

Can natural selection favour indiscriminate spite?

2 Matishalin Patel^{1†}, Stuart A. West¹ & Jay M. Biernaskie²

4 ¹Department of Zoology, University of Oxford
Zoology Research and Administration Building; 11a Mansfield Road; Oxford OX1 3SZ

6 [†]Author for correspondence (matishalin.patel@sjc.ox.ac.uk)

8 ²Department of Plant Sciences, University of Oxford
South Parks Road; Oxford OX1 3RB

10 **Abstract**

Spiteful behaviours occur when an actor harms its own fitness to inflict harm on the fitness of
12 others. Several papers have predicted that spite can be favoured in sufficiently small populations,
even when the harming behaviour is directed indiscriminately at others. However, it is not clear
14 that truly spiteful behaviour could be favoured without the harm being directed at a subset of
social partners with relatively low genetic similarity to the actor (kin discrimination). Using
16 mathematical models, we show that: (1) the evolution of spite requires kin discrimination; (2)
previous models suggesting indiscriminate spite involve scenarios where the actor gains a direct
18 feedback benefit from harming others, and so the harming is selfish rather than spiteful; (3)
extreme selfishness can be favoured in small populations (and in some cases small groups)
20 because this is where the feedback benefit of harming is greatest.

22 **Keywords:** competition, harming, inclusive fitness, kin selection, negative relatedness, social
evolution, super-territory, territory size.

24 **Introduction**

25 Spite is the hardest type of social trait to explain. Spiteful behaviour reduces the lifetime fitness
26 of both the recipient and the performer (actor) of that behaviour (Hamilton 1970). In terms of
Hamilton's rule, $-C + RB > 0$, spite represents the case where there is a fitness cost to the actor
28 (positive C), and a fitness cost to the harmed recipient (negative B), which can only be favoured
if the genetic relatedness term, R , is negative. Understanding the meaning of negative relatedness
30 is therefore crucial for explaining how and why spite evolves.

It has been argued that the evolution of spite requires kin discrimination, allowing the
32 actor to direct harm towards a subset of individuals with whom they share relatively low genetic
similarity (Foster & Ratineks 2000; Foster et al. 2001; Gardner & West 2004a,b, 2006; Gardner et
34 al. 2004, 2007; Lehmann et al. 2006; West & Gardner 2010). Specifically, spite can be favoured
when harming the less-similar individuals in a social group (primary recipients) reduces
36 competition and therefore benefits the unharmed individuals (secondary recipients). In this case,
negative relatedness arises because the actor's genetic similarity to primary recipients is less than
38 its genetic similarity to secondary recipients (Lehmann et al. 2006). In contrast, without kin
discrimination, harming behaviours could not be directed at individuals to whom the actor is
40 negatively related, so indiscriminate spite should be impossible.

However, a number of theoretical studies have suggested the possibility for indiscriminate
42 spite. Hamilton (1970) originally suggested that if genetic similarity is measured relative to the
entire population (including the actor), then there will be a negative relatedness between the actor
44 and all others in the population, especially in small populations. Consequently, several papers

have predicted that spiteful harming, directed indiscriminately at others, could be favoured in
46 sufficiently small populations (Hamilton 1970, 1971; Grafen 1985; Vickery et al. 2003; Taylor
2010; Smead & Forber 2012). As a specific example, Verner (1977) and Knowlton and Parker
48 (1979; Parker & Knowlton 1980) suggested that individuals could be favoured to hold territories
that are larger than needed for their own interest (“super-territories”), in order to spitefully
50 exclude others from resources. It is not clear, though, whether such indiscriminate harming traits
are truly spiteful.

