

Life-history theory, fertility and reproductive success in humans

Beverly I. Strassmann^{1*} and Brenda Gillespie²

¹Department of Anthropology, and Mammal Division, Museum of Zoology, and ²Department of Biostatistics, School of Public Health, and Center for Statistical Consultation and Research, University of Michigan, Ann Arbor, MI 48109, USA

According to life-history theory, any organism that maximizes fitness will face a trade-off between female fertility and offspring survivorship. This trade-off has been demonstrated in a variety of species, but explicit tests in humans have found a positive linear relationship between fitness and fertility. The failure to demonstrate a maximum beyond which additional births cease to enhance fitness is potentially at odds with the view that human fertility behaviour is currently adaptive. Here we report, to our knowledge, the first clear evidence for the predicted nonlinear relationship between female fertility and reproductive success in a human population, the Dogon of Mali, West Africa. The predicted maximum reproductive success of 4.1 ± 0.3 surviving offspring was attained at a fertility of 10.5 births. Eighty-three per cent of the women achieved a lifetime fertility level (7–13 births) for which the predicted mean reproductive success was within the confidence limits (3.4 to 4.8) for reproductive success at the optimal fertility level. Child mortality, rather than fertility, was the primary determinant of fitness. Since the Dogon people are farmers, our results do not support the assumptions that: (i) contemporary foragers behave more adaptively than agriculturalists, and (ii) that adaptive fertility behaviour ceased with the Neolithic revolution some 9000 years ago. We also present a new method that avoids common biases in measures of reproductive success.

Keywords: life-history trade-offs; fertility; reproductive success

1. INTRODUCTION

Hunter-gatherers such as the !Kung of Botswana and the Ache of Paraguay are often thought to behave more adaptively than other contemporary populations. Yet demographic studies report that, in both groups, the relationship between the number of offspring born to a given woman and her lifetime reproductive success was linear with a substantial slope (Pennington & Harpending 1988; Hill & Hurtado 1996). This linear relationship suggests that !Kung and Ache women might have achieved higher fitness by increasing fertility beyond observed levels, which implies that their fertility behaviour may have failed to maximize fitness (see Hill & Hurtado 1996; Hill & Kaplan 1999).

The linear relationship between fitness and fertility is also surprising from the point of view of life-history theory, which holds that inclusive fitness-maximizing organisms face a trade-off between the number of offspring born and the number successfully reared (reviewed in Lessells (1991) and Stearns (1992)). If the Ache and the !Kung were not under-reproducing, then life-history theory predicts diminishing returns to fitness from higher fertility, which would produce either a plateau or a downturn in fitness at high fertility. The classic example of this trade-off is the work carried out by Lack (1947) on clutch size, which shows that altricial birds can achieve higher reproductive success by laying an intermediate number of eggs per nest. Birds who produce too large a clutch are unable to feed each nestling adequately and fledge fewer young.

In a pioneering study, Blurton Jones & Sibly (1978)

tested for reproductive trade-offs among the !Kung. They concluded that !Kung birth spacing was adaptive because the modal interbirth interval yielded the most offspring who survived to the age of 10 years. Although this study has not been without controversy (Harpending 1994; Hill & Hurtado 1996; Blurton Jones 1997), it helped motivate subsequent tests of life-history theory in humans, including more recent work among the !Kung and Ache (Pennington & Harpending 1988; Hill & Hurtado 1996).

These later studies did not replicate the finding of optimal fertility and raise the possibility that most !Kung and Ache women did not maximize reproductive success. The evidence, however, is inconclusive: unmeasured heterogeneity among women may explain why these studies found a linear relationship between fitness and fertility (Borgerhoff Mulder 1992; Hill & Hurtado 1996; Blurton Jones 1997; Hill & Kaplan 1999). If people vary in health, wealth, kin support or other critical parameters, then the women who have the highest fertility may also have the greatest opportunity for rearing offspring. Each woman might reproduce at her own personal optimum, but the aggregate relationship between fertility and fitness for the population as a whole might be linear. Life-history trade-offs faced by individual women, such as the trade-off between offspring number and offspring quality, may be masked by the phenotypic correlation between resources and reproductive success in the population as a whole (van Noordwijk & de Jong 1986; Borgerhoff Mulder 1992; Hill & Hurtado 1996; Hill & Kaplan 1999). Animal behaviourists have overcome the problem of phenotypic correlation through controlled experiments in which fertility levels are manipulated (Gustaffson & Sutherland 1988; Pettifor *et al.* 1988; Dijkstra *et al.* 1990). In studies

* Author for correspondence (bis@umich.edu).

of humans, heterogeneity among women may be controlled by multiple regression and other statistical methods. Although Hill & Hurtado (1996) did attempt appropriate statistical controls, either these controls were incomplete or Ache women did not actually maximize fitness either in the precontact 'forest period' or after relocation to missions.

Since the Ache study is one of the most sophisticated examples of the application of life-history theory to human demography, the absence of conclusive results presents an interesting challenge. In testing predictions from evolutionary theory in contemporary environments, human behavioural ecologists assume that human behaviour is sufficiently flexible and sensitive to environmental contingencies to permit adaptive outcomes, so long as major novelties such as modern contraception are absent (Smith 2000). Evolutionary psychologists, by contrast, emphasize that humans are adapted to an ancestral forager lifestyle and assume that adaptive behaviour ceased with the onset of agriculture about 9000 years ago (Symons 1989; Barkow *et al.* 1992; Buss 1995).

Here, our goal is to test these alternative assumptions by examining the adaptiveness of fertility behaviour in a contemporary agricultural population, the Dogon of Mali, West Africa. In line with the behavioural ecological approach, we expect the relationship between female fertility and reproductive success to conform to predictions from inclusive fitness and life-history theory. Specifically, we test the hypothesis that there is a trade-off between the number of offspring born and the number who survive. This hypothesis predicts that the relationship between fertility and reproductive success is either inverse U-shaped or increases to a plateau, rather than being linear.

We also present a new method for examining the relationship between fertility and reproductive success that employs a form of survival analysis (Kaplan & Meier 1958) to avoid significant biases that are commonly ignored. This method gives fertility-specific measures of reproductive success and allows the inclusion of individuals who have incomplete reproductive histories. Hence it is particularly useful for long-lived species. It may also be helpful when fitness is measured in the grand-offspring generation, as, at the time of sampling, few individuals have a complete tally of their surviving grand-offspring.

