A Full Sibling is not as Valuable as an Offspring: On Hamilton's Rule

Xuhua Xia

NOTES AND COMMENTS

A FULL SIBLING IS NOT AS VALUABLE AS AN OFFSPRING: ON HAMILTON'S RULE

J. B. S. Haldane is said to have remarked that he would be willing to lay down his life for two brothers or eight cousins. This opinion was later formulated into the famous Hamilton's rule: $R B - C > 0$, which has been the guiding light for many studies in behavioral ecology. As a general guideline, Hamilton's rule is remarkably correct, but the quantitative aspect of the rule should not be over-stretched.

As an example of overstretching, Dawkins and Carlisle (1976), when criticizing a fallacy in Trivers (1972), employed the following argument based on Hamilton's rule: "Suppose a female has an orphaned baby brother, the same age as her own son. She has only enough food to keep one of the two infants alive. Which should she prefer? Intuition points to the son, but this is not necessarily correct. There are no genetic grounds for preferring either infant: the mother's relatedness to both is the same, 0.5" (Dawkins and Carlisle 1976, p. 131). While Dawkins and Carlisle (1976) were correct in criticizing Trivers (1972), they were wrong in a different way because there are good genetic grounds for preferring the son.

Let us first distinguish two different kinds of sibling helping. In the first kind, animals help their siblings directly so that the fitness of their current siblings is increased. This is the situation when the female in Dawkins and Carlisle's (1976) scenario helps her baby brother or when Haldane lays down his life for two brothers. In the second kind, animals help their siblings indirectly by enabling their parents to produce more or healthier siblings in the future. In both kinds the helper can potentially increase its inclusive fitness through its siblings, but, strictly speaking, only the first kind is true sibling helping. In this note, I will present a single-locus model to illustrate why helping offspring is evolutionarily more advantageous than helping full siblings. Specifically, I will show two costs, one deterministic and the other stochastic, associated with this first kind of sibling helping.

THE SINGLE-LOCUS MODEL

Suppose a locus has two alleles, designated O and S, in an animal population with discrete generations. An OO genotype will provide two doses of help to its

offspring, an SS genotype will provide two doses of help to its full siblings, and an OS genotype will provide one dose of help to its offspring and another dose to its full siblings. Otherwise, the three genotypes are identical to each other, and, consequently, their fitness increment (Z) through helping others, according to Hamilton's rule, follows the same equation:

$$Z = RB - C, \tag{1}$$

where B and C are constants and R is the probability of the recipient carrying the allele of its helper (Grafen 1991). Animals mate randomly, and each animal, regardless of genotype, has L offspring and $(L - 1)$ full siblings as potential recipients of its help. Let p and q be the allelic frequency of O and S, respectively, in the whole population.

DETERMINISTIC COST

Consider first an OO genotype in a population. Let $p_{L.oo}$ and $q_{L.oo}$ be the allelic frequency of O and S alleles, respectively, in offspring of an OO parent. These offspring will inherit one copy of O from the OO parent, and the probability that the other parent will carry an O allele is p. Therefore, an OO parent produces two kinds of offspring, OO offspring with frequency of p and an OS genotype with frequency of q. The expected allelic frequency of O among offspring of an OO genotype is then

$$p_{L.oo} = 0.5 + 0.5p = R_{oo}, \tag{2}$$

where R_{oo} can be used to replace R in equation (1) in calculating the increment of inclusive fitness for an O allele through an OO genotype helping its offspring.

Equation (2) makes two points: first, helping an offspring is better than helping a randomly chosen individual because $p_{L.oo} > p$, and, second, because R_{oo} is the same as R in equation (1), its dependence on p implies that the evolutionary benefit of helping depends on the frequency of the helping gene p in the population. In the case of an OO individual helping its offspring, its fitness increment is a linear function of p with an intercept of $(0.5B - C)$ and a slope of $(0.5B)$. Evidently, as long as $B \geq 2C$, the net benefit for an OO genotype from helping its offspring (Z_{oo}) will always be greater than zero.

Allelic frequencies of O in offspring of an OS individual are

$$p_{L.os} = 0.25 + \frac{p}{2} = R_{os1}, \tag{3}$$

where R_{os1} can be used to replace R in equation (1) in calculating the increment of inclusive fitness for an O allele through an OS genotype helping its offspring.