52 Here, we resolve this disagreement over indiscriminate spite. Many harming traits will be
costly to primary recipients ($B < 0$) but provide a direct fitness benefit to the actor, because they
54 reduce competition for the actor or its offspring. Consequently, the traits are selfish ($-C > 0$)
rather than spiteful ($-C < 0$) (Hamilton 1970; Keller et al. 1994; Foster et al. 2001; West &
56 Gardner 2010). We address the possibility that indiscriminate harming traits like territory size
have been misclassified as spiteful when they are actually selfish (Colgan 1979; Tullock 1979).
58 Our specific aims are to: (1) determine generally whether indiscriminate harming evolves as a
spiteful or a selfish trait; (2) examine how different modelling approaches can change the
60 meaning of negative relatedness and lead to misclassification of harming traits; (3) re-analyse the
Knowlton & Parker (1979) territory-size model to determine whether it predicts spiteful
62 behaviour.

Harming traits

64 We first modelled natural selection acting on a harming trait, following the approach of Lehmann
et al. (2006). The trait has a fitness effect on a focal actor ($-C$) and on two categories of

66 recipients: the harmed primary recipients and the unharmed secondary recipients who benefit
from reduced competition (fitness effects B_1 and B_2 , respectively). Crucially, we define an
68 individual's fitness as its number of offspring that survive to adulthood (not simply the number of
offspring produced), which is consistent with other definitions used for classifying social traits
70 (Hamilton 1964; Rousset 2004; Lehmann et al. 2006; West et al. 2007). We assume that fitness
effects on the actor, primary recipients, and secondary recipients must sum to zero because of
72 competition for finite resources (Rousset & Billiard 2000):

$$74 \quad -C + B_1 + B_2 = 0, \quad (1)$$

76 implying that any decrease in fitness for one category necessarily means an increase in fitness for
another. Our model could apply to any finite population of constant size or to a local “economic
78 neighborhood” (Queller 1994) in which there is a zero-sum competition for access to the next
generation. Key examples of such local competition include polyembryonic wasps competing for
80 resources inside a host (Gardner & West 2004a; Gardner et al. 2007) or male fig wasps
competing for females inside a fig (West et al. 2001).

82 To predict the direction of natural selection acting on the harming trait, we considered the
fate of a mutant harming allele in a population of individuals with a fixed, resident genotype. The
84 success of the mutant allele depends on its “inclusive fitness effect” (Hamilton 1964): the sum of
effects from a focal actor's mutant trait on its own fitness and on the total fitness of each recipient
86 category, weighted by their genetic similarity with the actor. Under the usual assumptions of
weak selection and additive gene action, the inclusive fitness effect for our model is

88

$$\Delta W_{IF} = -C + B_1 Q_1 + B_2 Q_2, \quad (2)$$

90

where Q_1 and Q_2 are probabilities of sharing identical genes between the focal actor and a random individual from the primary and secondary recipients, respectively. We note that the fitness effects in Equation 2 could alternatively be weighted by relatedness coefficients, where genetic similarity is measured with respect to a reference population (e.g., $R_i = \frac{Q_i - \bar{Q}}{1 - \bar{Q}}$, where \bar{Q} is the average genetic similarity in the entire population, including the actor; Hamilton 1970). However, doing this would not change any of the results given below. We therefore prefer the simpler approach used in Equation 2 and what follows below.

98

In the following sections, we examine two different ways of defining the category of secondary recipients and therefore partitioning the fitness effects of harming. Both methods correctly predict the direction of selection (they give the same sum as in Eq. 2). The first partitioning also maintains complete separation of direct and indirect fitness effects ($-C$ and RB , respectively), making it appropriate for classifying harming traits as selfish ($-C > 0$) or spiteful ($-C < 0$). In contrast, the second partitioning obscures the separation of direct and indirect fitness effects, making it inappropriate for classifying traits in this way.

100

Is indiscriminate harming spiteful or selfish?

106

We determined the conditions for a harming trait to be classified as spiteful or selfish. For this purpose, we assume that the focal actor, primary recipients, and secondary recipients are mutually exclusive categories. This ensures that the actor is not a recipient of its own behaviour, and so the

108

110 $-C$ term in the inclusive fitness effect (Eq. 2) captures all effects of the actor's harming behaviour
on its own fitness. From Equation 2, we derived the typical two-party version of Hamilton's rule
by eliminating the fitness effect on secondary recipients, using $B_2 = C - B_1$ (from Eq. 1). After
112 rearrangement, the inclusive fitness effect is positive, and the harming trait is favoured, when

$$114 \quad -C + \frac{Q_1 - Q_2}{1 - Q_2} B_1 > 0, \quad (3)$$

116 which is Hamilton's rule with the relatedness between actor and primary recipients given by
 $\frac{Q_1 - Q_2}{1 - Q_2} \equiv R_1$. This is the genetic similarity between the actor and an individual from the potential
118 primary recipients, measured relative to an individual from the potential secondary recipients.