The Dogon have lived along the Bandiagara cliff for over half a millennium and grow millet as their staple crop. In 1998, 68% of the married men in the study village adhered to the traditional religion, which entails ancestor worship, masked dancing and the seclusion of menstruating women in designated huts. Both indigenous and Western contraception were entirely absent and most reciprocity was among close kin. The region is remote, *ca.* 250 km south of Timbuktu, and despite a superficial influx of tourists, the Dogon have remained one of the most traditional peoples of Africa (Paulme 1940; Calame Griaule 1965; Pern 1982; Strassmann 1992, 1996, 1997*a,b*; Cazes 1993).

2. METHODS

(a) *The subjects and their reproductive histories*

The study population was the set of women, aged 20 years and older, who were resident in the village of Sangui (14°29' N,

3°19' W) at any time during a 10 year period from 1988 to 1998. Reproductive and marital histories of these women were conducted by one of the authors (B.I.S.) in 1988, 1994 and 1998. A total of 167 women were resident in the village at the time of the interviews in at least one of these 3 years. A few women were briefly married to men of Sangui at some point during the 10 year study period but could not be included in the analysis because they were not resident at the time of any of the interviews. Three elderly women, mean age 69 years, did not provide data on age at death for deceased children and were therefore excluded from the analyses, together with one woman who by the age of 37 years had never had a pregnancy. We estimate that the set of 167 women includes about 95% of the women in the entire study population.

Most of the women were present for interviews in more than 1 year and their data were updated to be as complete as possible. Interviews of the same women in different years were compared and found to be remarkably consistent. B.I.S. lived for nearly 3 years in the study village, knew most of the women individually and conducted the interviews in Dogon (Sangha-Sò). Data on stillbirths and miscarriages are sometimes useful for assessing the accuracy of reproductive data. Among 64 Dogon women whose youngest child was born at least 5 years prior to the last interview, the mean \pm s.d. for the number of stillbirths and miscarriages experienced per woman were 0.42 ± 0.97 and 0.63 ± 1.10 , respectively. High values such as these tend to reflect good subject cooperation and reporting accuracy.

(b) *Age estimates*

Dogon women born in Sangha can identify their 'age mates' and their relative age with respect to members of their own and other age cohorts. Sets of age mates never include more than one interbirth interval, as two consecutive siblings with the same mother are always assigned to different age classes. Thus the Dogon have a very precise scheme for ranking relative age. Through information on the known year of birth for some individuals, particularly men who had gone to school and women born during a salient year such as Malian Independence (1960), the year of the Sigi Dance (1969) and the Great Famine (1914), we were able to determine actual birth years with unusual accuracy for a non-literate population.

(c) *Reproductive success: lifetime method*

We used linear regression to model reproductive success as a function of fertility. We investigated the shape of this relationship by including fertility alone or fertility and fertility squared in the regression models (Kaplan *et al.* 1995). To further investigate functional forms that may be neither linear nor quadratic, we used loess regression (Cleveland 1993), which gives a non-parametric estimate for the shape of the function. In most cases, the loess and quadratic functions (Cleveland 1993) were almost identical, so we usually present only the latter. We calculated the women's reproductive success through three different methods that we call 'lifetime', 'Kaplan-Meier' and 'age adjustment'. In all analyses we used the statistical software Spss 10.0 (SPSS 1999).

For lifetime reproductive success, we included only women for whom 5 or 10 years had elapsed since their last live births ($n = 65$ and $n = 55$ women, respectively). Their reproductive success was measured by counting the number of offspring who survived to age 5 or 10 years, respectively.

In studies of lifetime reproductive success, researchers often neglect to restrict the sample to females whose offspring have

all had time to reach the age to which survival is measured. However, this procedure underestimates reproductive success if offspring who have not yet reached this age are not counted. Alternatively, scoring these offspring as having survived to the given age, or simply counting the number of offspring alive at the time of sampling, would exaggerate reproductive success because survival is as yet undetermined. An offspring who was born longer ago will be more likely to be deceased than an offspring born more recently. Treating two such offspring equivalently would inflate the reproductive success of mothers whose offspring were still young compared with that of mothers of older offspring.

(d) Reproductive success: the Kaplan–Meier method

Although the lifetime method should be viewed as the gold standard, it inevitably reduces sample sizes. We present a novel method that gives unbiased estimates of reproductive success for a given fertility level, even when the sample includes women who are not yet post-reproductive or whose youngest child has not yet reached the age to which survival is measured.

To obtain an estimate of cumulative survival to age 10 years using this method, the oldest offspring must have been born at least 10 years prior to the mother's last interview. Out of the 163 potential women, this method yielded sample sizes of 132 and 117 women for survival to age 5 and 10 years, respectively. A significant advantage of this method over the lifetime method is that the increase in sample size gives tighter confidence intervals for reproductive success.

The Kaplan–Meier method has two steps. First, for each woman we obtained Kaplan–Meier estimates (Kaplan & Meier 1958) of the expected proportion of her offspring surviving to age 5 or 10 years. Second, to estimate reproductive success, we multiplied the Kaplan–Meier estimates by the number of live births for each woman. Among women who were 5 or 10 years post-reproductive, the lifetime and Kaplan–Meier values are identical.

(e) Reproductive success: age-adjustment method

An alternative approach might adjust for mother's age in regression models of reproductive success on fertility. Like the Kaplan–Meier method, this approach permits the inclusion of women who have not completed their reproduction. However, the Kaplan–Meier method uses information on all offspring up to the point of death or censoring, whereas the age-adjustment method only counts offspring of age 10 years or older (or the age to which survival is measured). The number of offspring of age 10 years or older will generally increase with the mother's age, and the assumption of the age-adjustment model is that this increase is linear (during the reproductive years). However, the number of children over the age of 10 years depends not only on age and fertility, but also on the timing of births. Variation among women in birth timing will mean that age adjustment is at best approximate. Other problems with this method may include interaction and collinearity between age and fertility. Despite these potential flaws, we present results using this method for the purpose of comparison.

(f) Correction for the influence of mortality on fertility

Previous research has established that when an infant ceases to nurse and dies, a Dogon mother usually resumes both ovulation and menstruation within two months, but if the child

lives, the median duration of postpartum amenorrhoea is 20 months (Strassmann 1996; Strassmann & Warner 1998; see Perez *et al.* (1972) for Chilean women). Thus, when a correlation is observed between mortality and fertility, causation may go in both directions. To test the possibility that high maternal fertility increases offspring mortality, we controlled for the possibility that women who lose nursing infants get pregnant again more quickly and thus have higher fertility. We assumed that if a child died at age 2 years or older, (s)he was weaned or nearly so and death did not shorten the time until the next birth. We therefore subtracted from each woman's number of live-born offspring any child who died in less than 24 months unless the child who died young was a last-born offspring. The death of a last born would not hasten a subsequent birth. To help avoid over-correction, if two children in a row both died at under 2 years, we subtracted only one of them from the woman's fertility. In the Kaplan–Meier analyses of survival to age 10 years ($n = 117$ women), a mean (\pm s.d.) of 1.1 ± 1.1 live births were subtracted per woman, or a total of 15% of all live births. In the analyses of lifetime reproductive success ($n = 55$ women), 1.4 ± 1.2 live births were subtracted per woman, or a total of 16% of all live births. We then compared the results obtained in this fashion with results based on fertility without the correction. Whereas the correction is useful for understanding the mechanism that underlies the correlation between fertility and mortality, the actual fertility data may be more useful for testing the optimality of fertility behaviour.

