2000), higher fertility (Hill and Hurtado 1996; Marlowe 1999; Wiessner 2002), number of extramarital sex partners (Kaplan and Hill 1985), offspring survival (Hill and Hurtado 1996; Wiessner 2002), number of stepchildren (Marlowe 1999), number of coresident kin (Wiessner 2002), number of trade partners or household possessions (Bailey 1991; Wiessner 2002), status and coalitional support (Gurven and Von Rueden 2006), amount of time a man's wife spends gathering plant resources (Hurtado et al. 1992), and amount of time spent hunting itself (Bird, Smith, and Bird 2001; Hawkes et al. 1985; Hill and Hawkes 1983). Surprisingly, some studies report *no* relationship between hunting and likely associated variables. These include no relationship between hunting rates and food consumed by the hunter's family (Kaplan and Hill 1985), total amount of food received from others (Hawkes, O'Connell, and Blurton Jones 2001a), participation in child care (Marlowe 1999), offspring seasonal weight change (Hawkes, O'Connell, and Blurton Jones 2001b), probability of marriage, number of wives (Bailey 1991; Wiessner 2002), a man's age, number of offspring (Hawkes, O'Connell, and Blurton Jones 2001a), age at marriage, level of fertility, offspring survival, or other fitness measures (Dwyer and Minnegal 1993; Kent 1996). A lack of association for some of these might be expected, but others seem theoretically improbable but have nevertheless been employed to promote a complete revision of accepted anthropological ideas about men, women, sexual division of labor, and the role of hunting in society (e.g., Hawkes, O'Connell, and Blurton Jones 2001b). The surprising negative findings are often widely cited and incorporated into newly emerging dogma (e.g., Codding and Jones 2007; Hawkes and Bird 2002; O'Connell et al. 2002), with no assessment of the power of the statistical tests to detect differences, if they did exist, between men or acknowledgement of required methodological adjustments when the independent variable (hunting return rate) in a regression contains considerable error.

Here we examine the implications of error in assessing long-term hunting skill, given considerable daily variance in hunting success. First, we consider whether scientists can really distinguish individual hunters by skill, given typical sample sizes, and how correctly to employ noisy hunting data as the independent variable in regression analyses. Second, we consider whether native peers can assess each other's hunting

Table 1. The Ache hunting sample

Year	Hunter days	Days with kill	Successful days (%)	Game (kg)	Rate (kg/day)	Rate (kg/hour) ^a
1980	600	252	42	1,696.1	2.83	0.41
1981	264	140	53	1,214.9	4.60	0.67
1982	560	243	43	1,732.0	3.09	0.42
1984	22	12	55	87.5	3.98	0.54
1985	112	47	42	270.8	2.42	0.29
1994	77	37	48	228.5	2.97	0.79
1995	381	200	52	1,557.7	4.09	
1996	1,023	643	63	5,004.5	4.89	
1997	1,036	612	59	3,967.1	3.83	0.55
1998	1,556	819	53	5,846.1	3.76	0.56
1999	1,397	667	48	4,429.8	3.17	
2000	1,730	785	45	5,314.7	3.07	0.47
2001	1,633	669	41	4,299.4	2.63	
2002	931	340	37	1,928.2	2.07	
2003	609	258	42	1,449.4	2.38	
2004	620	281	45	1,724.1	2.78	
2005	902	436	48	2,761.2	3.06	
2006	792	358	45	2,156.9	2.72	
2007	119	60	50	366.1	3.08	
Total Mean ^b	14,364	6,859	48	46,035.0	3.20	0.47

^aOn observed days only.

^bAverages weighted by sample day.

skill accurately and implications for cultural patterns in hunter-gatherers.

The Ache Sample

Ache hunting has been extensively described and analyzed during the past quarter century. The sample of hunters analyzed here includes most adult men in three study communities from 1980 to 2007 (see CA+ online supplement B, table B1, available as an Excel file or a tab-delimited ASCII file). Some Ache men, who were over age 20 in 1980, have contributed to the database for the entire 27-year study period (CA+ online supplement A, figs. A1–A15), while others were monitored only in some years. Sample dates and sample size for individual men vary according to whether they resided in a particular community when we were monitoring hunting activities. The data reported here come from two sources: direct observation (1980–2000) and systematic interviews (1994–2007). The observational data were recorded by an anthropologist who accompanied hunters and recorded all game acquired and time spent hunting each day. These observational data have been described mostly in prior publications (Hill 1988, 2002; Hill and Hawkes 1983; Hill et al. 1984, 1987; Walker et al. 2002). The interview data were collected continuously through the year by Ache research assistants who recorded the species, age, and sex for all game acquired, as well as the number of days spent hunting regardless of whether prey were obtained. Game weights for the interview-based data were assigned on the basis of averages (by species, age, and sex) obtained from the observational

database. Some results using this method have also been published (Hill and Padwe 2000; Hill, McMillan, and Farina 2003; Hill et al. 1997).