Equation 3 implies that indiscriminate spite cannot evolve. This is because negative
120 relatedness (and hence an indirect fitness benefit of harming) will arise only if harm can be
directed at primary recipients who are less genetically similar to the actor than secondary
122 recipients are ($Q_1 < Q_2$). In contrast, if the actor were harming others indiscriminately—for
example, harming a random subset of a population or local economic neighbourhood—then its
124 expected similarity to these primary recipients would be the same as to the set of potential
secondary recipients ($Q_1 = Q_2$), and relatedness would be zero ($R_1 = 0$). This implies that
126 indiscriminate harming will be favoured when it is a selfish trait with a positive direct fitness
benefit ($-C > 0$).

128 **Why does misclassification occur?**

Misclassification of harming traits can occur because the fitness effects of social traits can be
130 partitioned in different ways (Frank 1998). An alternative way of partitioning the effects of
harming is to include the actor in the set of secondary recipients who may benefit from reduced
132 competition. In fact, it is often implicitly assumed that the set of potential secondary recipients is
the entire population (or economic neighbourhood), including the focal actor (Hamilton 1970,
134 1971; Grafen 1985; Vickery et al. 2003; Taylor 2010; Smead & Forber 2012). To make this
explicit, we re-write the inclusive fitness effect as

$$\Delta W_{\text{IF}} = -c + b_1 Q_1 + b_2 \bar{Q} , \quad (4)$$

136
138 using lower-case letters to indicate that the fitness effects no longer match those from Equation 2.
140 In particular, b_2 is now the benefit of reduced competition that may be experienced by all
individuals in population (including the actor), and \bar{Q} is the probability of genetic identity
142 between the focal actor and a random individual the entire population (including itself). It follows
that $-c$ is not a total direct fitness effect because it excludes the secondary benefit of harming that
144 feeds back to the focal actor (increased direct fitness due to reduced competition; Fig. 1).

We used Equation 4 to derive an analogue of Hamilton's rule, which reveals a different
146 version of negative relatedness. For example, in a population (or economic neighbourhood) of N
individuals, an actor could indiscriminately harm a random subset of individuals with genetic
148 similarity Q_1 to the actor. If the entire population is in the set of secondary recipients, then the
expected genetic similarity between the actor and these recipients is $\bar{Q} = \frac{1}{N} 1 + \frac{N-1}{N} Q_1$ (where the

150 first term accounts for the actor's similarity to itself). Eliminating the fitness effect on secondary
recipients (using $b_2 = c - b_1$), shows that indiscriminate harming is favoured when

152

$$-c + \frac{-1}{N-1} b_1 > 0 , \quad (5)$$

154

where $-1/(N-1)$ is the relatedness between actor and primary recipients, measured with respect
156 to the entire population ($\frac{q_1 - \bar{q}}{1 - \bar{q}} \equiv R_{1,p}$). This is the version of negative relatedness that has led to
predictions of indiscriminate spite in small populations (e.g., Hamilton 1971; Grafen 1985).

158 However, although the term $\frac{-1}{N-1} b_1$ resembles an indirect fitness benefit ($RB > 0$), it
actually accounts for the secondary fitness benefit of harming that feeds back to the focal actor.
160 This can be made more explicit by deriving an analogue of Hamilton's rule from Equation 4, this
time eliminating the fitness effect on primary recipients (using $b_1 = c - b_2$). For example, in a
162 well-mixed population of N individuals, indiscriminate harming is favoured when

164

$$-c + \frac{1}{N} b_2 > 0 , \quad (6)$$

166 where $1/N$ is the relatedness between the actor and the entire population (including itself),
measured with respect to primary recipients ($\frac{\bar{q} - q_1}{1 - q_1} \equiv R_{2,p}$). The term $(1/N)b_2$ accounts for the
168 fraction of the secondary benefit (reduced competition) that feeds back to the focal actor, which

gets larger as the actor makes up a larger fraction of the population.