3. RESULTS

Figure 1 shows the relationship between corrected fertility and reproductive success with offspring survival measured to age 10 years (figure 1*a,b*) or 5 years (figure 1*c,d*). Figure 1*a,c* uses the lifetime method and figure 1*b,d* employs the Kaplan–Meier method. In both methods, the x -axis is fertility corrected for the influence of mortality on fertility (see § 2). Figure 2 shows the same relationship as in figure 1*a*, but without the fertility correction.

(a) Lifetime method

According to the lifetime method, reproductive success (based on survival to age 10 years) increased with corrected fertility and then levelled off such that additional births did not enhance reproductive success (figure 1*a*). Reproductive success reached a maximum of 4.3 children at 8.6 births. From a frequency distribution of corrected fertility we found that 27% of women had nine or more births, and thus reached the predicted maximum. But a fertility of six births yielded a reproductive success within the confidence interval (3.7 to 4.8) for this maximum (figure 1*a*). Only 25% of the women had fewer than six births, which leads to the striking conclusion that most of the women achieved a fertility level for which mean reproductive success was statistically indistinguishable from reproductive success at the optimal fertility level. The regression equation was $y = -1.36 + 1.30x - 0.08x^2$ (y , reproductive success; x , fertility) with p -values of 0.002 and 0.022 for the linear and squared terms, respectively ($n = 55$, $R^2 = 0.32$). To control for secular trends, the mother's year of birth (z) was added to the model, and the equation became $y = -73.12 + 1.52x - 0.09x^2 + 0.04z$ with p -values of 0.001, 0.009 and 0.076, respectively ($n = 55$, $R^2 = 0.36$). Thus, when reproductive success was

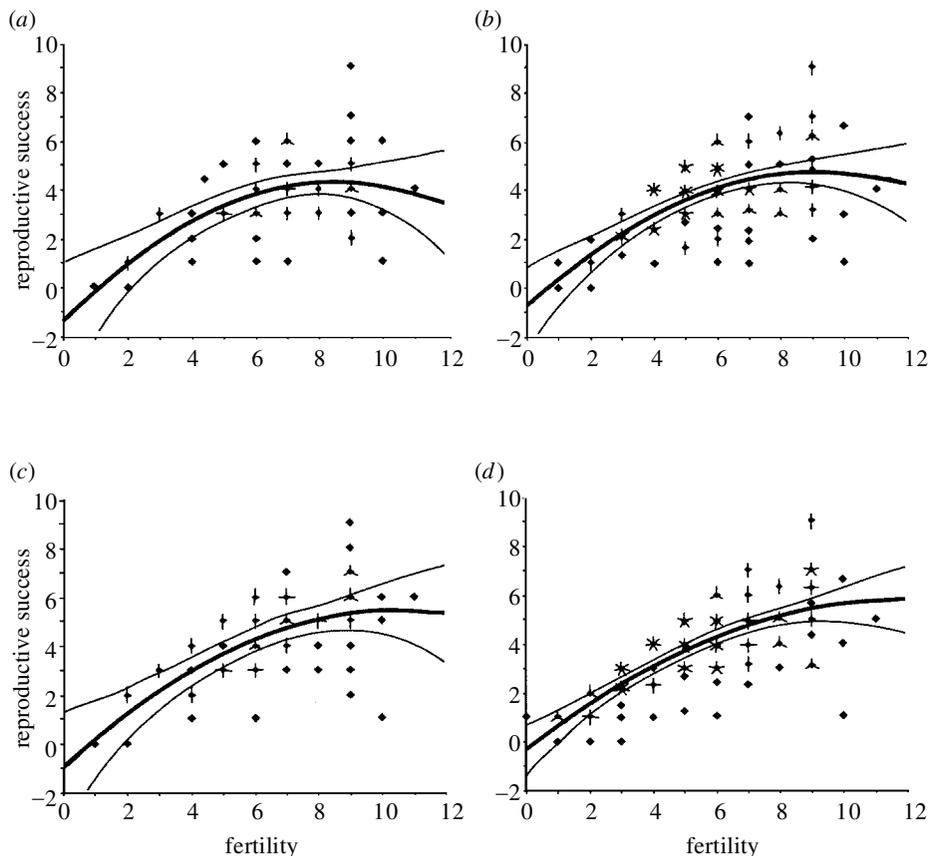


Figure 1. The relationship between fertility and reproductive success: regression line (bold) plus 95% confidence limits. Each 'petal' on a data point represents an additional case. (a) Lifetime method with survival to the age of 10 years: the regression equation is $y = -1.36 + 1.30x - 0.08x^2$ (y , reproductive success; x , fertility) with p -values of 0.002 and 0.022 for the linear and squared terms, respectively ($n = 55$, $R^2 = 0.32$). (b) Kaplan-Meier method with survival to the age of 10 years: the regression equation is $y = -0.91 + 1.23x - 0.07x^2$ with p -values of <0.001 and 0.008 for the linear and squared terms, respectively ($n = 117$, $R^2 = 0.32$). (c) Lifetime method with survival to the age of 5 years: the regression equation is $y = -0.99 + 1.23x - 0.06x^2$ with p -values of 0.002 and 0.059 for the linear and squared terms, respectively ($n = 65$, $R^2 = 0.41$). (d) Kaplan-Meier method with survival to the age of 5 years: the regression equation is $y = -0.37 + 1.03x - 0.04x^2$ with p -values of <0.001 and 0.016 for the linear and squared terms, respectively ($n = 132$, $R^2 = 0.52$). The loess function (Cleveland 1993) (not shown) was almost identical to the quadratic fit for graphs *a-d*.

measured as the number of offspring who survived to age 10 years, the relationship between corrected fertility and reproductive success was quadratic, both with and without adjustment for the mother's year of birth.

When offspring survival was measured only to age 5 years (figure 1c), the regression equation for the lifetime method was: $y = -25.52 + 1.28x - 0.06x^2 + 0.01z$ (y , reproductive success; x , fertility; z , mother's year of birth) and the p -values for the coefficients were 0.002, 0.051 and 0.435, respectively ($n = 65$, $R^2 = 0.41$). Regardless of whether survival was measured to age 5 or 10 years, the relationship was quadratic, but in the former case the coefficient for the squared term was slightly smaller.