We now calculate the allelic frequency of S among full siblings of an SS genotype (represented by $q_{L.ss}$). The parents of an SS genotype must be S? and S?, where "?" is unknown but has a probability q of being S and a probability p of being O. Given these conditions, the allelic frequency of S among offspring from
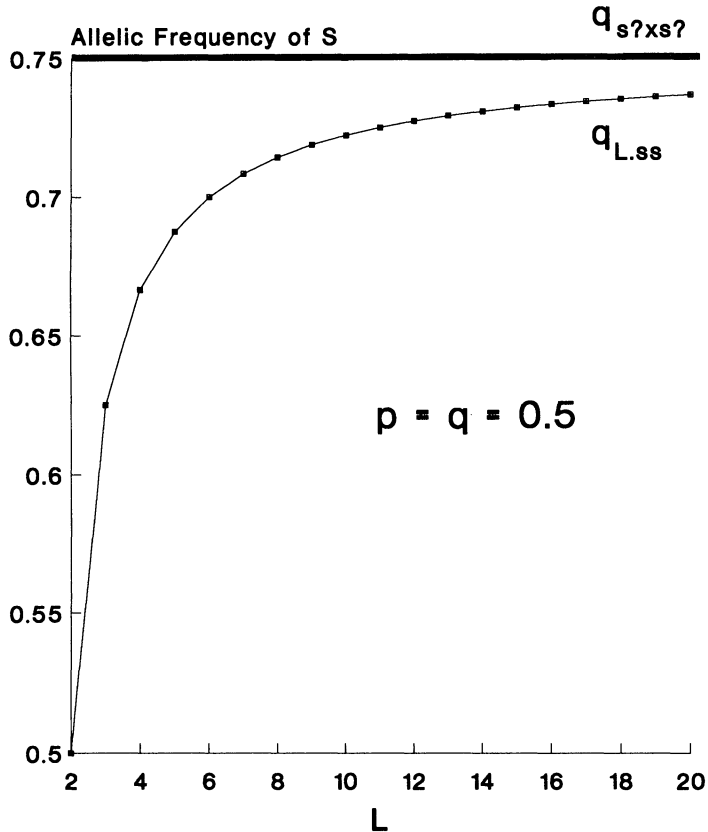
FIG. 1.—The allelic frequency of S in full siblings of an SS individual ($q_{L.ss}$). Note that it increases asymptotically toward $q_{S?\times S?}$ when L approaches infinity.

an S? × S? mating is

$$q_{S?\times S?} = 0.5 + 0.5q,\tag{4}$$

which is identical in form, as expected, to equation (1).

But $q_{S?\times S?}$ is not $q_{L.ss}$, which is the allelic frequency of S among full siblings of an SS individual, calculated after excluding this SS individual. The two are equal only when L, the number of offspring per family, is infinite. When L is finite,

$$q_{L.ss} = \frac{Lq_{S?\times S?} - 1}{L - 1} = \frac{L + Lq - 2}{2(L - 1)} = R_{ss},\tag{5}$$

where R_{ss} can be used to replace R in equation (1) in calculating the increment of inclusive fitness for an S allele through an SS genotype helping its full siblings.

Evidently, when L is small, R_{ss} is substantially smaller than $q_{S?\times S?}$ (fig. 1), the latter being identical to R_{oo} in form. Therefore, R_{ss} for any q is smaller than R_{oo} for $p = q$. This suggests that an O allele derives greater fitness gain through an

OO individual helping its offspring than an S allele would through an SS individual helping its full siblings.

The allelic frequencies of S in full siblings of an OS individual are

$$q_{\text{L.os}} = \frac{L + 2Lq - 2}{4(L - 1)} = R_{\text{os2}}, \tag{6}$$

where R_{os2} can be used to replace R in equation (1) in calculating the increment of inclusive fitness for an S allele through an OS genotype helping its full siblings.