Previous research showed that recall of prey obtained on a particular day is accurate for at least 2 weeks and that interview and observational data give very similar results when collected during the same time period (McMillan 2001, 60–73). For this reason, we combined all data into a single database from which we culled partial days and reservation day hunts, as well as occasional days hunting with guns or hunting during some other activity (e.g., game census work). This left 14,364 full hunter days on multiday forest treks where hunters used only traditional bow and arrow (dogs are not employed by Ache hunters). This sample includes 147 different hunters that were monitored between 1 and 1,049 days each over the study period (table 1). Because the number of hours hunted each day was not recorded in the interview-based data set, here we analyze the rate of kilograms of prey (whole carcass weight) acquired per day of hunting. This mean daily return rate corresponds closely to the measured mean hourly return rate of observed hunting trips for 114 men whose hourly hunting time was recorded between 1980 and 2000 ($r^2 = 0.98$; CA+ online supplement A, fig. A16). Thus, analyses of mean daily return rate are equivalent to analyses of mean hourly hunting return rate. The mean number of hours spent hunting on all observed hunting trips from 1980 to 2000 (2,314 days) was 6.84 hours/day. In contrast to our previous report (Hill and Hawkes 1983), there is no relationship between mean hours hunted per hunting day and

mean hourly return rate for the men in this sample; however, better hunters did contribute more hunting days to our sample (CA+ online supplement A, figs. A17, A18). Because better hunters contributed more hunting days to the sample, the day-weighted average return rate for the entire sample is 3.2 kg meat/day, whereas the mean hunter-weighted return rate is only 2.82 kg meat/day. Although variation across some years is evident in table 1 (decreasing and then increasing return rates from 1994 to 2007), there is no consistent trend of increase or decrease across the entire study period (slope = -0.0006 , $r^2 = 0.007$). Year-to-year variation seems to reflect mainly fluctuations in game densities and is at least partially due to weather patterns (Hill, McMillan, and Farina 2003).

Methods and Results

There is a strong convex age pattern to Ache hunting return rates over the life of a single individual because of learning and then senescence (Gurven, Kaplan, and Gutierrez 2006; Walker et al. 2002). As a consequence, we cannot simply average the age monitored and the observed return rate for each hunter during all time periods and then compare men to see whether good and poor hunters can be statistically distinguished. The averaged hunting returns for men who were observed at both young and old ages would always be low, despite an age average that would fall in the middle of the range where hunting returns are typically highest. To avoid this problem, we instead aggregated the data for each hunter into 5-year time blocks, beginning with 1980–1984 (hereafter referred to as a “hunter-period”). In this article, we consider whether “good” and “poor” hunters can be distinguished in a 5-year hunter-period, given a particular sample size of hunting days monitored. We define good and poor hunters as those who would be characterized by high and low mean

hunting return rates if we were able to monitor them long enough to gain an accurate measure of their “true” long-term return rate.

The frequency distribution of prey weight acquired per day shows a highly skewed pattern. No meat was obtained on just more than 50% of the >14,000 days hunted (table 1); successful hunting days show an approximately log-normal distribution of kilograms of meat obtained per day (fig. 1). The modal acquisition amount is 4–8 kg for days on which some meat was obtained. The sample also shows a low frequency of highly successful days (maximum = 191 kg because of a tapir kill). The extreme values in mean return rates for each hunter-period are mainly due to small sample size. For example, there is an 80-fold difference between the best hunters and the worst hunters for the set of men who acquired some meat but were sampled less than 10 days in a hunter-period (~35% of them got no meat at all). However, there is only about a fivefold range in mean return rates for men who were sampled at least 100 days in a hunter-period, and all men acquired some meat (see fig. 2). This suggests that large differences between “best” and “worst” hunters reported in populations that have been monitored for short time periods are often due to sampling error.¹

Because the frequency distribution of daily meat acquisition is far from normal, parametric statistical procedures are not

1. Sample size differences between men in our study were primarily due to some communities being sampled continuously while others were sampled sporadically. However, some small sample sizes are probably due to men who hunted infrequently *because* they were poor hunters (see CA+ online supplement A, fig. A18). While this could explain the lower minimum observed return rate at small sample sizes in figure 2, it cannot explain the 2.5-fold increase in maximum observed return rates at small sample sizes.

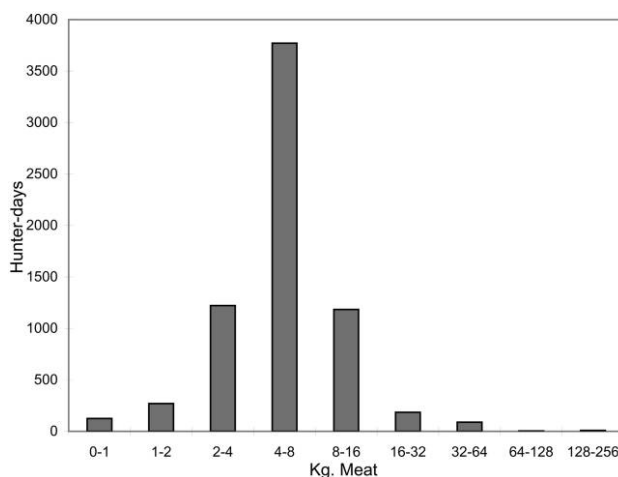


Figure 1. Frequency distribution of kilograms of meat obtained per day of hunting by Ache men on days when at least some meat was acquired (nonzero days).