The results for actual fertility without the correction were very similar to those for corrected fertility, regardless of whether we used survival to age 10 years (figure 2) or to age 5 years (data not shown). Reproductive success (based on survivorship to age 10 years) reached a maximum of 4.1 surviving offspring at 10.5 births. From a frequency distribution of actual fertility, we found that 35% of women had 10 or more births, and 10 was the modal fertility. Median fertility was somewhat lower at 8.5 births. Eighty-three per cent of women attained a fertility

level (7 to 13 births) for which the predicted reproductive success was within the confidence limits (3.4 to 4.8) for reproductive success at the maximum. No woman had more than 13 births.

(b) *Kaplan-Meier method*

The Kaplan-Meier method gave tighter confidence intervals (figure 1b,d), but otherwise gave results that were similar to those for the lifetime method. Using corrected fertility, reproductive success (based on survival to age 10 years) reached a maximum of 4.7 surviving offspring at 9.1 births. A fertility of 7 to 11 births yielded a reproductive success that was within the confidence limits (4.3 to 5.1) for the maximum (figure 1b). According to this method, 56% of the women reached a fertility level (7 to 11 births) for which the predicted mean reproductive success was statistically indistinguishable from reproductive success at the optimal fertility level. The regression equation for the relationship between corrected fertility and reproductive success was $y = -57.30 + 1.24x - 0.06x^2 + 0.03z$ (y , reproductive success; x , fertility; z , mother's year of birth) with p -values of <0.000 , 0.014 and 0.004

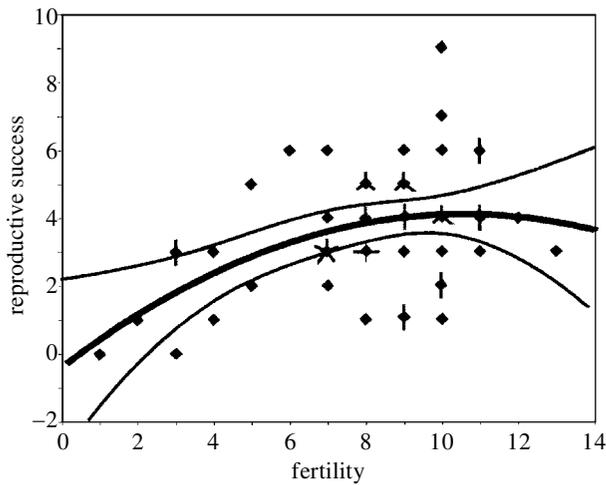


Figure 2. The relationship between lifetime fertility and reproductive success (based on survivorship to the age of 10 years): regression line (bold) plus 95% confidence limits. In contrast with figure 1, this graph shows actual fertility (uncorrected for the influence of mortality). Each ‘petal’ on a data point represents an additional case. Controlling for birth cohort, the regression equation is $y = -72.87 + 1.03x - 0.05x^2 + 0.04z$ (y , reproductive success; x , fertility; z , mother’s year of birth) with p -values of 0.010, 0.071 and 0.112 for fertility, fertility squared and the mother’s year of birth, respectively ($n = 55$, $R^2 = 0.25$).

for fertility, fertility squared and the mother’s year of birth, respectively ($n = 117$, $R^2 = 0.36$).

When survival was measured only to age 5 years (figure 1d) the regression equation for the Kaplan–Meier method was: $y = -31.79 + 1.09x - 0.04x^2 + 0.02z$ with p -values for the coefficients of 0.000, 0.016 and 0.071, respectively ($n = 132$, $R^2 = 0.53$). Thus, both the lifetime and Kaplan–Meier methods indicate a quadratic relationship between corrected fertility and reproductive success, regardless of the age (5 or 10 years) to which survivorship was measured.

Using actual rather than corrected fertility, reproductive success at the optimum fertility (11.7 births) was 4.7 ± 0.4 offspring surviving to age 10 years. The predicted reproductive success for women who had seven births was 3.9, which was within the confidence limits (3.8 to 5.5) for reproductive success at the maximum. As with the lifetime method, 83% of women attained a fertility level (7–13 births) that was statistically indistinguishable from the optimum fertility level.

(c) Age-adjustment method

The results for the age-adjustment method are illustrated in figure 3 with a separate loess fit (Cleveland 1993) for three age groups (under 30, 30 to 40, and over 40 years). This method shows that the curve for the relationship between corrected fertility and reproductive success among women over the age of 40 increases to a plateau (figure 3). For women aged 30–40 years, the relationship appears more linear because many of these women had not finished reproducing. For women under 30 years, there is no relationship between fertility and reproductive success because the highest corrected fertility was only three births and few offspring had had time to reach an age of 10 years. Figure 3 also shows that, for a given fertility level, the older women had higher reproductive

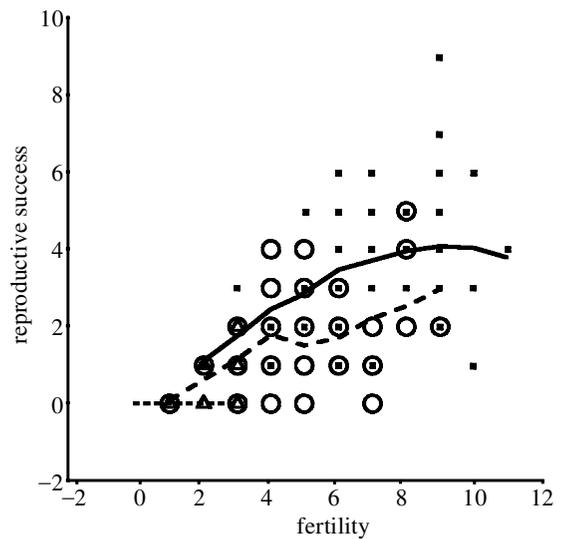


Figure 3. The relationship between fertility and reproductive success: loess line (Cleveland 1993) for each of three age categories (bold line and squares, over 40 years; dashed line and circles, 30–40 years; dotted line and triangles, less than 30 years; $n = 117$).

Table 1. Predictors of reproductive success (the number of offspring who survived to the ages of 5 and 10 years) by the Kaplan–Meier method.

predictor	age of 5 years ($n = 119^d$, $R^2 = 0.61$)		age of 10 years ($n = 104^d$, $R^2 = 0.46$)	
	β	p	β	p
mother’s year of birth	0.03	0.005	0.05	0.000
1st wife ^{a,d}	-0.61	0.032	-0.73	0.027
2nd, 3rd, 4th wife ^{a,d}	-0.84	0.003	-0.88	0.010
arranged wife ^{b,d}	-1.68	0.003	-1.72	0.006
non-arranged wife ^{b,d}	-1.48	0.008	-1.30	0.033
fertility ^c	1.13	0.000	1.25	0.000
fertility squared ^c	-0.05	0.006	-0.06	0.007
constant	-56.73	0.006	-89.27	0.000

^a Versus the omitted category, sole wife.