The part of inclusive fitness of an O allele due to helping, according to Hamilton (1964, 1972) and Grafen (1991), can be calculated as follows:

$$W_{\text{o}} = \frac{P_{\text{oo}}\left[\left(0.5 + \dfrac{p}{2}\right)B - C\right] + \dfrac{P_{\text{os}}\left[\left(0.25 + \dfrac{p}{2}\right)B - C\right]}{2}}{P_{\text{oo}} + \dfrac{P_{\text{os}}}{2}}. \tag{7}$$

Similarly, the part of inclusive fitness of S allele due to helping can be written as

$$W_{\text{s}} = \frac{P_{\text{ss}}\left\{\left[\dfrac{L + Lq - 2}{2(L - 1)}\right]B - C\right\} + \dfrac{P_{\text{os}}\left\{\left[\dfrac{L + 2Lq - 2}{4(L - 1)}\right]B - C\right\}}{2}}{P_{\text{ss}} + \dfrac{P_{\text{os}}}{2}}. \tag{8}$$

If the selection on sibling groups and parent-offspring groups is not strong, then we can approximate P_{oo}, P_{os}, and P_{ss} by p^2, $2pq$, and q^2. Now we have

$$W_{\text{o}} = \frac{3Bp}{4} + \frac{B}{4} - C \tag{9}$$

and

$$W_{\text{s}} = \frac{4BL - 3BLp - 4B - 4CL + 4C + 2Bp}{4(L - 1)}. \tag{10}$$

If L is infinite, equation (10) becomes

$$W_{\text{s}} = \frac{3Bq}{4} + \frac{B}{4} - C, \tag{11}$$

which is identical in form to equation (9), as one would expect.

We can see from equations (9) and (10) that both W_{o} and W_{s} depend on B, C, and p, but W_{s} is additionally dependent on L, the effect of which is shown in figure 2. When the frequency of a certain allele is almost one, then virtually all the help provided by individuals carrying the allele is used to benefit the replication of the allele, and little assistance is wasted to help increase the alternative allele. Thus, the maximum benefit for O and S alleles is attained when p and q, respectively, reach one (fig. 2). When the frequency of a certain allele, say S, decreases, some help provided by individuals carrying S may be wasted in helping individu-
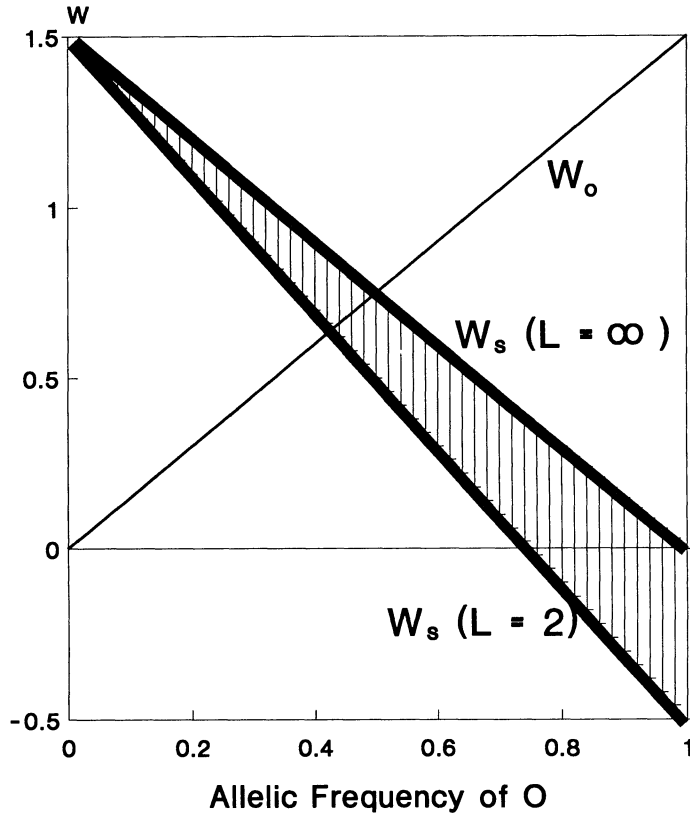
FIG. 2.—Increment of fitness through helping full siblings (W_s) and through helping off-spring (W_o) for $L = 2$ and $L = \infty$, respectively. The variables B and C are set to two and one, respectively. The area between the W_s lines for $L = 2$ and $L = \infty$ is the deterministic cost of helping full siblings instead of helping offspring for species with $L = 2$.

als carrying O alleles, and this benefit to O alleles may not be compensated for by the help received by S-carrying individuals from O-carrying individuals. When the frequency of O allele is almost one, and if L = 2, then equations (5) and (6) yield a $q_{L.ss}$ and a $q_{L.os}$ close to zero. Thus, virtually all help provided by S-carrying individual is wasted in helping O-carrying siblings. On the other hand, an O-carrying individual will have an R no smaller than 0.25, as one can see from equations (2) and (3).