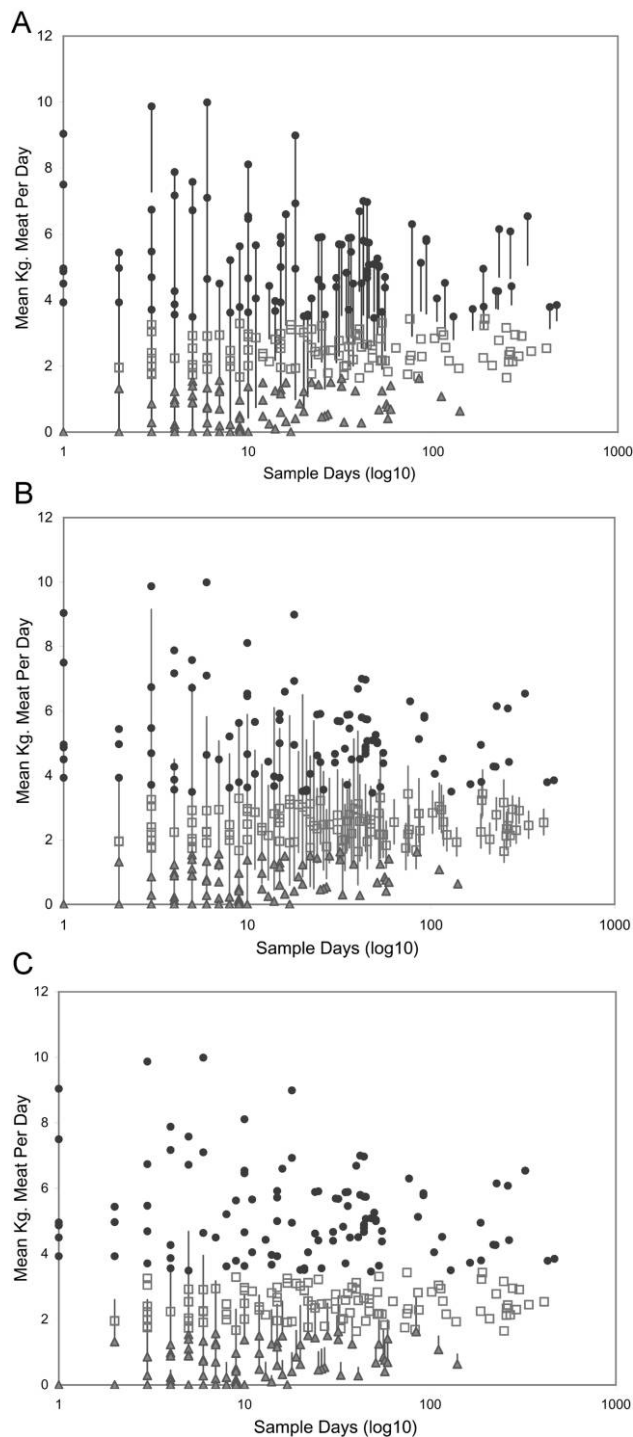


Figure 2. Measured mean daily hunting returns and confidence intervals during 5-year periods by sample size (days) for each of 283 hunter-periods where at least one kill was observed. The 95% confidence intervals (CIs) around the mean are shown for the top third (A; lower CI only), middle third (B), and lower third (C; upper CI only) of measured hunter-periods. *Circles*, top third; *squares*, middle third; *triangles*, lower third. A color version of this figure appears in the online edition.

appropriate for obtaining a confidence interval (CI) around each hunter-period mean return rate (cf. Hill and Hurtado 1996, 333). In order to estimate the CI around each man's measured return rate, we instead employed a resampling procedure. For each hunter-period, we drew 10,000 samples whose size was dictated by the number of days that individual hunted in that period. The samples were drawn, with replacement, from the hunter's empirical set of daily hunting return values for that interval. For each of these 10,000 samples, a return rate was determined. Using the set of 10,000 bootstrapped return rates, the 95% CI for the mean return rate of the hunter-period ranges from the 2.5 percentile value of the return rates to the 97.5 percentile value in the cumulative frequency distribution of the 10,000 results. For each of the 332 hunter-periods in our sample, the mean return rate from the 10,000 samples was virtually identical to the mean measured return rate. Figure 2 shows the 95% CI by sample size for all hunter-periods. The top, middle, and lower thirds of measured return rates are shown separately for better visual resolution. The CIs of the top and bottom thirds of hunters show that these men can be statistically distinguished by hunting skill after about 25 sample days in a 5-year period. But neither the top third nor the bottom third can be statistically distinguished from the middle third of monitored hunters with less than about 100 sample days per man.

This database also allows us to estimate the correlation coefficient between return rate estimated from a small sample (X_{measured}) and return rate calculated from a large sample of days (X_{true}) for individual men. This provides a way of assessing the reliability of return rate measures based on different-sized samples. We sampled blocks of 10–170 days of hunting return rate data and correlated these with the overall measures of return rate for the same hunter periods based on at least 100 sample days and at least twice the total sample size as the block size. Figure 3 shows that the correlation between shorter and longer measures of return rates increases steeply until about 80 data days and then slowly levels off to an r^2 value of about 0.9 at highest sample sizes.