^b Versus the omitted category, levirate wife.

^c Corrected for the influence of mortality on fertility as described in the text.

^d Women who changed marital status during the study were assigned their marital status during two of three interviews. Thirteen widows were excluded from the analysis because they were not married at the time of the interview and therefore had missing values for the marital status variables.

success. This occurs simply because there was more time for their offspring to have reached the age of 10 years.

(d) Fertility and reproductive success after controlling for covariates

Table 1 shows the relationship between corrected fertility and Kaplan–Meier reproductive success after

Table 2. Predictors of reproductive success (the number of offspring who survived to the age of 10 years) by the lifetime and age-adjustment methods.

predictor	lifetime method ($n = 42^f$, $R^2 = 0.52$)		age-adjustment method ($n = 104^f$, $R^2 = 0.48$)	
	β	p	β	p
mother's age ^a	—	—	0.06	0.014
mother's year of birth ^b	0.05	0.071	—	—
1st wife ^{c,f}	-0.22	0.687	-0.67	0.041
2nd, 3rd, 4th wife ^{c,f}	-0.58	0.325	-0.71	0.033
arranged wife ^{d,f}	-2.42	0.019	-1.64	0.008
non-arranged wife ^{d,f}	-1.45	0.107	-1.20	0.046
fertility ^e	1.50	0.001	0.59	0.050
fertility squared ^e	-0.08	0.019	-0.02	0.295
constant ^f	-93.51	0.073	-0.37	0.767

^a In the age-adjustment method, the mother's age at last interview was truncated at 50 years because women older than 50 years did not gain additional offspring who survived to the age of 10 years.

^b In the lifetime method, there was no need to adjust for mother's age because all the women were post-reproductive, but we adjusted for mother's year of birth to model birth cohort effects.

^c Versus the omitted category, sole wife.

^d Versus the omitted category, levirate wife.

^e Corrected for the influence of mortality on fertility as described in the text.

^f Women who changed marital status during the study were assigned their marital status during two of three interviews. Thirteen widows were excluded from the analysis because they were not married at the time of the interviews and therefore had missing values for the marital status variables in those years.

controlling for other predictors of reproductive success using multiple linear regression. Regardless of whether we measured survival to age 5 or to age 10, the same set of predictors emerged as significant. These included: the woman's year of birth, her marital status (first wife, second/third/fourth wife, versus the reference category, sole wife), and whether she was an arranged, non-arranged, or levirate wife (a woman who was married to her deceased husband's brother).

The women in the study were born between 1908 and 1969 ($n = 119$). For each decade later that a woman was born, reproductive success based on survival to age 10 increased by 0.5 offspring. First wives had 0.73 fewer surviving offspring and second or higher-order wives had 0.88 fewer surviving offspring than sole wives. Women in arranged marriages had 1.72 fewer surviving offspring and women in non-arranged marriages had 1.30 fewer surviving offspring than levirate wives. Arranged and non-arranged wives did not differ significantly from each other, so levirate wives were chosen for the reference category. Other covariates were not significant, including the wealth of the husband's family (measured as metric tons of grain harvested or relative wealth rank). The significant

Table 3. The contribution of fertility^a and mortality^b to reproductive success^c ($n = 55$ women).

reproductive success	n (women)	mean fertility	s.d. fertility	mean mortality	s.d. mortality
0	2	2.0	1.41	1.0	0.00
1	6	7.0	3.22	0.7	0.36
2	4	8.0	2.45	0.7	0.08
3	16	7.5	2.66	0.5	0.27
4	12	9.4	1.51	0.6	0.07
5	7	8.0	1.41	0.4	0.08
6	6	9.0	2.10	0.3	0.21
7-9	2	10.0	0.00	0.2	0.14
all women	55	8.1	2.56	0.5	0.26

^a Fertility is defined as the actual number of live births.

^b Mortality is defined as one minus the proportion surviving to the age of 10 years.

^c Reproductive success is defined as the number of offspring who survived to the age of 10 years for women who last gave birth to a live-born offspring at least 10 years prior to the interview.

covariates explain 46% of the variance in Kaplan–Meier reproductive success to the age of 10 years and 61% of the variance in Kaplan–Meier reproductive success to the age of 5 years. Fertility and fertility squared remained significant ($p < 0.01$) after controlling for the other covariates, which provides further evidence for a trade-off between offspring number and offspring survival.

Consistent with the smaller sample size, the p -values for the covariates were less significant when the lifetime method (table 2) instead of the Kaplan–Meier method (table 1) was employed. The coefficients, however, were in the same direction and similar in magnitude. After controlling for the same covariates as above, fertility and fertility squared were significant predictors of the number of offspring who survived to the age of 10 years, with p -values of 0.001 and 0.019, respectively (table 2).

The covariates that were significant in the age-adjustment method were the same as those in the Kaplan–Meier method, except that fertility was only borderline significant ($p = 0.05$) and fertility squared was not significant ($p = 0.29$; table 2). By including younger women, the age-adjustment method obscures the diminishing returns to reproductive success at high fertility observed among women over the age of 40 years (figure 3). This method assumes parallel curves for different ages, with equal intervals between ages, which is an assumption not supported by figure 3. It also has the problem of collinearity between fertility and the mother's age at the last interview ($r = 0.60$, $p < 0.001$, $n = 117$), which makes it difficult to assess the relationship between fertility and reproductive success independent of age. Collinearity between the mother's age and year of birth means that this method does not permit adjustment for cohort effects.

(e) *Relative contribution of fertility and mortality*

Table 3 explores the relationship between fertility (total uncorrected number of live births), mortality (one minus the proportion surviving to the age of 10 years), and reproductive success (number of offspring who survived to the age of 10 years). In this table we used the

post-reproductive sample of 55 women whose last live births occurred at least 10 years prior to their last interviews. Two women who did not raise any offspring to 10 years had one and three live births. At the other end of the spectrum, the two women who raised seven and nine children to the age of 10 years, both had 10 live births and an offspring mortality rate of only 20%. In between these two extremes, reproductive success ranged from 1 to 6 offspring, fertility from a mean (\pm s.d.) of 7.00 ± 3.22 to 9.00 ± 2.10 live births, and mortality varied from 0.72 ± 0.36 to 0.26 ± 0.21 .