In short, S is fitter than O when p is less than p^*, the p value at which lines W_o and W_s cross each other (fig. 2), but less fit than O when p is greater than p^*. Figure 2 shows that p^* is always smaller than 0.5 when L is finite and that W_o − W_s for any given p is greater than W_s − W_o for $q = p$. The area between the lines for $L = \infty$ and $L = 2$ is the cost of helping full siblings relative to helping offspring when $L = 2$. This leads us to predict that species in which sibling helping is prevalent should have a relatively large L. Honey bees, termites, and African mole rats are examples consistent with this prediction.

STOCHASTIC COST

In equations (9) and (10), W_o and W_s are expected gains of inclusive fitness of O and S through helping offspring and siblings, respectively. These expected fitness gains have their respective variances, which also affect the reproductive performance of O- and S-carrying individuals over generations (Gillespie 1972, 1973a, 1973b, 1974, 1975, 1977). We now examine whether W_o and W_s differ in variance and whether the difference, if present, is evolutionarily significant in its effect on long-term performance of O- and S-carrying individuals.

Let us first rewrite equations (9) and (11) in the following form:

$$W_o = BpR_{oo} + BqR_{os1} - C \tag{12}$$

and

$$W_s = BqR_{ss} + BpR_{os2} - C. \tag{13}$$

Because B and C are constants and because the variance of p (or q) due to sampling error is at least $2N$ times smaller than p (or q), we can write the variance of W_o and W_s as follows:

$$\sigma^2_{W_o} = B(p\sigma^2_{R_{oo}} + q\sigma^2_{R_{os1}}) \tag{14}$$

and

$$\sigma^2_{W_s} = B(q\sigma^2_{R_{ss}} + p\sigma^2_{R_{os2}}). \tag{15}$$

Note that I have ignored the covariance term because R_{oo} and R_{os1} represent independent sampling points and should not co-vary given p. Similarly, we should not expect R_{ss} and R_{os2} to co-vary given p.

The variance term of R_{oo}, R_{os1}, R_{os2}, and R_{ss} in equations (14) and (15) can be written as