Finally, it would be useful to know the number of sample days required in order for the upper and lower confidence limits to be within a specified percentage of a hunter's estimated mean return rate. For example, how many sample days do we need for an Ache hunter in order to be 95% certain that his true mean return rate is $\pm 20\%$ of the measured mean? The answer to this can be determined by regression. Because the bootstrap method produces asymmetrical confidence limits, we define the relative size of the upper and lower limits of the CI, ϕ , for a hunter as

$$\phi = \frac{|\text{mean} - \text{CI limit}|}{\text{mean}}.$$

A log-log plot of ϕ by sample size for all hunter-periods shows a linear relationship (fig. 4; R^2 for $\phi_{\text{upper}} = 0.68$, R^2 for $\phi_{\text{lower}} = 0.72$). The analyses show that observations of ≥ 43

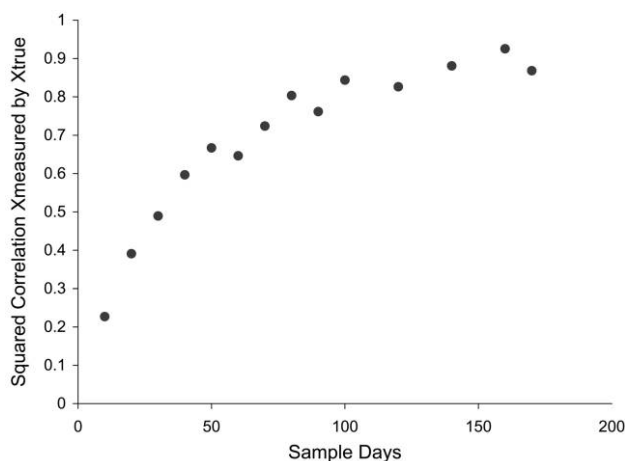


Figure 3. Squared correlation (r^2) between short-term measures of hunting return rate (X_{measured}) and long-term measures during the same 5-year period (X_{true}) for all men who contributed at least 100 data days in each time period. Long-term measures are based on a sample of 100 days and a sample that is more than double the size of the short-term measure. The X-axis indicates the number of sample days in each short-term measure (X_{measured}).

days are required, on average, in order for both 95% upper and lower confidence limits to be within $\pm 50\%$ of the measured mean return rate for a hunter. This is clearly problematic. Because about half of all Ache men sampled show a measured daily return rate within $\pm 50\%$ of the population mean, none of these men could be reliably distinguished from each other without a sample of more than 43 days of hunting. Sample size must be very large if we desire even more accurate knowledge about each man's true return rate. For example, the sample size necessary for CI limits $\pm 20\%$ of a hunter's mean is approximately 260 days of hunting data on each man (fig. 4).

The log-log plot of ϕ by sample size shows a slope of $-1/2$ (ϕ_{upper} slope = -0.50 , ϕ_{lower} slope = -0.49). This is consistent with expectations from statistical theory if we assume that the central limit theorem applies to large bootstrap sample sizes. Recall that the formula for standard error (SE) is $SE = SD \times n^{-1/2}$ (where n = sample size and SD = standard deviation) and that approximately 2 SE above and below the mean defines the 95% CI. Thus, the relative 95% CI limits, ϕ , for any hunter should simply be $k \times n^{-1/2}$ (where $k = [2 \times SD]/\text{mean}$). Thus, a log-log plot of ϕ by sample size should produce a linear fit whose height is $\log k$, a function of the variation in that sample relative to the mean, and whose slope should be approximately $-1/2$. This mathematical relationship allows us to quickly estimate sample sizes needed in order to achieve specified relative CIs in other populations with daily hunting variance higher or lower than that of the Ache.

A Comparison with the !Kung

The sample sizes required in order to distinguish good, medium, and poor hunters are determined by levels of daily variation in hunting return relative to the mean return rate. However, daily variance is determined by characteristics of prey, hunting techniques, and technology; thus, the required sample sizes for distinguishing hunting skill are likely to be specific to each ethnographic group. Other hunter-gatherers that specialize in larger and more dispersed game show significantly higher daily hunting variance relative to their daily mean acquisition rate, and thus even larger samples would be required in order to confidently distinguish men in those groups by their hunting abilities. For example, Hawkes, O'Connell, and Blurton Jones (2001a, 121) report that most Hadza game obtained is from large mammals (mean = 142 kg) and that hunters typically kill only one large-game animal per month on average.

One published database from the Dobe !Kung (Ju/'hoansi) can be used to make a quick comparison with the Ache. John Yellen (1977, app. B) followed !Kung hunters on forays from the Dobe water hole in 1967–1968 and published man-by-man and day-by-day hunting data on 213 person days of hunting taken from 10 different men. In that sample, !Kung hunters failed to acquire any meat on 73% of all hunting days, but they still obtained a mean daily return rate of 4.6 kg/day as the result of occasional kills of large prey. The days on which !Kung men obtained some meat show a frequency distribution considerably more spread out than that among the Ache (even when logged), with a higher frequency of

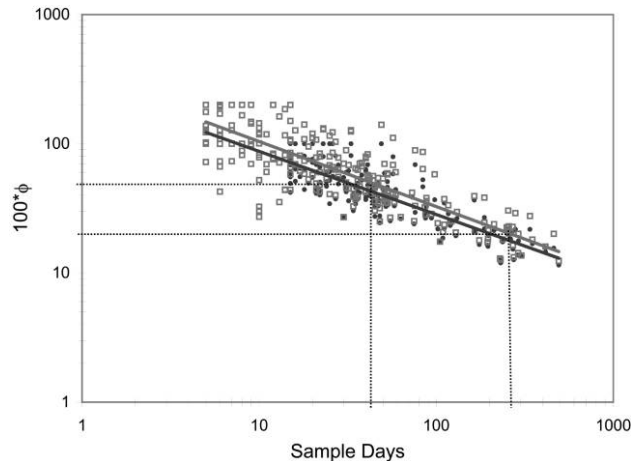


Figure 4. Estimated relative size of 95% confidence limits, ϕ , plotted by number of observed hunting days (N) in each hunter-period. Because the lower confidence interval (CI_{lower}) cannot take a value less than 0 ($CI < 100\%$ of the mean), only sample sizes of $N \geq 15$ days are shown for ϕ_{lower} . The slope for ϕ_{upper} is $Y = 331.79 \times N^{-0.5036}$ and for ϕ_{lower} is $Y = 268.9 \times N^{-0.4885}$. Dotted lines show linear models and required sample sizes to achieve hunting return rate estimates of mean $\pm 50\%$ and mean $\pm 20\%$ for individual men. *Squares*, values for upper CI; *circles*, values for lower CI. A color version of this figure appears in the online edition.

