



**Cite this article:** Rogers AR. 2013 Genetic relatedness to sisters' children has been underestimated. *Proc R Soc B* 280: 20121937. <http://dx.doi.org/10.1098/rspb.2012.1937>

Received: 18 August 2012

Accepted: 5 November 2012

**Subject Areas:**

behaviour, evolution, genetics

**Keywords:**

avunculate, genetic relatedness, paternity threshold

**Author for correspondence:**

Alan R. Rogers

e-mail: [rogers@anthro.utah.edu](mailto:rogers@anthro.utah.edu)

# Genetic relatedness to sisters' children has been underestimated

Alan R. Rogers

Department of Anthropology, University of Utah, 270 S 1400 E, Salt Lake City, UT 84112, USA

Males of many species help in the care and provisioning of offspring, and these investments often correlate with genetic relatedness. For example, many human males invest in the children of sisters, and this is especially so where men are less likely to share genes with children of wives. Although this makes qualitative sense, it has been difficult to support quantitatively. The prevailing model predicts investment in children of sisters only when paternity confidence falls below 0.268. This value is often seen as too low to be credible; so investment in sisters' children represents an unsolved problem. I show here that the prevailing model rests on a series of restrictive assumptions that underestimate relatedness to sisters' children. For this reason, it understates the fitness payoff to men who invest in these children. This effect can be substantial, especially in societies with low confidence in paternity. But this effect cannot be estimated solely from confidence in paternity. One must also estimate the probability that two siblings share the same father.

## 1. Introduction

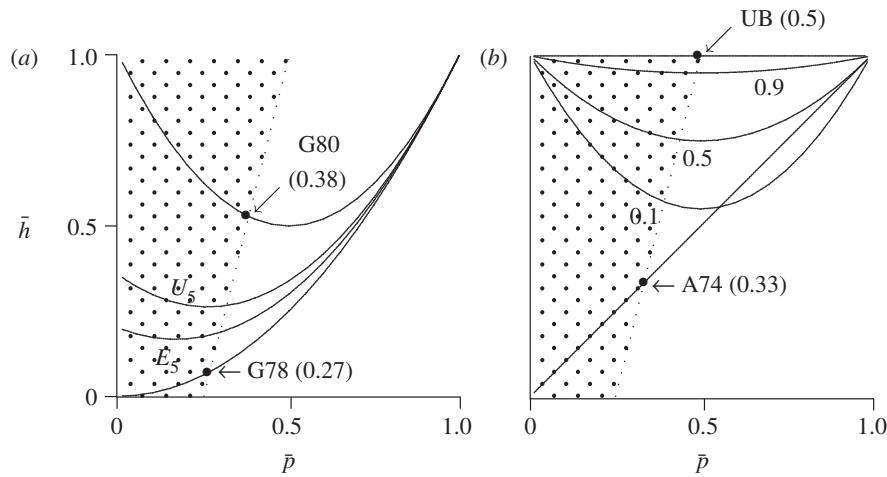
In species with male parental care, it makes sense that males would direct that care towards offspring with whom they share genes. Male dunnocks, for example, often provision the young of several females, and the rate at which they provision each clutch closely matches their likely share of its paternity [1]. Similarly, many human males invest in the children of sisters rather than in those of wives, and this practice may be most widespread where extramarital mating is common [2–6] (but see [7–9]).

This human example was first approached in a quantitative way during the 1970s. In that decade, quantitative models were introduced by Alexander [10] and Greene [11,12]. I will refer to these as models A74, G78 and G80, respectively. Their common goal was to specify conditions under which men are genetically closer to sisters' children than to those of wives. The point at which they are equally close is called the 'paternity threshold' [11]. These authors assumed that selection would favour investment in sisters' children only when paternity confidence is below this threshold. The three models disagreed about its numerical value. The lowest value—0.268—was that of the G78 model. This model has become enshrined in the literature.

The others have been less influential. The A74 model was used in two publications [4,13] and the G80 model (without attribution) in one [14], but neither has been used since the 1980s.

The G78 model, on the other hand, continues to influence thinking. Anthropologists have extended it to various types of relative [2,15] and across several generations [5,6]. It also shows up in economic literature ([14]; [16], pp. 1923–1924). It is used in recent literature to argue that transfers to sisters' children compromise the reproductive interests of husbands ([17], p. 109; [18], p. 157).