The fertility of women who had six surviving offspring was only 29% higher, on average, than that of the women who had only one surviving offspring. Child mortality in the latter group, however, was 2.8-fold higher. Although both fertility and mortality contribute to reproductive success, mortality evidently played a larger role.

4. DISCUSSION

In contrast with other species (reviewed in Stearns (1992)), there is little evidence for optimal fertility behaviour in humans. Reproductive trade-offs, such as the trade-off between fertility and offspring survivorship, have also been difficult to document. One exception is the evidence for strong selection for singleton as opposed to twin births, which explains why humans usually have a litter size of only one child (Haukioja *et al.* 1989; Anderson 1990). Other attempts to test for optimal fertility behaviour have led to negative or ambiguous results. For example, Kaplan *et al.* (1995) tested the hypothesis that low fertility in Anglo and Hispanic men in urban New Mexico (USA) maximized the production of grand-offspring. Instead, they found a linear relationship between male fertility and the number of grand-offspring born. A trade-off was observed, but only between parental fertility and offspring education and income (Kaplan *et al.* 1995).

The failure of optimality models in societies that have undergone the demographic transition is not altogether surprising. The negative results for hunter-gatherers, however, contradict the assumption that humans flexibly adjust fertility levels so as to maximize inclusive fitness (Smith 2000). Human behavioural ecologists assume that adaptive fertility behaviour is not confined to the Palaeolithic, or any other specific period of the human evolutionary past. Instead they predict that it occurs in a wide range of environments, as long as modern contraception and other radical departures from past human experience are absent.

Despite theoretical reasons for expecting adaptive fertility behaviour among foragers, the first test has been difficult to interpret (Blurton Jones & Sibly 1978; Blurton Jones 1986, 1987) and two subsequent studies reported no evidence for optimization (Pennington & Harpending 1988; Hill & Hurtado 1996). The strengths and limitations of the original study have been thoroughly explored (Blurton Jones 1994, 1997; Harpending 1994; Hill & Hurtado 1996), but of particular concern is the failure of the attempts at replication. Regardless of whether these negative results reflect differences in ecology or issues of methodology, there is surprisingly little empirical support

for optimal fertility behaviour in the forager populations so far studied (Hill & Hurtado 1996, p. 396).

A recent investigation of the Kipsigis, who are agropastoralists of Kenya, examined the yield in grandchildren for women who had different numbers of surviving offspring (Borgerhoff Mulder 2000). As data on female fertility were not included in the analysis, this study does not address the relationship between female fertility and fitness, as measured in either the offspring or the grand-offspring generation. However, it does address reproductive trade-offs. Specifically, it reports a nonlinear relationship between the number of offspring who survived to the age of 5 years and the number of surviving grand-offspring ($n = 64$ women). Women who had 8–11 surviving offspring had slightly fewer grand-offspring than women who had six or seven surviving offspring. Although this difference was not quite significant ($p = 0.08$), it is consistent with diminishing returns to fitness from raising additional offspring. The Kipsigis study examines the effects of offspring competition after the age of 5 years. Our study, by contrast, examines competition from birth up to the age of 10 years, which is when most child mortality takes place. The former approach is more useful for exploring competition among adult siblings for heritable resources such as land or cattle. Our approach is more useful for exploring the optimality of fertility behaviour.

In summary, none of the foregoing studies reported that the observed number of live births maximized fitness. Our results demonstrate a clear-cut trade-off between the number of offspring born and the number who were successfully reared. Thus, they address a significant gap between theory and data. Unlike studies of the Ache and the !Kung, women at the highest observed fertility level did not produce more surviving offspring than did women at intermediate fertility levels.

After the correction for the influence of mortality on fertility, the predicted maximum lifetime reproductive success of 4.3 ± 0.3 surviving offspring was attained at a fertility of 8.6 births. Seventy-five per cent of the post-reproductive women ($n = 55$) achieved a fertility level (6–11 births) for which the predicted reproductive success was within the confidence limits (3.7 to 4.8) for reproductive success at the optimal fertility level. Optimal fertility for the Kaplan–Meier method was 9.1 births. Using this method, 56% of the women ($n = 117$) achieved a fertility level (7–11 births) for which the predicted reproductive success was within the confidence limits (4.3 to 5.1) for reproductive success at the maximum. The lower percentage of women displaying optimal fertility with the Kaplan–Meier method (56% versus 75%) is partly due to the higher estimate for optimum births (9.1 versus 8.6) and partly due to the narrower width of the confidence interval for reproductive success (0.8 versus 1.1).

Using the lifetime method and actual fertility without the correction, optimum fertility was estimated to be 10.5 births, which agreed closely with the modal fertility of 10 births and was somewhat higher than the median fertility of 8.5 births. Reproductive success at this optimum was 4.1 ± 0.3 surviving offspring. Eighty-three per cent of the women attained a fertility level (7–13 births) for which the predicted reproductive success was within the confidence limits (3.4 to 4.8) for reproductive success at the maximum. The Kaplan–Meier method also found that

83% of the women attained a fertility level (7–13 births) that was statistically indistinguishable from the optimum.

We realize that whether any individual woman reproduced at her own personal optimum is impossible to say. Moreover, some women reproduced more efficiently than others, with fewer births and fewer deaths. However, the overall fertility behaviour of the women, as observed in the aggregate, was fitness maximizing. Aside from Blurton Jones' analysis of interbirth intervals (Blurton Jones & Sibly 1978; Blurton Jones 1986, 1997), this is, to our knowledge, the first demonstration of optimal fertility behaviour for a human population.

Regardless of whether we used the lifetime or Kaplan–Meier estimates of reproductive success, measured survival to the age of 5 or 10 years, or based our analysis on corrected fertility or actual fertility, the relationship between fertility and reproductive success was best described by a quadratic function. Although a quadratic function models a downturn in reproductive success with high fertility levels, our estimated quadratic functions were all flattened near the maximum. The women in our dataset did not have fertility levels significantly beyond the estimated maxima, so it was not possible to statistically distinguish between a plateau and an inverse U-shape. The fact that few women had greater than optimal fertility implies that few women made the mistake of over-reproducing.

To make sure that the relationship we observed was not an artefact of a cohort effect or some other source of heterogeneity in the women sampled, we controlled for other significant predictors of reproductive success. Even after controlling for these covariates, the relationship between fertility and reproductive success remained quadratic.