acquisition of large prey (cf. fig. 5 with fig. 1). Figure 6 shows the mean daily return rate and 95% CI (using the resampling method) for each of the eight !Kung men who killed at least one animal during the sample period. Despite a 10-fold range in hunting returns for these men, only the worst and the second-best hunters can be statistically distinguished from any of the others. This sample includes a man who was lauded by Lee (1979, 249) as a great hunter, but, even with 60 observation days, his mean return rate was statistically distinguishable from only four of the seven other hunters.

Using the power relationship between ϕ and sample size explained above, we can estimate the number of observation days for !Kung hunters required in order to obtain 95% confidence limits less than $\pm 50\%$ of their mean return rate. Given the height of the regression line on the log-log plot and projecting a line with a $-1/2$ slope through the eight available points (fig. 7), we estimate that approximately 110 days of data for a !Kung hunter would be needed for CIs $\pm 50\%$ of the mean and that about 600 days of data are needed for CIs $\pm 20\%$ of mean return rate. This is more than twice the sample size required for the same level of precision obtained monitoring Ache men.

Discussion

Our results have both methodological and theoretical implications. Although many evolutionary/sociocultural theories suppose that more skilled hunters will gain certain advantages over less skilled hunters, those theories are difficult to test

because reliable measures of hunting skill are not easily obtained. In this article, we have considered only whether men in an intensively monitored hunter-gatherer population can be distinguished from each other, given daily variation typical in the ecological context of our study group. Other problems would complicate things further. For example, if each man hunts only when conditions meet a critical threshold of individual success, then poor hunters may not hunt on days likely to give low returns, but good hunters might hunt on all days. In that case, voluntary measures of return rates would compress the true difference in ability between the best and the worse hunters. Likewise, even greater fluctuations due to seasonal and yearly effects, as well as long-term age-related changes in hunting success, might make distinguishing hunters by skill even more problematic than we observed in the Ache.

Hypothesis testing about the adaptive and cultural significance of hunting should be conducted with caution. Paired or grouped comparisons, such as the better half of hunters versus the worse half, will be problematic if a high fraction of men are incorrectly assigned to each skill level. Our data suggest that comparisons between the top third and the bottom third of hunters are less prone to noise from sample error. Furthermore, ordinary least squares (OLS) regressions using hunting return rate as the independent variable will produce incorrect slopes if the data points of greatest leverage (best and worst hunters) reflect mainly sampling error or when the hunting measure used for men contains a great deal of measurement error. Some previous studies are probably based on sample sizes too small and too noisy to achieve a statistically significant relationship. For example, Bailey (1991, table 6.2) reports no association between hunting and marriage parameters on the basis of a sample of only 19 hunters and as few as 15 days per hunter,² and Hawkes, O'Connell,

2. Assuming about 6 hours hunting per day.

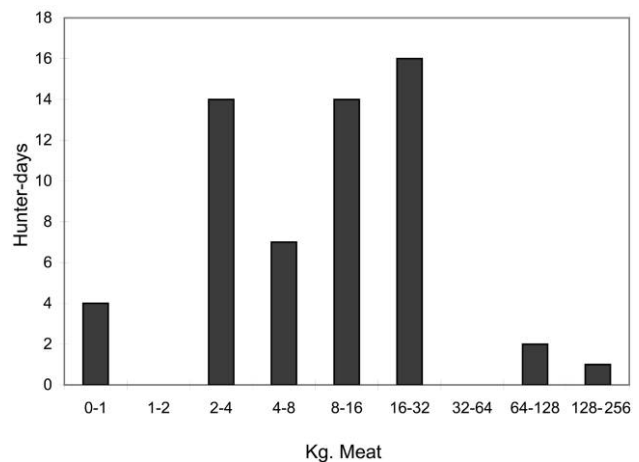


Figure 5. Frequency distribution of kilograms of meat obtained per day of hunting by !Kung men on days when at least some meat was acquired (data from Yellen 1977).

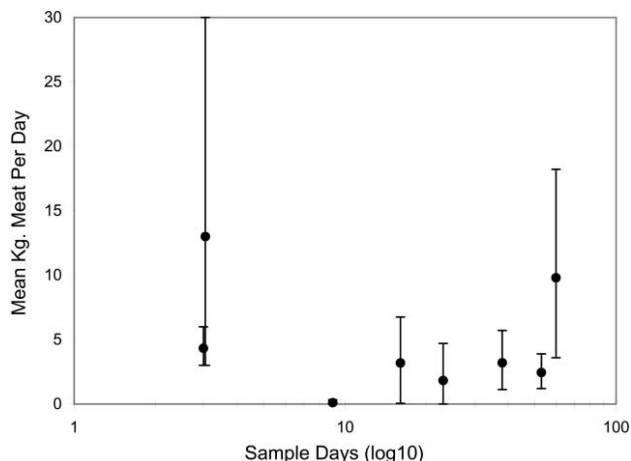


Figure 6. Measured mean daily hunting returns and 95% confidence intervals for each of eight !Kung men that made at least one kill during the sample period (data from Yellen 1977).

and Blurton Jones (2001*b*) present regression results on approximately 14 married men with 21 or more days of sampled hunting returns. With such small samples, we must be careful not to accept the null because we are unable to reject it. This is especially relevant to studies that conclude that there is no evidence of direct gains from successful hunting and that instead posit other hypothetical gains (e.g., signaling benefits) as the primary explanation for men's hunting.