It is not clear, however, that the paternity threshold provides a useful way to think about paternal investment. In the first place, it is hard to believe that paternity certainty is often as low as the model requires [4,5]. Furthermore, many forms of parental investment may exhibit decreasing returns to scale. In other words, the benefit from an additional unit of investment in any given offspring may decline with each unit invested. Where this is so, we might expect men to distribute



**Figure 1.** Effect of average paternity confidence ( $\bar{p}$ ) on the average probability ( $\bar{h}$ ) that two sibs share a father. In the shaded regions, men share more genes with the child of a sister than with that of a wife. Paternity thresholds of several models are indicated by filled circles. For these thresholds, the value of  $\bar{p}$  is shown in parentheses. (a) Paternity confidence constant across families. Curves G78 and G80 represent the models of Greene [11,12]. The other two curves assume that women have five extrapair mates, with whom the frequency of mating is either even ( $E_5$ ) or uneven ( $U_5$ ). (b) Paternity confidence varies among families. Curve A74 shows the Alexander model. The other curves assume that  $b = 1$  and that, for each family,  $p$  is drawn from a Beta distribution with mean  $\bar{p}$ . Labels show the variance of this distribution as a fraction of the maximum possible variance,  $\bar{p}(1 - \bar{p})$ . Curve UB shows the upper bound, at which the paternity threshold reaches its maximal value, 0.5.

investment among multiple offspring, weighting that investment in favour of those with whom they are most likely to share genes. Furthermore, the optimal behaviour of males will depend in part on the responses of females. Neither of these complexities is accommodated by the paternity threshold model.

They are central, however, to recent game theoretic analyses [19,20]. Those show that selection can favour investment in sisters' children even when paternity confidence is well above the paternity threshold. Yet, these models also use the G78 model in calculating fitness payoffs ([19], supplementary material, p. 8; [20]). That model thus continues in importance.

In what follows, I argue that all three models (A74, G78 and G80) make restrictive assumptions. Some of these have been acknowledged [4,11,12], but others seem not to have been noticed. All of them distort our views about the relatedness of men to sisters. These distortions may have biased the conclusions of all the work mentioned earlier.

This article will not try to explain investment by males in sisters' children. It will deal only with questions of relatedness, a more limited project that may prove useful in the larger one.

## 2. Material and methods

Most quantitative work on this subject has used the *coefficient of relatedness*. I work instead with the *coefficient of kinship* ([21], p. 121)—the probability that two genes, drawn at random from each of two relatives, are copies of the same gene in some given ancestral generation. In the absence of inbreeding, the coefficient of kinship equals half the coefficient of relatedness. We are interested in two of these coefficients, which describe the genetic kinship of a man to the children of (i) his wife and (ii) his sister.

These coefficients depend on two probabilities, which may vary among families. The first of these— $p$ , or paternity confidence—is the probability that a child's father is his mother's husband. The other is  $h$ , the probability that two siblings share the same father. The value of  $h$  is relevant because it affects the genetic relatedness of a man (Ego) to his sister's children. This value depends on paternity confidence, but not that provided by Ego's wife. It depends instead on the paternity confidence provided by Ego's mother to her husband. The value of  $h$  also

depends on the number of a woman's husbands [20]. I assume throughout that women have only one. Finally, I assume that Ego does not know values specific to his own family and must rely instead on population averages,  $\bar{p}$  and  $\bar{h}$ .

Within a family, the coefficient of kinship of Ego with his wife's child is  $\phi_1 = p/4$ , whereas that with his sister's child is  $\phi_2 = (1 + h)/16$ . A man shares more genes with the child of a sister than with that of a wife when  $\phi_2 > \phi_1$ , or equivalently when  $h > 4p - 1$ . This condition refers to an individual family, but also holds on average if  $p$  and  $h$  are replaced by  $\bar{p}$  and  $\bar{h}$ . The shaded regions in figure 1 show the combinations of  $\bar{h}$  and  $\bar{p}$  that satisfy this inequality.