In future research it might be helpful to ask mothers whether menstruation had resumed at the time of death for all children who died before the age of 3 years. The return of the post-partum menses closely coincides with the resumption of ovulation (Perez *et al.* 1972), so this information would permit a more fine-tuned correction for the influence of mortality on fertility. In the analyses of corrected fertility, we found that when survivorship was estimated only to the age of 5 years and no covariates were controlled, the quadratic relationship was less pronounced. Mean survivorship to the age of 10 years ($n = 55$ women) was only 6% lower than to the age of 5 years ($n = 65$ women); nonetheless, mortality between these ages increased the coefficient for the quadratic term. This result may imply that most mortality between the ages 5 and 10 years occurred among women with high fertility. It is also consistent with the interpretation that the quadratic relationship between corrected fertility and reproductive success was caused by the effect of fertility on mortality, rather than the converse. However, in the regressions with actual fertility, the quadratic relationship was more pronounced when survival was measured only to the age of 5 years.

Since the Dogon people are farmers, our results are at odds with the assumptions that: (i) contemporary foragers behave more adaptively than agriculturalists, and (ii) that adaptive fertility behaviour ceased with the Neolithic revolution some 9000 years ago. The tendency to assume that adaptation ceased with agriculture overlooks a number of issues (reviewed in Strassmann & Dunbar (1999)). One

of the more salient is that agriculture spread in Europe because Neolithic farmers out-reproduced the Neolithic hunter-gatherers (Menozzi *et al.* 1978; Sokal *et al.* 1991). Thus, agriculture appears to have been an adaptive change in mode of subsistence.

Humans did pass more generations as foragers than as farmers, but this alone does not establish the significance of the transition, which was only one of many transitions in our evolutionary past (Strassmann & Dunbar 1999). For example, most primates eat fruit, leaves and insects and very few species regularly hunt mammals (Boyd & Silk 1997, p. 397). Thus, most of our primate ancestors were probably vegetarians and insectivores who did not engage in cooperative foraging for meat. A social and dietary shift was required for the emergence of hunting and gathering, but it does not follow that this shift automatically induced maladaptive fertility behaviour. Similarly, in the absence of empirical evidence, it cannot be assumed that the life histories of agriculturalists show less evidence for adaptive design than those of foragers. Our data support the hypothesis of optimal fertility behaviour in a contemporary farming population, which contradicts the notion that agriculture is at the juncture of a crude dichotomy between adaptation and maladaptation.

Dogon first wives and higher-order wives had lower reproductive success than women in monogamous marriages. Research on the proximate mechanisms that underlie this cost of polygyny suggests a role for reduced paternal investment (Strassmann 1997*a,b*, 2002). Another significant finding is that arranged wives and non-arranged wives had lower reproductive success than did levirate wives. Widows who have many surviving offspring may be more likely to choose to become levirate wives upon their husband's death. Marrying the husband's close relative permits a widow to stay with her children since they belong to their father's patrilineage. A widow who has many surviving offspring may be more likely to remain with the lineage that was good for reproductive success and where she can continue to protect her prior reproductive investment. Due to increased statistical power, most covariates were more significant in the Kaplan–Meier analyses than in the analyses using the other methods, indicating that the Kaplan–Meier approach is particularly useful for measuring covariate effects.

Another significant finding is that offspring mortality, not fertility, was the primary determinant of variation in reproductive success in our sample. Most of the effect of mortality on reproductive success was direct, rather than mediated by the effect of mortality on fertility. In the Ache, by contrast, variation in fertility explained a greater proportion of the variation in reproductive success than did child mortality (Hill & Hurtado 1996, p. 393). We hypothesize that the situation found in the Dogon, in which it is offspring mortality that is the major determinant of fitness, is more typical of the human evolutionary past. This hypothesis assumes that, like the Dogon and the Ache, most of our ancestors did not live in populations characterized by high levels of venereal disease as in the African infertility belt (Caldwell & Caldwell 1983) or by high levels of permanent celibacy, as in the rural Irish (Strassmann & Clarke 1998).

One caveat is that our results may be biased towards women who lived long enough to be present during at

least one of our three interview sessions. However, this problem of selection bias was mitigated by our use of a semi-prospective design. We conducted three interview sessions over a 10-year period rather than a single cross-sectional one and therefore were able to interview about 95% of the eligible women. During the 10 years of this study, only three women in the study village died after their first live birth but prior to menopause. In a population with very high adult mortality, such as forest-dwelling Ache, the variation in female fertility would be largely caused by premature deaths. This type of pattern probably typifies much of the human past and, even if accompanied by higher child mortality, might augment the fertility component of fitness variance.

Our study focused on the relationship between female fertility and reproductive success. Two previous studies of the Dogon addressed the proximate mechanisms that underlie variation in female fertility and child mortality, but did not explore fitness effects (Strassmann 1997a,b; Strassmann & Warner 1998). Taken together, the three studies suggest that the quadratic relationship between fertility and reproductive success was probably caused by offspring competition for parental investment. In the mortality study, the chance of death for any given child increased by 26% as the number of children (aged 10 years or under) in the work-eat group (the set of people who are economically interdependent) increased by one additional child. As the dependency ratio, or the number of children per married adults, increased by one child, the likelihood of death increased 3.4-fold (Strassmann 2000).

In summary: (i) Dogon women face a trade-off between offspring quantity and offspring quality, and (ii) observed in the aggregate, ca. 83% of the women in our sample experienced lifetime fertility levels that were approximately optimal for maximizing fitness.

We thank the village of Sangui for continued hospitality and the women of Sangui, in particular, for their generous participation. Kim Hill, Claudius Vincenz and two anonymous reviewers offered helpful comments on the manuscript. This research was supported by the LSB Leakey Foundation, NSF BNS-8612291, and NSF SBR-9727229.