170 Our key distinction here is that harming behaviours can be either beneficial or costly to
the actor ($-C > 0$ or $-C < 0$), whereas spiteful behaviours are strictly costly to the actor ($-C < 0$).
172 We showed that indiscriminate harming is always favoured because it is beneficial to the actor—
it has a positive effect on the actor’s number of surviving offspring ($-C > 0$). Moreover,
174 indiscriminate harming can be favoured most in small populations (or small economic
neighbourhoods) because this is where the focal actor can benefit most from the reduced
176 competition that results from its harming behaviour.

Re-visiting “super-territories”

178 We next re-examined the territory size model from Knowlton & Parker (1979; Parker &
Knowlton 1980). We first analysed the model to fully separate direct and indirect fitness effects
180 (applying Eq. 2), asking whether the model predicts selfish behaviour, as expected. We then used
the alternative approach (applying Eq. 4) to illustrate why previous studies have interpreted
182 territory size as a spiteful trait.

We considered a finite, deme-structured population (“island model”; Wright 1943) with d
184 demes (assuming $d > 1$) and n individuals competing for territory in each deme (total population
size is $N = dn$). Individuals that secure a territory have offspring and then die before a fraction m
186 of their offspring disperse independently to a random deme in the entire population. All
individuals have a genetically-determined strategy for the size of territory that they try to obtain
188 (a continuous trait). Taking over a larger territory has three key effects: (1) it incurs a fecundity
cost for the actor (we assume a linear cost with increasing trait size, with slope $-a$ and $a \in [0,1]$);

190 (2) it harms the actor's deme mates by taking resources away and reducing their fecundity; (3) it
reduces the competition faced by all remaining offspring in the population to secure a territory in
192 the next generation.

We first assumed that the actor, primary recipients, and secondary recipients are mutually
194 exclusive categories (as in Eq. 2). In the Appendix, we derive an expression for the fitness, W , of
a focal actor as a function of its own territory-size strategy, x ; the average strategy of its deme
196 mates (primary recipients), y ; and the average strategy of individuals in other demes (secondary
recipients), z . We used this "neighbour-modulated" fitness function to derive the inclusive fitness
198 effect, by taking partial derivatives with respect to the strategies of the different categories of
individuals (Taylor & Frank 1996; Rousset & Billiard 2000):

200

$$\begin{aligned}\Delta W_{\text{IF}} &= \frac{\partial W}{\partial x} + \frac{\partial W}{\partial y} Q_1 + \frac{\partial W}{\partial z} Q_2, \\ &= -C + B_1 Q_1 + B_2 Q_2\end{aligned}\tag{7}$$

202

where all partial derivatives are evaluated in a monomorphic population ($x = y = z$). We derive
204 expressions for Q_1 and Q_2 in the Appendix, and with these we determined the equilibrium of the
model (\hat{z} , where directional selection stops) by solving $\Delta W_{\text{IF}} = 0$. We also checked that the
206 equilibrium is a convergence-stable strategy, denoted z^* , meaning that if the population is
perturbed from the equilibrium then natural selection will push it back ($\left. \frac{d\Delta W_{\text{IF}}}{dz} \right|_{z=\hat{z}} < 0$).