Consider a family within which the probability of paternity is  $p$ . With probability  $p^2$ , two random siblings were fathered by the husband, and with probability  $(1 - p)^2$  neither was. In the second case, both may have been fathered by the same extrapair<sup>1</sup> male. Let  $b$  represent the conditional probability of this event, given that neither sib was fathered by the husband. With these definitions,

$$h = p^2 + (1 - p)^2 b. \quad (2.1)$$

For an average family,

$$\begin{aligned} \bar{h} &= \overline{p^2} + \overline{(1 - p)^2 b} \\ &= \overline{p^2} + V_p + \overline{(1 - p)^2 b} \end{aligned} \quad (2.2)$$

where the overbars represent averaging over families, and  $V_p = \overline{p^2} - \bar{p}^2$  is the variance among families in  $p$ .

## 3. Results

This section will (a) derive the assumptions that underlie the three published models, and then relax assumptions involving (b) the distribution of sexual access among extrapair mates and (c) the variation of paternity confidence among families.

### (a) Assumptions of the classical models

The three models make differing claims about  $\bar{h}$ :

$$A74: \bar{h} = \bar{p}, \quad (3.1)$$

$$G78 : \bar{h} = \bar{p}^2 \tag{3.2}$$

and

$$G80 : \bar{h} = \bar{p}^2 + (1 - \bar{p})^2. \tag{3.3}$$

Each of these results can be derived by setting  $\bar{b}$  and  $V_p$  to values at the edges of their feasible ranges—that is, by setting each parameter either to its highest or its lowest feasible value.

Because  $b$  is a probability, it must lie between 0 and 1, and so must its average,  $\bar{b}$ . This average equals 0 when women never mate with the same extrapair male twice, or (equivalently) when each woman has an infinite number of such mates. On the other hand,  $\bar{b} = 1$  when no woman has more than a single extrapair mate. The variance,  $V_p$ , must lie between 0 and  $\bar{p}(1 - \bar{p})$ . It attains the lower value when all women have the same fraction of extrapair matings. The upper value occurs when  $p = 1$  for a fraction  $\bar{p}$  of women (who always mate with their husbands), and  $p = 0$  for the rest (who never do). In summary,  $\bar{b}$  lies within the range  $[0, 1]$  and  $V_p$  within  $[0, \bar{p}(1 - \bar{p})]$ .

Consider what happens when we substitute these upper and lower feasible values into equation (2.2). The result equals equation (3.1) when  $\bar{b} = 0$  and  $V_p = \bar{p}(1 - \bar{p})$ , equals equation (3.2) when  $\bar{b} = V_p = 0$ , and equals equation (3.3) when  $\bar{b} = 1$  and  $V_p = 0$ . Thus, models A74, G78 and G80 can each be derived by setting  $\bar{b}$  and  $V_p$  equal to values at the limits of their feasible ranges. The A74 model also holds if  $p \leq \frac{1}{2}$  and  $b = p/(1 - p)$  within each family, for then equation (2.1) reduces to  $h = p$ , and averaging over families gives  $\bar{h} = \bar{p}$  (equation (3.1)).

These results are summarized in table 1. For the G78 model, the assumptions discussed above are necessary as well as sufficient. Those for the A74 and G80 models are sufficient, but may not be necessary: they may hold also under other assumptions. None, however, have ever been described. The assumptions discussed earlier are the only ones under which these models are known to hold. Let us ask now what happens when these assumptions are relaxed. §3b considers the possibility that siblings share paternity through an extrapair male.

### (b) Extrapair paternity

This section focuses on the probability,  $b$ , that two siblings share a father, given that neither was fathered by their mother’s husband. It will be useful to simplify the other parts of the model; so let us follow Greene [11,12] in assuming that  $p$  and  $b$  are constant across families. In this context, there is no distinction between values for families and averages across families; so I omit the overbars.

It is easy to derive the minimum and maximum feasible values of  $h$ . Equation (2.1) implies that  $h$  increases with  $b$ , for any given  $p$ . Consequently, the minimum  $h$  occurs when  $b = 0$  and the maximum when  $b = 1$ . These correspond to curves G78 and G80 in figure 1a. These curves provide lower and upper bounds on the value of  $h$ , provided that  $p$  and  $b$  are constant across families. When families vary, however, we will see below that  $\bar{h}$  can be even larger than G80 would imply.