$$\sigma^2_{R_{oo}} = \frac{q}{4LNp}, \tag{16}$$

$$\sigma^2_{R_{os1}} = \frac{4pq + 1}{32LNpq}, \tag{17}$$

$$\sigma^2_{R_{os2}} = \frac{L(L + 2Lpq - 2 + 2pq)}{32(L - 1)^3 Npq}, \tag{18}$$

and

$$\sigma^2_{R_{ss}} = \frac{Lp(2L - Lp - 2 - p)}{8(L - 1)^3 Nq^2}, \tag{19}$$

where N is population size (number of individuals providing help), L is the number of offspring each individual has, and $(L - 1)$ is the number of full siblings each individual has. These variance terms are derived in the following way. First, the variance of R within each offspring group or full sibling group is calculated for each of the three genotypes OO, OS, and SS. For example, the variance of R for an OO individual is $pq/(4L)$. These within-group variance terms of R are then divided by the number of respective genotypes. For example, $pq/(4L)$ is
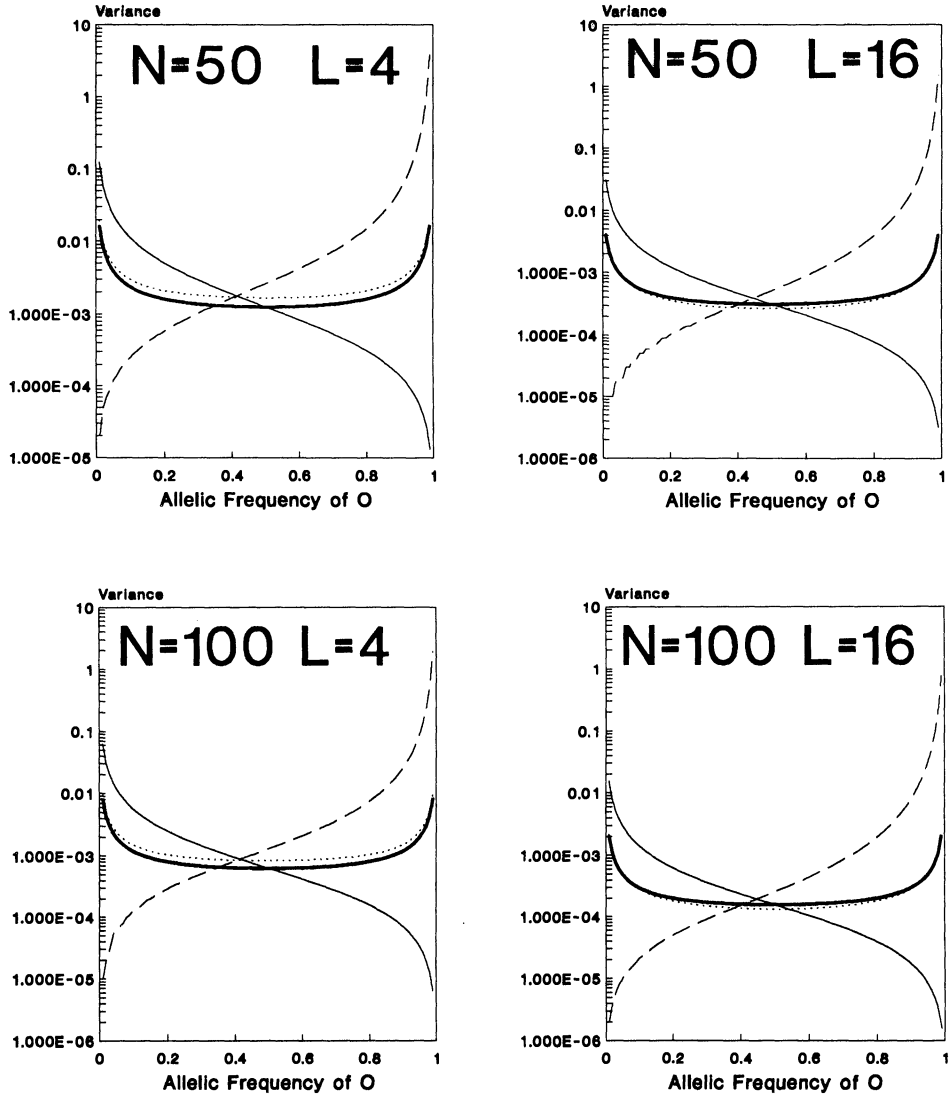
FIG. 3.—The variance of R terms in relation to allelic frequency of O for different N and L values. *Solid thin line,* R_{oo}; *dashed line,* R_{ss}; *solid thick line,* R_{os1}; *dotted line,* R_{os2}.

divided by the number of OO individuals, which is again approximated by Np^2, to obtain the variance of R for all OO individuals in the population for one generation. The resulting four variance terms are graphed in figure 3 for two different N values (50 and 100, respectively) and two different L values (four and 16, respectively).