208 We found that the equilibrium of our model, $z^* = 1/(aN)$, is identical to that originally
predicted by Parker & Knowlton (1980); however, our analysis shows that the optimal territory

210 size strategy is selfish rather than spiteful. Territory size cannot be spiteful in this model because
the actor's genetic similarity to individuals in other demes is always equal to or less than the
212 similarity to deme mates ($Q_1 \geq Q_2$). Accordingly, the relatedness to primary recipients (measured
relative to secondary recipients) is never negative ($R_1 \geq 0$), and so there is no indirect benefit of
214 larger territory size. Moreover, when offspring dispersal is limited ($m < 1$) and deme mates are
positively related ($R_1 > 0$), there is no indirect benefit of smaller territory size (as a form of
216 helping). This is because limited dispersal increases competition among offspring within the
deme, which promotes harming and exactly cancels the effect of positive relatedness (as in Taylor
218 1992). Territory size therefore evolves for its direct benefit only, with larger territories promoted
by a smaller fecundity cost to the actor (smaller a) and smaller population size (smaller N).
220 Specifically, the direct fitness effect at equilibrium ($z = z^*$) is

$$222 \quad -C = \frac{a(d-1)d(m-1)^2}{N-1}, \quad (8)$$

224 which is either positive (when $m < 1$) or zero (when $m = 1$). In the case of full offspring dispersal
($m = 1$), the equilibrium is the point where the fecundity cost to the actor is exactly balanced by
226 the feedback benefit experienced by its offspring (reduced competition for space in the next
generation). As the population approaches this equilibrium, however, direct fitness is always
228 positive ($-C > 0$), confirming that territory size evolves as a selfish trait (Fig. 2).

We next assumed that the set of secondary recipients is the entire population, including
230 the focal actor (as in Eq. 4). In this case, the inclusive fitness effect is

232
$$\Delta W_{IF} = \frac{\partial W}{\partial x} + \frac{\partial W}{\partial y} Q_1 + \frac{\partial W}{\partial z_p} \bar{Q}, \quad (9)$$

$$= -c + b_1 Q_1 + b_2 \bar{Q}$$

234 where z_p is the average territory size strategy in the entire population (including the focal actor),
and all partial derivatives are evaluated at $x = y = z_p$. As expected, solving for the equilibrium of
236 Equation 9 gives the same answer as before, $z^* = 1/(aN)$.

This version of the model shows, however, why territory size could be misclassified as
238 spiteful. For example, in a fully mixing population at the equilibrium ($m = 1$; $z_p = z^*$), the first
term in Equation 9 is

240
$$-c = -\frac{aN}{N-1}, \quad (10)$$

242 which is always negative. This term reflects the fecundity cost of the focal actor's territory size
strategy; however, it is not the total direct fitness effect because it excludes the feedback benefit
244 experienced by the actor's offspring (reduced competition). As noted above, when $m = 1$ this
feedback benefit should exactly balance the fecundity cost at equilibrium. Following Equations 5
246 or 6, we can calculate the feedback benefit as $(-1/[N-1])b_1$ or $(1/N)b_2$ (both evaluated at $z_p = z^*$),
248 which gives the expected result, $aN/(N-1)$. The partitioning in Equation 9 therefore splits the
total direct fitness effect of territory size into two separate terms, $-c + (-1/[N-1])b_1$ or $-c +$

250 $(1/N)b_2$, which could be misinterpreted as a direct fitness cost ($-C < 0$) and an indirect fitness
benefit ($RB > 0$).

252 **Discussion**

We examined a general model of harming traits and a specific model where larger territory size is
254 an indiscriminate harming trait. In both models we found that: (1) the evolution of spite requires
kin discrimination, where the actor harms only a subset of other individuals (those with relatively
256 low genetic similarity); (2) without kin discrimination, harming can be favoured but only when
there is a sufficient direct, feedback benefit to the actor (reduced competition for the actor or its
258 offspring); (3) indiscriminate harming can be favoured most in small populations (or small
economic neighbourhoods), where the feedback benefit to the actor is greatest; (4) previous
260 studies have misclassified indiscriminate harming as spite, partly because they misinterpret the
feedback benefit as an indirect (kin-selected) benefit ($RB > 0$). Overall, our analyses illustrate
262 why indiscriminate harming traits are selfish rather than spiteful.