Between the two extremes,  $b = 0$  and  $b = 1$ , the value of  $b$  will vary in response to the number of extrapair partners and the distribution of matings among them. To model this effect, let  $g_i$  represent the fraction (among all extrapair matings of a

**Table 1.** Assumptions that underlie each model. The models differ with respect to assumptions about two parameters:  $V_p$  (the variance across families in paternity confidence) and  $\bar{b}$  (the mean probability that two siblings share an extrapair father, if neither was fathered by the mother’s husband). The first of these ranges from 0 to  $\bar{p}(1 - \bar{p})$  and the second from 0 to 1. Each model can be derived by setting each parameter either to its highest or its lowest feasible value. The A74 model can also be derived from assumptions that imply intermediate values of the two parameters. ‘Upper bound’ (UB) is the model with the highest possible paternity threshold.

model	abbreviation	assumption	
		$V_p$	$\bar{b}$
Greene [11]	G78	0	0
Greene [12]	G80	0	1
Alexander [10]	A74	$\bar{p}(1 - \bar{p})$	0
upper bound	UB	$\bar{p}(1 - \bar{p})$	1

given woman) of matings with the  $i$ th male. Then  $g_i^2$  is the probability that two sibs were both fathered by this male, if neither was fathered by the husband. In these terms,

$$b = \sum_{i=1}^k g_i^2 = \frac{1}{k} + kV_g, \tag{3.4}$$

where  $k$  is the number of males,  $1/k$  is the mean of the  $g_i$  and  $V_g = k^{-1} \sum_i g_i^2 - 1/k^2$  is the variance.

If the wife visits extrapair partners with equal frequency, then  $b = 1/k$ , and  $h = p^2 + (1 - p)^2/k$ . This reduces to model G80 if the wife has just a single extrapair partner, and to model G78 if she has an infinity of them. In addition to these curves, figure 1a also includes model  $E_5$ , in which the wife allocates matings evenly among five extrapair partners. For a given value of  $p$ , the figure shows that  $h$  declines as sexual partners become more numerous, because offspring are then less likely to share paternity through an extrapair male.

When extrapair males receive uneven allocations,  $V_g$  is large, increasing both  $b$  and  $h$ , and making siblings more similar. This makes intuitive sense: variation in  $g_i$  implies that a small number of males enjoy disproportionate mating success; so random pairs of offspring are likely to share paternity through one of these favoured males. To illustrate this effect, figure 1a includes curve  $U_5$ , representing the case of five extrapair males who get sexual access in proportion to  $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}$  and  $\frac{1}{16}$ . Because of this unevenness, curve  $U_5$  is higher than  $E_5$ .

As each of the curves in figure 1a passes from the shaded to the unshaded region,  $\bar{p}$  passes what Greene ([11], p. 153) called the ‘paternity threshold’—the ‘ $p$  below which a male is more related to his sister’s offspring than to his spouse’s’. Because of the slope of the boundary, the smallest and largest thresholds are those for the lowest and highest curves on the page. These two extremes—models G78 and G80—imply thresholds of 0.268 [11] and 0.382 [12].

So far, we have seen that relatedness to sisters (and thus to sisters’ offspring) increases if the mother has a small number of extrapair partners or allocates matings unevenly among them. Furthermore, the figure shows that these effects can be quite large. §3c considers another influence—variation among families in paternity confidence.

### (c) Variation in paternity confidence

The fraction of children fathered by the current husband will ordinarily vary among families. In such cases,  $p$  and  $h$  are random variables, and attention turns to their averages,  $\bar{p}$  and  $\bar{h}$ . To obtain an upper bound on  $\bar{h}$ , assume that  $b = 1$  for all women. Then equation (2.2) becomes

$$\bar{h} = 1 - 2\bar{p}(1 - \bar{p}) + 2V_p.$$

When the variance ( $V_p$ ) among families is large,  $\bar{h}$  will also be large. To illustrate this effect graphically, I assume that  $p$  is drawn randomly for each family from a beta distribution with mean  $\bar{p}$ . In figure 1b, the curves labelled 0.1, 0.5 and 0.9 refer to models with increasing variance among families. The larger the variance, the higher the probability ( $\bar{h}$ ) that two siblings share a father.