When such regressions fail to achieve a threshold P value, it is critical to report the regression coefficient as well. Regression coefficients, not P values, are the best guide as to the magnitude of a likely effect if there is one (P values provide information only about whether the regression slope might be due to sample error). But, with significant error in measured return rates, even the regression slope will generally underestimate the effect of hunting skill on other long-run variables of theoretical interest. Because the regression coefficient is calculated as $\beta = \text{Cov}(Y, X) / \text{Var}(X)$, increased random noise and, hence, increased scatter in the X dimension (hunting return rate) will cause a lowered regression coefficient. This problem can be corrected. If the "true return rate" for a subset of men is known, we can then estimate the relation between the measured return rate (X_{measured}) and the actual return rate (X_{true}) by assuming that very large samples yield good approximations of X_{true} (as we illustrate in fig. 3). The β coefficient obtained using the noisy observed hunting data as the independent variable can then be corrected by dividing by the square of the correlation between X_{true} and X_{measured} , which is also equal to the ratio of variances $\text{Var}(X_{\text{true}}) / \text{Var}(X_{\text{measured}})$, known as the reliability ratio (Angrist and Krueger 1999, 1,344). Likewise, when the relationship between X_{measured} and X_{true} is unknown but the measurement error in X is likely to be considerable, some corrective procedure, such as reduced major axis (RMA) regression (McArdle 1988; Sokal and Rohlf 1981), is required in order

to estimate the true slope of the relationship between long-term hunting differences and some independent variable.

Because of noisy hunting measures, we suggest that previous regression coefficients of the association between hunting return rates and other dependent variables have been significantly underestimated in most anthropological studies. For example, a simple OLS regression of lifetime reproductive success by age-corrected hunting return rate for a cohort of Ache men that have finished reproducing gives an OLS β coefficient of only 0.20 when based on 36 older men with hunting sample sizes from 10 to 100 days (CA+ online supplement A, fig. A19). If we run the same regression using only 19 men with between 100 and 1,000 days of hunting data, the regression β coefficient is 0.83 (CA+ online supplement A, fig. A20), and if the β of the small-sample OLS regression is corrected by the reliability ratio for 10-day samples versus 100-day samples, we get an estimate of $\beta = 0.86 (= 0.20/0.227)$. Finally, the β coefficient for RMA regression on the full 55-man data set is 2.41. Although the three correction methods (employing larger samples, dividing β by the reliability ratio, and RMA regression) give different estimated regression coefficients, they all imply that the slope of the relationship is much steeper than that estimated by employing OLS regression on the uncorrected hunting data set with considerable noise due to small sample size.

Given these problems, one might consider whether informant rankings are a better way to distinguish hunters by skill. Unfortunately, most reported associations between informant rank and measured hunting returns are statistically significant only because of a few exceptionally good or poor hunters (e.g., Blurton Jones, Hawkes, and O'Connell 1997; Hill and Hurtado 1996; Marlowe 1999, 2000). However, informant rankings are often based on longer periods of observational

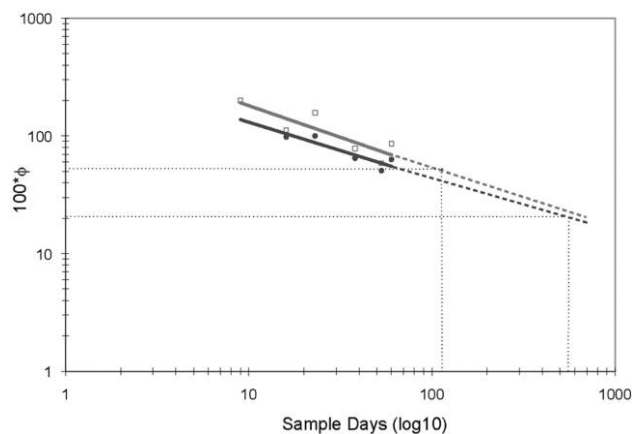


Figure 7. Log-log plot of estimated relative size of the 95% confidence limits, ϕ , plotted by number of observed hunting days for !Kung men (as in fig. 4). The slope for ϕ_{upper} is $Y = 621.4 \times N^{-0.537}$ and for ϕ_{lower} is $Y = 395.6 \times N^{-0.481}$. Dotted lines show extrapolated linear models and required sample sizes to achieve estimates of mean $\pm 50\%$ and mean $\pm 20\%$. A color version of this figure appears in the online edition.

experience than the sample size in most anthropological databases. Thus, informant rankings might be expected to show stronger associations with related dependent variables than would regressions using measured return rate data (see, e.g., Gurven and Von Rueden 2006; Hill and Hurtado 1996, 328).