That the variance of R_{oo} should in general be smaller than that of R_{os} (fig. 3) is intuitively obvious because an SS individual can have siblings with all three genotypes, one of which (OO) does not have a single copy of S, whereas an OO individual can have only two genotypes (OO and OS) in its offspring, with each
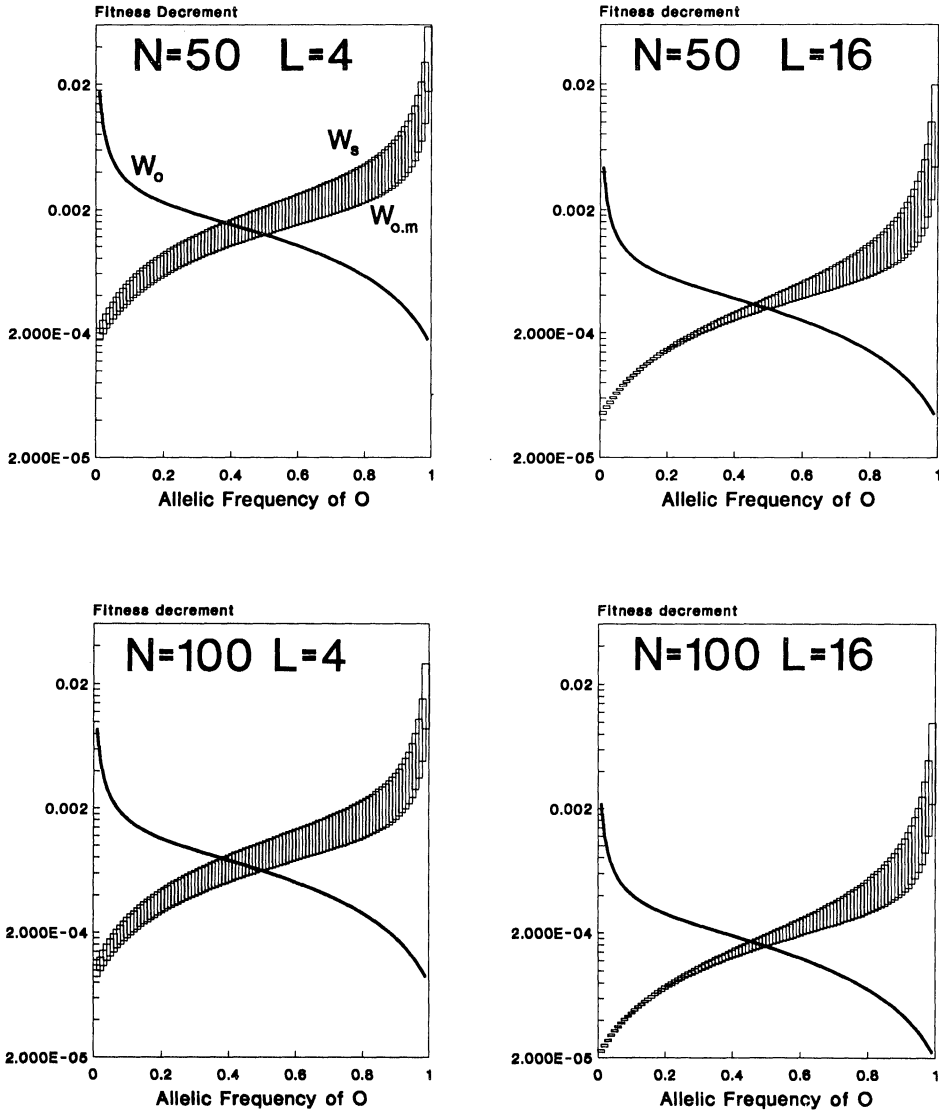
FIG. 4.—Decrement of fitness due to variation in W_o and W_s in relation to allelic frequency of O for different N and L values. The line $W_{o.m}$ is a mirror image of W_o. The area between lines W_s and $W_{o.m}$ is the stochastic cost of helping full siblings instead of helping offspring. The variable B is set to two.

genotype having at least one O allele. In addition, the within-group variance R_{ss} is divided by $(L - 1)$ whereas that of R_{oo} is divided by L.

Given the variance terms of R_{oo}, R_{os1}, R_{os2}, and R_{ss} expressed in equations (16)–(19), the variance terms of W_o and W_s in equations (14) and (15) are the variances of W_o and W_s over generations with N helpers in each generation. The fitness decrement due to variation of fitness is approximately half of the variance (Gillespie 1977) and is graphed against different p values in figure 4 with $B = 2$

and two different L values (four and 16, respectively) and two different N values (50 and 100, respectively). Note that population size N, which does not play a role in determining the expected value of W_o and W_s in the previous section, now becomes important in determining fitness decrement due to variance of W_o and W_s over generations. When N is infinitely large, both variance of W_o and variance of W_s become insubstantial in terms of affecting fitness, although one may still be many times greater than the other.