The maximal value of  $\bar{h}$  occurs when  $b = 1$  for all women, and  $V_p$  is at its maximal value,  $\bar{p}(1 - \bar{p})$ . In this case,  $\bar{h} = 1$ , whatever the value of  $\bar{p}$ . The paternity threshold attains its largest possible value,  $\frac{1}{2}$ . This model is shown as curve UB in figure 1b.

## 4. Discussion and conclusions

It has been understood for four decades that paternity certainty increases the relatedness of men to the children of wives and sisters [10]. But relatedness to sisters' children also responds to other influences, which have not been appreciated. It is greater when women have few extrapair partners and allocate matings unevenly among them, and when paternity confidence varies among families. Published models involve restrictive assumptions about all these influences.

These assumptions underlie an old debate about the correct form of the relationship between paternity certainty and relatedness to sisters' children. Various authors disagreed about which functional form was correct ([11], p. 153; [2], pp. 151–152; [13], p. 321; [4], pp. 443–444). The present work shows that all are correct—they simply involve different assumptions.

When spelled out, these assumptions seem remarkably restrictive. They include: (i) that paternity confidence is the same in each family (G78, G80); (ii) that women never mate twice with the same extrapair male (A74, G78); (iii) that some women always mate with their husbands, but the rest never do (A74); and (iv) that no woman mates with more than one extrapair male (G80). Each of these assumptions is unrealistic, and each biases estimates of genetic relatedness. The assumptions (i) and (ii) of the G78 model both bias results downward. Consequently, this model provides only a lower bound on the relatedness of men to sisters' children. Current theory relies on this model, and thus underestimates the fitness

payoff to males who invest in such children. The upper bound on relatedness occurs when assumptions (iii) and (iv) both hold. At this upper bound, all siblings share paternity, because no sibship has more than one biological father.

Presumably, real populations lie somewhere between these extremes. This range of uncertainty implies that there is no single paternity threshold—no single value of  $\bar{p}$  at which men are equally related to children of wives and of sisters. Instead, the paternity threshold varies among populations between 0.268 (the lower bound) and 0.5 (the upper). Even under the most generous conditions, the paternity threshold requires a very low confidence in paternity.

The present results also bear on recent game-theoretic analyses. Fortunato & Archetti [19] studied the evolution of monogamous marriage and 'vertical transfer' (i.e. investment in the children of wives). In calculating fitness payoffs, they used the assumptions of the G78 model. As we have seen, this minimizes the payoff from investing in sisters' children. Consequently, their model may overstate the stability of vertical transfers. Similarly, Fortunato [20] shows that selection can favour investment in sisters' children even when  $p > \frac{1}{2}$ . This conclusion is conservative because it also relies on the G78 model. Investment in sisters' children would evolve even more easily if the assumptions of this model were relaxed.

We can estimate men's kinship,  $\frac{\bar{p}}{4}$ , to wives' offspring directly from the population-wide average paternity confidence,  $\bar{p}$ . But we cannot estimate kinship to sisters' children from this value alone. It may lie anywhere between the lower and upper bounds,  $(1 + \bar{p}^2)/16$  and  $\frac{1}{8}$ . The difference between these bounds decreases with increasing values of  $\bar{p}$ . Among the Himba, for example, Scelza [22] estimates that  $\bar{p} = 0.83$ . This implies that kinship to sisters' children is between 0.106 and 0.125, a range of 18 per cent. Had  $\bar{p}$  been lower—say  $\frac{1}{2}$ —the range of uncertainty would have been 60 per cent.

Without measuring, one cannot know where within this range of uncertainty any real population lies. To remove the uncertainty, we need separate estimates of paternity confidence ( $\bar{p}$ ) and of the probability ( $\bar{h}$ ) that two siblings share a father.

I am grateful for comments from Ryan Bohlender, Monique Borgerhoff Mulder, Elizabeth Cashdan, Mark Flinn, Laura Fortunato, Henry Harpending, Kristen Hawkes and Michael Lewis.

## Endnote

<sup>1</sup>I use the term 'extrapair' rather than 'extramarital', because relationships that produce children need not be sanctioned as marriages. Nonetheless, I assume that one male is primary and refer to him as the 'husband'.