Our study also has important implications for cultural patterns in hunting societies. If scientists cannot easily determine who is a better or a worse hunter for the majority of men in a sample, then the members of the population may not be able to either. Although exceptional hunters (or inept hunters) probably stand out easily, the vast majority of the population often will appear equivalent in skill if judged by direct observation. In band-level societies, various ecological and social forces regularly bring people together who have not observed each other extensively and who must have a very noisy assessment of each other's hunting skill. This inability to distinguish most hunters by skill levels may be important for understanding the evolution of cultural transmission biases that emphasize reputation and prestige of potential models rather than observational assessments of skill (Henrich and McElreath 2003).

The difficulties of determining relative hunting skill may also facilitate the emergence and spread of cultural sharing rules applied to hunted game. It is often suggested that skilled hunters should resist bandwide food redistribution without some direct compensation for their above-average provisioning over long time periods. But, most of the individuals who reside together long enough to truly ascertain differences in hunting skill are close kin, who are expected to share extensively because of kin selection. Those who are not kin may reside together on a less permanent basis and have ambiguous information about each other's relative hunting skills. One of the easiest ways to avoid conflict over food in such groupings is through cultural conventions that impose groupwide redistribution of game. Such conventions reduce the negative effects of daily intake variance to everyone's benefit (Smith 1988; Winterhalder 1997) and also eliminate potentially costly conflicts that arise when food is obtained asynchronously. Because few hunters can be certain how their hunting production will compare to that of their compatriots over some period of coresidence, most men might accept and adhere to culturally imposed groupwide distribution patterns. The ubiquitous presence of meat-sharing conventions in hunting societies may indicate that the benefits of such rules are greater than the individual temptation to consume and defend only one's own prey items when differential hunting success is difficult to ascertain.

In short, it is not easy for anthropological researchers to distinguish a population of hunters in small-scale society by absolute or even relative hunting skill during a moderate period of observation. Difficulties of distinction become even greater over longer time periods because the age trajectory of men's improvement and decline in hunting skills through the life span is heterogeneous and because illness and injury can turn an otherwise excellent hunter into a poor one overnight

(K. Hill, M. Barton, and A. M. Hurtado, unpublished manuscript; Sugiyama 2004; Sugiyama and Chacon 2000). Measuring the true fitness benefits of higher hunting productivity using short-term measures of hunting returns may be possible only with a large sample size of men and appropriate correction of the *X* variable in regression analyses. The difficulty of accurately assessing differences in hunting skill in small-scale human societies with fluctuating temporal patterns of coresidence has important implications for our understanding of cultural transmission and food-sharing patterns in our ancestral past.

Acknowledgments

K.H. thanks the Ache men of Arroyo Bandera, Chupa Pou, and Kue Tuvy for 30 years of cooperation. This article benefited from discussions with Mike Gurven, Eric Smith, and Tom Hertz.

References Cited

- Angrist, J. D., and A. B. Krueger. 1999. Empirical strategies in labor economics. *In* Handbook of labor economics, vol. 3A. O. Ashenfelter and D. Card, eds. Pp. 1277–1365. Amsterdam: Elsevier.
- Bailey, R. C. 1991. The behavioral ecology of Efe pygmy men in the Ituri Forest, Zaire. *Anthropological Papers, Museum of Anthropology*, no. 86. Ann Arbor: University of Michigan.
- Bird, R. B., E. A. Smith, and D. W. Bird. 2001. The hunting handicap: costly signaling in human foraging strategies. *Behavioral Ecology and Sociobiology* 50:9–19.
- Blurton Jones, N. G., K. Hawkes, and J. F. O'Connell. 1997. Why do Hadza children forage? *In* *Uniting psychology and biology: integrative perspectives on human development*. N. L. Segal, G. E. Weisfeld, and C. C. Weisfeld, eds. Pp. 297–331. New York: American Psychological Association.
- Codding, B. F., and T. L. Jones. 2007. Man the show-off? or the ascendance of a just-so-story: a comment on recent applications of costly signaling theory in American archaeology. *American Antiquity* 72(2):349–357.
- Dwyer, P. D., and M. Minnegal. 1993. Are Kubo hunters "show-offs"? *Ethology and Sociobiology* 14:53–70.
- Gurven, M., H. Kaplan, and M. Gutierrez. 2006. How long does it take to become a proficient hunter? implications for the evolution of extended development and long life span. *Journal of Human Evolution* 51:454–470.
- Gurven, M., and C. Von Rueden. 2006. Hunting, social status and biological fitness. *Social Biology* 53:81–99.
- Hawkes, K., and R. Bird. 2002. Showing off, handicap signaling, and the evolution of men's work. *Evolutionary Anthropology* 11:58–67.
- Hawkes, K., J. F. O'Connell, and N. G. Blurton Jones. 2001a. Hadza meat sharing. *Evolution and Human Behavior* 22: 113–142.

