



The measurement and meaning of inclusive fitness

DAVID C. QUELLER

Department of Ecology and Evolutionary Biology, Rice University

*(Received 9 May 1995; initial acceptance 12 May 1995;
final acceptance 20 July 1995; MS. number: AS-1161)*

Evolutionary adaptation is a genetic process, but fortunately we can often understand it, to a first approximation, without knowing the genetic details. We use what Grafen (1982) called the phenotypic gambit: traits with the highest fitness will tend to be favoured. An important exception is in certain kinds of social behaviour, such as altruism, where population genetic models demonstrate that traits lowering fitness can be favoured by selection. In this case, the phenotypic gambit can often be preserved by introducing one piece of genetic information, the relatedness of the altruist to its beneficiary. Although this seemingly pollutes the phenotypic gambit with genetics, the main advantage is preserved, for we still assume that we can ignore the genetics underlying the particular trait. The concept of inclusive fitness therefore provides a crucial shortcut for studying social behaviour, and understanding how to apply it is an important goal.

Lucas et al. (1996) showed that a method of applying inclusive fitness suggested by Creel (1990) seems to lead to certain anomalies. In this note, I argue that such anomalies are perhaps not surprising for two reasons. First, Creel's (1990) method has not been adequately justified, and is likely to lead to errors. Second, the paradox that led Creel to this method has another resolution.

A little history is required to set the stage. Hamilton (1964) defined inclusive fitness as the individual's

production of adult offspring . . . stripped of all components which can be considered as due to the individual's social environment, leaving the fitness he would express if not exposed to any of the harms or benefits of that environment . . . aug-

mented by certain fractions of the quantities of harm and benefit which the individual himself causes to the fitness of his neighbors. The fractions in question are simply the coefficients of relationship [page 8, my italics]

Grafen (1982, 1984) showed that problems arise if one omits the part I have emphasized about stripping all components due to the individual's social environment. The result of this omission is double-counting of fitness components and a wrong result.

Creel (1990), however, noticed that applying this definition seems to lead to a paradox. Imagine an obligately social species: solitary reproduction is impossible. Creel's dwarf mongooses, *Helogale parvula*, approach this condition, but assume for simplicity that the condition is met absolutely. Assume further that one individual does all the reproducing in the group. Now if we calculate this individual's inclusive fitness, we find that it seems to be zero; all of its reproduction comes from the efforts of its helpers, and it must all therefore be stripped from the fitness of the reproductive. Only the helpers have positive inclusive fitness because they are assigned the indirect effects on the dominant's reproduction (devalued by their relatedness, r). The apparent paradox lies in the fact that the reproductive has a lower inclusive fitness, which seems to predict that individuals ought to compete to be helpers. This prediction goes contrary to empirical observations that the reproductive position is sought after, and it clashes with the obvious fact that more of the reproductive's genes get transmitted than the helper's.

Creel's solution, which is followed by Lucas et al. (1996), was to alter the quantity that is supposed to be stripped from direct reproduction. Instead of stripping the effect of that individual's particular social environment, he recommended stripping the same amount, the average social effect, from all individuals. That something is

Correspondence: D. C. Queller, Department of Ecology and Evolutionary Biology, Rice University, 6100 Main St., Houston, TX 77251-1892, U.S.A. (email: queller@pop.rice.edu).

		Individual 1	
		Help	Repr.
Individual 2	Repr.	0	0
	Help	0	0

		Individual 1	
		Help	Repr.
Individual 2	Repr.	1	0
	Help	1	0

Figure 1. Payoff matrices for two individuals, each of which can choose to try to reproduce or to help its partner reproduce. Group living is considered to be obligatory. (a) Reproduction is successful only if one individual chooses the reproductive role and the other chooses to help. (b) The total reproduction of the group is divided equally if both individuals choose the same role.

amiss with this approach can be deduced from prior results. Suppose we begin with a model in which we neglect to strip off the social effect, a model that we know leads to incorrect results (Grafen 1982, 1984). Now, follow Creel's suggestion and strip the identical amount from everyone's direct fitness. With respect to the difference between two strategies, which is how we reason about inclusive fitness, these two models must give the same result. It is the same as if we had not stripped anything at all so we must be retaining the double-counting error that fitness-stripping was supposed to eliminate.

Creel (1990) arrived at his method through a modification of Hamilton's (1964) results. Both papers will need to be consulted to confirm what follows, which is a sketch of why I believe that the modification was unjustified. Creel began with the equation $a^* = T^*$. These quantities both refer to population averages, as stated in the text, so let's write it as $\bar{a} = \bar{T}$. The meanings of the two variables can be ignored; what matters is that this equation is required for the subsequent steps, and the conclusion (Creel's equation 8) is only justified as a statement about population means: $e^\circ = \bar{dT}^\circ$. This states that the average amount stripped away (e) must equal the average effect of an individual on others (dT). Creel asserted that this equality holds for each individual, $e^\circ = \bar{dT}^\circ$, but this has not been justified mathematically. So when Hamilton argues that e° should be stripped from the fitness of each individual, we have not justified substituting the average, \bar{dT}° .