Qualitatively, the effect of variance of W_o and W_s is similar to the effect of W_o and W_s (figs. 2 and 4) . When p is less than $p*'$, which is the p value at which lines W_o and W_s cross each other (fig. 4), the decrement of fitness for S allele due to variance of W_s is less than that for O due to variance of W_o. The opposite is true when p is greater than $p*'$. The area between lines W_s and $W_{o.m}$, which is a mirror image of line W_o, is the stochastic cost of helping siblings instead of helping offspring (fig. 4). The area between the lines for W_o and W_s in figure 4 for $p > p*$ is the fitness differential that the S allele must overcome before selection for reduced fitness variation acts in its favor. Figure 4 also shows that the difference between the O and S alleles decreases as L increases, which again leads us to predict that species in which sibling helping is prevalent should have a relatively large L.

In summary, the expected value of W_o for any p is greater than that of W_s for $q = p$ when L is finite. In addition, the fitness decrement due to variance of W_o for any p is less than the fitness decrement of S due to variance of W_s for $q = p$. Both results imply that helping offspring is selectively more advantageous than helping full siblings. I conclude that trading one's whole reproductive life for two brothers or giving up one's son for a baby brother is unlikely to be favored by natural selection.

DISCUSSION

My results above suggest that, if we are given n populations all with an allelic frequency of O equal to 0.5, we should end up with more populations fixed for O than populations fixed for S, especially when L and N are small. In reality, O must have evolved first because there would be no life if parents did not invest in their offspring. Thus, the evolutionary question of interest is whether it is likely for an S allele to invade a population fixed for O. My results show that such an invasion is very unlikely, not only because S suffers from both a deterministic cost and a stochastic cost but also because, once the O or S allele reaches a high frequency in a population, it is very difficult for the other allele to increase (figs. 2 and 4).

From the results that both the deterministic and the stochastic costs of helping siblings decrease with increased L, we can predict that sibling-helping behavior should be more prevalent in species with a large L. This prediction accords well with nature as species known to exhibit much altruism among siblings are those with a large L (e.g., eusocial hymenopterans, termites, African mole rats, mongooses).

If the altruistic behavior is very effective (large B), if the number of siblings per individual ($L - 1$) is few, and if population size is small, then altruistic

behavior toward full siblings should be very unlikely unless one does not have the option of helping one's own offspring (e.g., when one is reproductively inhibited by social and environmental pressure). In populations in which each individual has a lengthy prereproductive life (e.g., human population), the option of helping offspring is not available during this prereproductive period. Animals during this period may choose between helping siblings and not helping. On the other hand, if the age at which an individual becomes capable of helping its siblings is also the age at which it becomes capable of reproduction, then the animal can choose between helping siblings and helping offspring.

The original models of kin selection by Hamilton (1964, 1972) allow the option of helping one's siblings as well as the option of helping one's offspring and are perhaps the most realistic among all models of kin selection. Subsequent models of altruism among siblings (see, e.g., Charlesworth 1978) usually compare fitness of sibling helpers with that of nonhelpers. These models are appropriate only in situations in which the option of helping one's own offspring is not available.

What I find most uncomfortable are the models of group selection (e.g., Eshel 1972; Wilson 1975, 1977, 1980). All of these models compare the fitness of helpers with nonhelpers and are appropriate only when the option of helping one's relatives is not available. If the fitness of those blind helpers (altruists) in those models is compared to the fitness of individuals (selfish ones) that direct their helping behavior toward their relatives, then these models all break down. It seems to me that, although the option of helping one's offspring may sometimes be unavailable, the option of helping one's relatives (including relatives besides offspring and parents) is readily available in all biological realities. Thus, those works of group selection appear to model a situation that is biologically very unrealistic. We should keep in mind that the trait of eating unhealthy food could be selectively advantageous if the trait is pitted against a trait the carrier of which does not eat at all. If the option of eating healthy food is readily available and if we have an alternative trait of eating healthy food in the population, then the trait of eating unhealthy food becomes immediately disadvantageous.

The last point I wish to make is that the essence of the model is the negative correlation between the tendency to help one's offspring and the tendency to help one's full siblings. These two alternative strategies could be quantitative traits and controlled by multiple loci instead of just a single locus. If the contribution of these multiple loci to the quantitative traits is additive, then the negative correlation between the tendency to help offspring and the tendency to help full siblings is maintained. The result of selection would be the same as the single-locus model I presented.