- . 2001*b*. Hunting and nuclear families: some lessons from the Hadza about men's work. *Current Anthropology* 42:681–709.
- Hawkes, K., J. O'Connell, K. Hill, and E. Charnov. 1985. How much is enough? hunters and limited needs. *Ethology and Sociobiology* 6:3–16.
- Henrich, J., and R. McElreath. 2003. The evolution of cultural evolution. *Evolutionary Anthropology* 12:123–135.
- Hill, K. 1988. Macronutrient modifications of optimal foraging theory: an approach using indifference curves applied to some modern foragers. *Human Ecology* 16(2):157–197.
- . 2002. Cooperative food acquisition by Ache foragers. *Human Nature* 13(1):105–128.
- Hill, K., and K. Hawkes. 1983. Neotropical hunting among the Ache of eastern Paraguay. *In* Adaptive responses of native Amazonians. R. Hames and W. Vickers, eds. Pp. 139–188. New York: Academic Press.
- Hill, K., K. Hawkes, H. Kaplan, and A. Hurtado. 1987. Foraging decisions among Ache hunter-gatherers: new data and implications for optimal foraging models. *Ethology and Sociobiology* 8:1–36.
- Hill, K., and A. M. Hurtado. 1996. Ache life history: the ecology and demography of a foraging people. New York: Aldine de Gruyter.
- Hill, K., H. Kaplan, K. Hawkes, and A. Hurtado. 1984. Seasonal variance in the diet of Ache hunter-gatherers of eastern Paraguay. *Human Ecology* 12:145–180.
- Hill, K., G. McMillan, and R. Farina. 2003. Changes in large vertebrate densities over a five-year period in the Mbaracayu Reserve, Paraguay: hunting depletion or natural factors. *Conservation Biology* 17:1312–1323.
- Hill, K., and J. Padwe. 2000. Sustainability of Ache hunting in the Mbaracayu Reserve, Paraguay. *In* Sustainability of hunting in tropical forests. J. Robinson and E. Bennet, eds. Pp. 79–105. New York: Columbia University Press.
- Hill, K., J. Padwe, C. Bejyvagi, A. Bepurangi, F. Jakugi, R. Tykuarangi, and T. Tykuarangi. 1997. Monitoring hunting impact on large vertebrates in the Mbaracayu Reserve, Paraguay, using native research assistants. *Conservation Biology* 11(6):1339–1353.
- Hurtado, A. M., K. Hill, H. Kaplan, and I. Hurtado. 1992. Trade-offs between female food acquisition and child care among Hiwi and Ache foragers. *Human Nature* 3:185–216.
- Kaplan, H., and K. Hill. 1985. Hunting ability and reproductive success among male Ache foragers. *Current Anthropology* 26(1):131–133.
- Kaplan, H., K. Hill, J. Lancaster, and A. M. Hurtado. 2000. The evolution of intelligence and the human life history. *Evolutionary Anthropology* 9(4):156–184.
- Kent, S. 1996. Hunting variability at a recent sedentary Kalahari village. *In* Cultural diversity and twentieth-century foragers: an African perspective. S. Kent, ed. Pp. 125–156. Cambridge: Cambridge University Press.
- Lee, R. B. 1979. The !Kung San: men women and work in a foraging society. Cambridge: Cambridge University Press.
- Marlowe, F. 1999. Show-offs or providers? the parenting effort of Hadza men. *Evolution and Human Behavior* 20: 391–404.
- . 2000. The patriarch hypothesis: an alternative explanation of menopause. *Human Nature* 11:27–42.
- McArdle, B. H. 1988. The structural relationship: regression in biology. *Canadian Journal of Zoology* 66:2329–2339.
- McMillan, G. 2001. Ache residential grouping and social foraging. PhD dissertation, University of New Mexico, Albuquerque.
- O'Connell, J. E., K. Hawkes, K. Lupo, and N. G. Blurton Jones. 2002. Male strategies and Plio-Pleistocene archaeology. *Journal of Human Evolution* 43:831–872.
- Robinson, J. G., and E. Bennett, eds. 2000. Hunting for sustainability in tropical forests. New York: Columbia University Press.
- Smith, E. A. 1988. Risk and uncertainty in the "original affluent society": evolutionary ecology of resource sharing and land tenure. *In* Hunters and gatherers: history, evolution, and social change. T. Ingold, D. Riches, and J. Woodburn, eds. Pp. 222–252. Oxford: Berg.
- . 2004. Why do good hunters have higher reproductive success? *Human Nature* 15(4):343–364.
- Sokal, R. R., and F. J. Rohlf. 1981. *Biometry*. 2nd edition. New York: W. H. Freeman.
- Stanford, C. T., and H. T. Bunn. 2001. Meat-eating and human evolution. Oxford: Oxford University Press.
- Sugiyama, L. 2004. Illness, injury, and disability among Shiwiari forager-horticulturalists: implications of health-risk buffering for the evolution of human life history. *American Journal of Physical Anthropology* 123:371–389.
- Sugiyama, L., and R. Chacon. 2000. Effects of illness and injury on foraging among the Yom and Shiwiari: pathology risk as adaptive problem. *In* Human behavior and adaptation: an anthropological perspective. L. Cronk, N. Chagnon, and W. Irons, eds. Pp. 371–395. New York: Aldine de Gruyter.
- Walker, R., K. Hill, H. Kaplan, and G. McMillan. 2002. Age dependency of strength, skill, and hunting ability among the Ache of Paraguay. *Journal of Human Evolution* 42: 639–657.
- Wiessner, P. 1996. Leveling the hunter: constraints on the status quest in foraging societies. *In* Food and the status quest. P. Wiessner and W. Schiefenhovel, eds. Pp. 171–192. Providence, RI: Berghahn.
- . 2002. Hunting, healing and hxaro exchange: a long-term perspective on !Kung (Ju/'hoansi) large game hunting. *Evolution and Human Behavior* 23:407–436.
- Winterhalder, B. 1997. Gifts given, gifts taken: the behavioral ecology of nonmarket, intragroup exchange. *Journal of Archaeological Research* 5:121–168.
- Yellen, J. 1977. Archeological approaches to the present: models for reconstructing the past. New York: Academic Press.