REFERENCES

- Anderson, D. J. 1990 On the evolution of human brood size. *Evolution* **44**, 438–440.
- Barkow, J. H., Cosmides, L. & Tooby, J. (eds) 1992 *The adapted mind: evolutionary psychology and the generation of culture*. New York: Oxford University Press.
- Blurton Jones, N. 1986 Bushman birth spacing: a test for optimal interbirth intervals. *Ethol. Sociobiol.* **7**, 91–105.
- Blurton Jones, N. 1987 Bushman birth spacing: direct tests of some simple predictions. *Ethol. Sociobiol.* **8**, 183–203.
- Blurton Jones, N. 1994 A reply to Dr Harpending. *Am. J. Phys. Anthropol.* **93**, 391–397.
- Blurton Jones, N. 1997 Too good to be true? Is there really a trade-off between number and care of offspring in human reproduction? In *Human nature: a critical reader* (ed. L. Betzig), pp. 83–86. New York: Oxford University Press.
- Blurton Jones, N. G. & Sibly, R. M. 1978 Testing adaptiveness of culturally determined behaviour: do Bushman women maximize their reproductive success by spacing births widely and foraging seldom? *Human behaviour and adaptation*, pp. 135–158. Society for the Study of Human Biology, Symposium 18. London: Taylor & Francis.
- Borgerhoff Mulder, M. 1992 Reproductive decisions. In *Evolutionary ecology and human behavior* (ed. E. A. Smith & B. Winterhalder), pp. 339–374. New York: Aldine de Gruyter.
- Borgerhoff Mulder, M. 2000 Optimizing offspring: the quantity–quality tradeoff in agropastoral Kipsigis. *Evol. Hum. Behav.* **21**, 391–410.
- Boyd, R. & Silk, J. B. 1997 *How humans evolved*. New York: Norton.
- Buss, D. M. 1995 Evolutionary psychology: a new paradigm for psychological science. *Psychological Inquiry* **6**, 1–30.
- Calame Griaule, G. 1965 *Ethnologie et langage: la parole chez les Dogon*. Paris: Editions Gallimard.
- Caldwell, H. C. & Caldwell, P. 1983 The demographic evidence for the incidence and cause of abnormally low fertility in tropical Africa. *World Hlth Statist. Q.* **36**, 2–34.
- Cazes, M. H. 1993 *Les Dogon de Boni: approche demo-genetique d'un isolat du Mali*. Paris: Presses Universitaires de France.
- Cleveland, W. S. 1993 *Visualizing data*. Summit, NJ: Hobart Press.
- Dijkstra, C., Bult, A., Bijlsma, S., Daan, S., Meijer, T. & Zijlstra, M. 1990 Brood size manipulations in the kestrel (*Falco tinnunculus*): effects on offspring and parent survival. *J. Anim. Ecol.* **59**, 269–286.
- Gustafsson, L. & Sutherland, W. J. 1988 The costs of reproduction in the collared flycatcher *Ficedula albicollis*. *Nature* **335**, 813–815.
- Harpending, H. 1994 Infertility and forager demography. *Am. J. Phys. Anthropol.* **93**, 385–390.
- Haukioja, E., Lemmetyinen, R. & Pikkola, M. 1989 Why are twins so rare in *Homo sapiens*? *Am. Nat.* **133**, 572–577.
- Hill, K. & Hurtado, M. 1996 *Ache life history: the ecology and demography of a foraging people*. New York: Aldine de Gruyter.
- Hill, K. & Kaplan, H. 1999 Life history traits in humans: theory and empirical studies. *A. Rev. Anthropol.* **28**, 397–430.
- Kaplan, E. L. & Meier, P. 1958 Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.* **53**, 457–481.
- Kaplan, H. S., Lancaster, J. B., Bock, J. A. & Johnson, S. E. 1995 Does observed fertility maximize fitness among New Mexico men? A test of an optimality model and a new theory of parental investment in the embodied capital of offspring. *Hum. Nature* **6**, 325–360.
- Lack, D. 1947 The significance of clutch size. *Ibis* **89**, 302–352.
- Lessells, C. M. 1991 The evolution of life histories. In *Behavioural ecology: an evolutionary approach*, 3rd edn (ed. J. Krebs & N. Davies), pp. 32–68. Oxford: Blackwell Scientific.
- Menozi, P., Piazza, A. & Cavalli-Sforza, L. L. 1978 Synthetic maps of human gene frequencies in Europeans. *Science* **201**, 786–792.
- Paulme, D. 1940 *Organisation sociale des Dogon*. Paris: Les Editions Domat-Montchrestien.
- Pennington, R. & Harpending, H. 1988 Fitness and fertility among Kalahari !Kung. *Am. J. Phys. Anthropol.* **77**, 303–319.
- Perez, A., Vela, P., Masnick, G. S. & Potter, R. G. 1972 First ovulation after childbirth: the effect of breastfeeding. *Am. J. Obstet. Gynecol.* **114**, 1041–1047.
- Pern, S. 1982 *Masked dancers of West Africa: the Dogon*. Peoples of the Wild Series. Amsterdam: Time-Life Books.
- Pettifor, R. A., Perrins, C. M. & McCleery, R. H. 1988 Individual optimization of clutch size in great tits. *Nature* **336**, 160–162.
- Smith, E. A. 2000 Three styles in the evolutionary analysis of human behavior. In *Adaptation and human behavior: an anthropological perspective* (ed. L. Cronk, N. Chagnon & W. Irons), pp. 27–46. New York: Aldine de Gruyter.

- Sokal, R. R., Oden, N. L. & Wilson, C. 1991 Genetic evidence for the spread of agriculture in Europe by demic diffusion. *Nature* **351**, 143–145.
- SPSS 1999 *SPSS advanced models 10.0*. Chicago, IL: SPSS, Inc.
- Stearns, S. C. 1992 *The evolution of life histories*. Oxford University Press.
- Strassmann, B. I. 1992 The function of menstrual taboos among the Dogon: defense against cuckoldry? *Hum. Nature* **3**, 89–131.
- Strassmann, B. I. 1996 Menstrual hut visits by Dogon women: a hormonal test distinguishes deceit from honest signaling. *Behav. Ecol.* **7**, 304–315.
- Strassmann, B. I. 1997a Polygyny as a risk factor for child mortality among the Dogon. *Curr. Anthropol.* **38**, 688–695.
- Strassmann, B. I. 1997b The biology of menstruation in *Homo sapiens*: total lifetime menses, fecundity, and nonsynchrony in a natural fertility population. *Curr. Anthropol.* **38**, 123–129.
- Strassmann, B. I. 2000 Polygyny, family structure, and child mortality: a prospective study among the Dogon of Mali. In *Adaptation and human behavior: an anthropological perspective* (ed. L. Cronk, N. Chagnon & W. Irons), pp. 49–67. New York: Aldine de Gruyter.
- Strassmann, B. I. 2002 The paternal contribution to child mortality in Africa. (In preparation.)
- Strassmann, B. I. & Clarke, A. L. 1998 Ecological constraints on marriage in rural Ireland. *Evol. Hum. Behav.* **19**, 33–55.
- Strassmann, B. I. & Dunbar, R. I. M. 1999 Human evolution and disease: putting the Stone Age in perspective. In *Evolution in health and disease* (ed. S. Stearns), pp. 91–100. New York: Oxford University Press.
- Strassmann, B. I. & Warner, J. 1998 Predictors of fecundability and conception waits among the Dogon of Mali. *Am. J. Phys. Anthropol.* **105**, 167–184.
- Symons, D. 1989 A critique of Darwinian anthropology. *Ethol. Sociobiol.* **10**, 131–144.
- van Noordwijk, A. J. & de Jong, G. 1986 Acquisition and allocation of resources: their influence on variation in life history tactics. *Am. Nat.* **128**, 127–142.

As this paper exceeds the maximum length normally permitted, the authors have agreed to contribute to production costs.