But if Creel's solution doesn't work, what is the solution to his paradox? I believe that there are

two answers, depending upon the expression of the gene in question. First, I will consider a case in which the altruism gene is expressed in all group members, irrespective of their likelihood of becoming the dominant individual. Then, I will consider a conditionally expressed gene in which one individual is granted first choice over whether it will be the reproductive. This right to choose is conferred by some external factor, which may be related to ability to win a contest (e.g. size of fighting ability) or may even be uncorrelated with its ability to win (e.g. age or order of arrival). The question then becomes: should this individual accept or reject the opportunity to be the reproductive?

First consider the unconditionally expressed gene. The key feature to bear in mind is that obligate sociality is a non-additive, game-theoretic situation. The payoffs to each individual depend non-additively on what strategy its partner chooses. Figure 1 gives two examples of possible payoff matrices for the simplest case, where a group consists of two individuals (still assuming that solitary living is impossible). In Fig. 1a, successful reproduction occurs only if one individual adopts the reproductive role and the other adopts the helper role. The payoffs are non-additive, which means that the payoff for changing roles always depends on the role choice of the partner (or equivalently, that the payoff for changing both row and column is not the same as the sum of the separate payoff changes). They remain non-additive even in Fig. 1b, where the reproduction is equally divided when both players choose the same role. Hamilton's (1964) model,

and most other inclusive fitness models, assume additivity of fitness effects. These models fail, to a greater or lesser degree, when fitness effects are non-additive (see Queller 1984, 1985, and references therein), which amounts to saying that Creel's paradox is not really a paradox that needs resolution. Instead it is a case that is recognized to fall outside the domain of standard inclusive fitness theory. For those who want to apply inclusive fitness theory this is not a happy resolution, but fortunately we can often do somewhat better.

The first approach might be to consider versions of inclusive fitness theory that take non-additivity into account (e.g. Uyenoyama & Feldman 1982; Karlin & Matessi 1983; Queller 1985, 1992). For the simple two-player game, a valid version of Hamilton's rule can be derived (Queller 1992): $\beta_{WG \cdot G'} + \beta_{G'G} \beta_{WG' \cdot G} > 0$. Here $\beta_{G'G}$ is the familiar genetic regression definition of relatedness (Hamilton 1972). The other β s are partial regression coefficients, and are interpreted as new definitions for the fitness components (costs and benefits) of Hamilton's rule. Partial regressions provide a statistical way of parcelling the non-additive interaction terms into the main terms. This result is theoretically valid, but its generality comes at a cost that can be appreciated by closely examining the partial regressions. $\beta_{WG \cdot G'}$ is the effect on an individual's fitness (W) of its own genes (G, technically the additive genetic value for the behaviour in question), while holding constant the effect of its partner's genes (G'). Similarly, $\beta_{WG' \cdot G}$ is the effect of its partner's genes, holding constant the effect of its own genes. These are the effects on fitness of genes, not phenotypes, and to apply them correctly we must know the underlying genetics. In other words, we have had to abandon the phenotypic gambit.

The phenotypic gambit can be salvaged, however, when the genes have the right kind of conditional behaviour (Charlesworth 1980; Parker 1989). Consider first the case where one individual of each pair is clearly stronger than its partner, and can assume whichever role it chooses. If individual 1 chooses to be the reproductive, then individual 2 is left with being a helper. There is now no question of both being reproductives or both being helpers. Individual 1 has the choice of having two offspring as the reproductive or helping its partner have two offspring. Clearly the former is better, but we again seem to be faced

with Creel's paradox: the reproductive has those two offspring only because of its helper, so should they not be stripped from its direct fitness, leaving zero? The answer is no. The reason for this stripping was to avoid double-counting the effects of genes. Here that reason does not apply, because the helper has never had the opportunity to express the genes under consideration, which are genes for accepting or rejecting the reproductive role, given that the bearer is the stronger partner. Indeed we have no basis for attributing anything to a partner who has not expressed the genes in which we are interested. We can only predict the helper's genes for this trait via its relatedness to the reproductive, not the other way around. The help that the weaker partner gives the reproductive must get counted to the reproductive's inclusive fitness, because it cannot be attributed to the helper. The correct inclusive fitness analysis for the trait expressed by individual 1 is simply that it should be a reproductive because $2 > r \cdot 2$ (the second 2 is the cost that the reproductive imposes on the helper by not letting it be the reproductive). So there is no paradox, and the inclusive fitness extension of the phenotypic gambit is preserved.

The conclusion is not altered if individual 2 has some other option besides helping, such as leaving, provided that this choice is controlled by genes that are independent of those controlling individual 1's choice. Individual 2's choice might now affect individual 1's payoffs, but there is still no non-additive genetic interaction acting on either set of genes because each is expressed in only one individual.