Now imagine that we have the following situation. One locus is already fixed for O and a second locus is fixed for a neutral allele that does not have any effect on anything. If a mutation S arises in this second locus that predisposes its carrier to allocate a certain portion of resources to help its full siblings, will this S have a selective advantage over its functionless alternative allele? The answer may not be intuitively obvious. If we consider this locus in isolation and if the organism under consideration has surplus resources that cannot be invested in offspring, then we can easily come up with some selection scenarios that favor S over its alternative. On the other hand, if helping full siblings necessitates channeling

away resources that would otherwise have been invested in offspring, then a negative correlation between the tendency to help offspring and the tendency to help full siblings is restored. Now if the ratio, B/C, for S were not greater than B/C for O (i.e., if S were not more efficient in helping than O), then we would find the S allele selectively worse than its functionless alternative.

What would happen if we had perfectly linked O-S and its alternative o-s, whose carrier provides no help to offspring and siblings? Now the tendencies to help one's offspring and to help one's full siblings are positively correlated. That is, helping one's offspring and helping one's full siblings are now no longer the two alternative strategies. Instead, we have one genotype that helps both its offspring and full siblings and another genotype that helps neither. This is a special selection regime of helpers (or altruists) versus nonhelpers (or nonaltruists). Two points should be made clear. First, a scenario that favors O-S over o-s does not support group selection but only suggests the plausibility of kin selection. Second, as soon as mutation gives rise to O-s, we would find it favored by selection over O-S.

In summary, the value of a son, in an evolutionary sense, is greater than, or at least equal to, the value of a full brother, being equal when L and N are infinite. If a female is made to choose between an orphaned baby brother and a son of equal age, we can predict that she is more likely to choose the son. This is because, given the small L in human populations, the probability of the female's being of OO genotype is much greater than the probability of the female's being of SS genotype. Interestingly, Dawkins and Carlise (1976) noted that intuition would suggest the choice of the son. It seems that intuition may sometimes lead to more accurate predictions than the vigorous reasoning of ethologists equipped with an insightful but often overstretched evolutionary theory.

LITERATURE CITED

Charlesworth, B. 1978. Some models of the evolution of altruistic behaviour between siblings. Journal of Theoretical Biology 72:297–319.
Dawkins, R., and T. R. Carlisle. 1976. Parental investment, mate desertion and a fallacy. Nature (London) 262:131–132.
Eshel, I. 1972. On the neighbor effect and the evolution of altruistic traits. Theoretical Population Biology 3:258–277.
Gillespie, J. H. 1972. The effects of stochastic environments on allele frequencies in natural populations. Theoretical Population Biology 3:241–248.

———. 1973*a*. Natural selection with varying selection coefficient—a haploid model. Genetical Research 21:115–120.

———. 1973*b*. Polymorphism in random environments. Theoretical Population Biology 4:193–195.

———. 1974. Natural selection for within-generation variance in offspring number. Genetics 76:601–606.

———. 1975. Natural selection for within-generation variance in offspring number. II. Discrete haploid model. Genetics 81:403–413.

———. 1977. Natural selection for variance in offspring numbers. American Naturalist 111:1010–1014.

Grafen, A. 1991. Modelling in behavioral ecology. Pages 5–31 *in* J. R. Krebs and N. B. Davies, eds. Behavioral ecology: an evolutionary approach. Blackwell Scientific, Oxford.

Hamilton, W. D. 1964. The genetical theory of social behaviour. I, II. Journal of Theoretical Biology 7:1–52.

———. 1972. Altruism and related phenomena, mainly in the social insects. Annual Review of Ecology and Systematics 3:193–232.

Trivers, R. L. 1972. Parental investment and sexual selection. Pages 136–179 *in* B. Campbell, ed. Sexual selection and the descent of man, 1871–1971. Aldine, Chicago.

Wilson, D. S. 1975. A theory of group selection. Proceedings of the National Academy of Sciences of the USA 72:143–146.

———. 1977. Structured demes and the evolution of group-advantageous traits. American Naturalist 111:157–185.

———. 1980. The natural selection of populations and communities. Benjamin/Cummings, Menlo Park, Calif.

XUHUA XIA

DIVISION OF LIFE SCIENCES
  SCARBOROUGH CAMPUS
    UNIVERSITY OF TORONTO
      SCARBOROUGH, ONTARIO, M1C 1A4
        CANADA