## References

- Davies N. 1992 *Dunnock behavior and social evolution*. Oxford, UK: Oxford University Press.
- Kurland JA. 1979 Paternity, mother's brother and human sociality. In *Evolutionary biology and human social behavior: an anthropological perspective* (eds N Chagnon, W Irons), pp. 145–180. North Scituate, MA: Duxbury Press.
- Gaulin SJ, Schlegel A. 1980 Paternal confidence and paternal investment: a cross cultural test of a sociobiological hypothesis. *Ethol. Sociobiol.* **1**, 301–309. (doi:10.1016/0162-3095(80)90015-1)
- Flinn M. 1981 Uterine vs. agnatic kinship variability and associated cousin marriage preferences: an evolutionary biological analysis. In *Natural selection and social behavior: recent research and new theory* (eds RD Alexander, DW Tinkle), pp. 439–475. New York, NY: Chiron Press.
- Hartung J. 1981 Paternity and inheritance of wealth. *Nature* **291**, 652–654. (doi:10.1038/291652a0)

6. Hartung J. 1985 Matrilineal inheritance: new theory and analysis. *Behav. Brain Sci.* **8**, 661–688. (doi:10.1017/S0140525X00045520)
7. Wolfe L, Gray JP. 1981 Comment on Gaulin and Schlegel (1980). *Ethol. Sociobiol.* **2**, 95–98. (doi:10.1016/0162-3095(81)90039-X)
8. Gray JP, Wolfe L. 1982 A note on brother inheritance. *Ethol. Sociobiol.* **3**, 103–105. (doi:10.1016/0162-3095(82)90006-1)
9. Hartung J. 1982 Comment: reply to Gray and Wolfe. *Ethol. Sociobiol.* **3**, 101. (doi:10.1016/0162-3095(82)90005-X)
10. Alexander RD. 1974 The evolution of social behavior. *Annu. Rev. Ecol. Syst.* **5**, 325–383. (doi:10.1146/annurev.es.05.110174.001545)
11. Greene PJ. 1978 Promiscuity, paternity, and culture. *Am. Ethnol.* **5**, 151–159. (doi:10.1525/ae.1978.5.1.02a00110)
12. Greene PJ. 1980 Paternity and the avunculate. *Am. Anthropol.* **82**, 381–382. (doi:10.1525/aa.1980.82.2.02a00110)
13. Alexander R. 1977 Natural selection and the analysis of human sociality. In *Changing scenes in natural sciences, 1776–1976* (ed. C Goulden), pp. 283–337. Philadelphia, PA: Academy of Natural Sciences.
14. Diamond A, Locay L. 1989 Investment in sister's children as behavior towards risk. *Econ. Inquiry* **27**, 719–735. (doi:10.1111/j.1465-7295.1989.tb00798.x)
15. Gaulin S, McBurney D, Brakeman-Wartell S. 1997 Matrilateral biases in the investment of aunts and uncles. *Hum. Nat.* **8**, 139–151. (doi:10.1007/s12110-997-1008-4)
16. Bergstrom TC. 1996 Economics in a family way. *J. Econ. Lit.* **34**, 1903–1934.
17. Holden C, Sear R, Mace R. 2003 Matriliney as daughter-biased investment. *Evol. Hum. Behav.* **24**, 99–112. (doi:10.1016/S1090-5138(02)00122-8)
18. Nolin DA. 2011 Kin preference and partner choice: patrilineal descent and biological kinship in Lamaleran cooperative relationships. *Hum. Nat.* **22**, 156–176. (doi:10.1007/s12110-011-9113-9)
19. Fortunato L, Archetti M. 2010 Evolution of monogamous marriage by maximization of inclusive fitness. *J. Evol. Biol.* **23**, 149–156. (doi:10.1111/j.1420-9101.2009.01884.x)
20. Fortunato L. 2012 The evolution of matrilineal kinship organization. *Proc. R. Soc. B* **279**, 4939–4945. (doi:10.1098/rspb.2012.1926)
21. Gillespie JH. 2004 *Population genetics: a concise guide*, 2nd edn. Baltimore, MD: Johns Hopkins University Press.
22. Scelza B. 2011 Female choice and extra-pair paternity in a traditional human population. *Biol. Lett.* **7**, 889–891. (doi:10.1098/rsbl.2011.0478)