So, the resolution of the paradox depends on the nature of the trait. If it is unconditionally expressed, we transgress on the assumptions of inclusive fitness theory. If the trait is conditionally expressed, then the trait becomes additive and we can apply inclusive fitness theory, but we must remember that we are only interested in the effects of that conditional trait. I further suggest that we might expect the conditional situation to be more common, partly because the conditional strategy described above is a more sophisticated strategy that takes into account more information. We might expect the unconditional strategy to be replaced by the one conditioned on relative strength for roughly the same reason that the strategy of taking a drink every half hour is replaced by the strategy of drinking when thirsty. But there is still stronger reason to expect

conditional strategies, even when the winner of a fight is not easily predicted (Maynard Smith & Parker 1976; Pollock, in press, shows this specifically for related partners). A pair of partners do not primarily compete against each other. Rather, they compete against the whole population, and for that reason, even arbitrary asymmetries can usefully assign roles. To put it another way, the arbitrary strategy, 'be a reproductive when you are older than your partner, otherwise be a helper', can often successfully replace an unconditional strategy, because it avoids the penalty of two partners choosing the same strategy (Fig. 1a) or the penalty of escalated contests.

This argument has an interesting implication for the generality of the inclusive fitness concept. It implies that inclusive fitness is trait-specific. Unlike traditional fitness, there is no such thing as an individual's inclusive fitness, independent of the trait being considered. Consider individual 2 in the conditional example above. Her partner (individual 1) has chosen the reproductive role, leaving only two options, help or go try to reproduce alone. The inclusive fitness of the latter strategy is zero (no direct or indirect reproduction). The inclusive fitness of helping, given that the stronger partner has chosen to reproduce, is r^*2 , because of the indirect reproduction. This assignment of help to the inclusive fitness of individual 2 seems to contradict the earlier conclusion that it should be attributed to individual 1. In fact, it is not a contradiction because we are considering the evolution of two different traits, one for what to do when one is stronger, expressed only in individual 1, and one for what to do when one is weaker, expressed only in individual 2. Our inclusive fitness accounting assigns the help to individual 1 in the first case and to individual 2 in the second, but there is no double-counting of the same genes. Two separate counts are required for two separate traits controlled by two different sets of genes.

Hamilton's (1964) definition of inclusive fitness, quoted above, seems a bit too broad. His model was for a particular trait, and the mathematics only justifies stripping away effects of helpers that arise from their expression of that particular trait.

It did not rigorously justify the concept of an individual's inclusive fitness that was independent of the trait being considered, and in this sense, inclusive fitness differs from standard individual fitness. Strictly speaking, we should avoid talking about the inclusive fitness of individuals and instead speak only of the inclusive fitness of traits or strategies, or of the inclusive fitness of individuals for some particular trait. Supported by NSF grants BIR-19451, IBN-9507515 and DEB-9510126.

REFERENCES

- Charlesworth, B. 1980. Models of kin selection. In: *Evolution of Social Behavior: Hypotheses and Empirical Tests* (Ed. by H. Markl), pp. 229–231. Weinheim: Verlag Chemie.
- Creel, S. 1990. How to measure inclusive fitness. *Proc. R. Soc. Lond. Ser. B*, **241**, 229–231.
- Grafen, A. 1982. How not to measure inclusive fitness. *Nature, Lond.*, **298**, 425.
- Grafen, A. 1984. Natural selection, kin selection and group selection. In: *Behavioral Ecology*. 2nd edn (Ed. by J. R. Krebs & N. B. Davies), pp. 62–84. Sunderland, Massachusetts: Sinauer.
- Hamilton, W. D. 1964. The genetical evolution of social behaviour. I. *J. theor. Biol.*, **7**, 1–16.
- Hamilton, W. D. 1972. Altruism and related phenomena, mainly in the social insects. *A. Rev. Ecol. Syst.*, **3**, 193–232.
- Karlin, S. & Matessi, C. 1983. Kin selection and altruism. *Proc. R. Soc. Lond. Ser. B*, **219**, 327–353.
- Lucas, J. R., Creel, S. R. & Waser, P. M. 1996. How to measure inclusive fitness, revisited. *Anim. Behav.*, **51**, 225–228.
- Maynard Smith, J. & Parker, G. A. 1976. The logic of asymmetric contests. *Anim. Behav.*, **24**, 159–175.
- Parker, G. A. 1989. Hamilton's rule and conditionality. *Ethol. Ecol. Evol.*, **1**, 195–211.
- Pollock, G. B. In press. Kin selection, kin avoidance, and correlated strategies. *Evol. Ecol.*
- Queller, D. C. 1984. Kin selection and frequency dependence: a game theoretic approach. *Biol. J. Linn. Soc.*, **23**, 133–143.
- Queller, D. C. 1985. Kinship, reciprocity and synergism in the evolution of social behavior. *Nature, Lond.*, **318**, 366–367.
- Queller, D. C. 1992. A general model for kin selection. *Evolution*, **46**, 376–380.
- Uyenoyama, M. K. & Feldman, M. 1982. Population genetic theory of kin selection. II. The multiplicative model. *Am. Nat.*, **120**, 614–627.