Classifying harming traits

264 For the purposes of classifying harming traits, we found that it is easiest to treat the actor,
primary recipients, and secondary recipients as separate categories. This makes it straightforward
266 to separate the total direct and indirect fitness effects of harming ($-C$ and RB , respectively) and
ensures that non-zero relatedness will always be associated with an indirect fitness effect. For
268 example, spiteful harming ($-C < 0$, $B < 0$) requires that harm is directed at primary recipients to
whom the actor is negatively related (with respect to secondary recipients; $Q_1 < Q_2$ and $R_1 < 0$),
270 resulting in a positive indirect fitness effect ($R_1B > 0$) (Lehmann et al. 2006). In contrast, when

harming is indiscriminate, the actor has zero relatedness to primary recipients (with respect to
272 secondary recipients; $Q_1 = Q_2$ and $R_1 = 0$), and so harming can be favoured as a selfish trait only
($-C > 0$, $B < 0$).

274 We showed that misclassification of indiscriminate harming is due to an implicit
assumption that the focal actor is a secondary recipient of its own behaviour (Hamilton 1970,
276 1971; Grafen 1985; Vickery et al. 2003; Taylor 2010; Smead & Forber 2012). This means that
some of the actor's direct benefit of harming has been accounted for by a fraction of the fitness
278 effects on recipients, giving the appearance of an indirect benefit ($RB > 0$). For example, in a
well-mixed population where all individuals (including the actor) are considered secondary
280 recipients, a fraction of the fitness effect on primary recipients ($-1/[N - 1] B_1$) actually
contributes to the direct benefit of indiscriminate harming.

282 Others have suggested that harming traits should be classified based on their primary
effects only, rather than their total fitness effects (Krupp 2013). This means that indiscriminate
284 harming traits like larger territory size, which may be associated with a survival or fecundity cost
($-c < 0$ in the terms of our model), would be classified as spiteful, despite the feedback benefit to
286 the focal actor. We argue, however, that a classification based on total fitness effects ($-C$ and RB)
is more useful (Hamilton 1964; West et al. 2007). This is because it emphasises the fundamental
288 distinction between spiteful harming, which is favoured by indirect fitness benefits and requires
kin discrimination, versus selfish harming, which is favoured by direct fitness benefits and does
290 not require kin discrimination (West & Gardner 2010). Similar arguments have been made for
maintaining the distinction between altruistic helping ($-C < 0$, $B > 0$) and mutually-beneficial
292 helping ($-C > 0$, $B > 0$) (West et al. 2007).

Indiscriminate harming in nature

294 We found that selfish indiscriminate harming can be favoured most in small populations or small
economic neighbourhoods (e.g., small groups with relatively local competition). This is because
296 harming primary recipients leads to reduced competition for all individuals in the population or
group, and a focal actor receives a larger fraction of this secondary benefit when it makes up a
298 larger fraction of the population or group. Indiscriminate harming can therefore be thought of as
producing a type of public good for secondary recipients (Tullock 1979), analogous to
300 indiscriminate helping, which is often thought of as a public good for primary recipients. A key
difference is that indiscriminate helping is inhibited by local competition (Taylor 1992; Griffin et
302 al. 2004); in contrast, indiscriminate harming requires local competition so that the focal actor
can actually benefit the reduced competition that results from its harming (Gardner & West
304 2004b).

So where can we expect to find the most extreme examples of selfish harming? As
306 recognised by Hamilton (1970), very small populations will tend to extinction, so harming traits
in these populations are unlikely to be observed. But examples of extreme selfishness should be
308 found in small groups with relatively local competition, such that harming other individuals
significantly reduces competition for the actor. One potential example is in fig wasps, where
310 males fight for access to females, and the intensity of fighting increases sharply as the number of
males in the fig declines (Murray et al. 1989; Reinhold 2003; West et al. 2001). Further potential
312 examples include competition among female honey bees for a colony and other cases where
males engage in local competition for mates (e.g., *Melittobia* parasitoids; Griffin & West 2002).

314

316 **Acknowledgements:** We thank Guy Cooper, Asher Leeks, Alan Grafen, and Tom Scott for
comments on the manuscript.

318 **Author contributions:** All authors conceived and designed the study, MP drafted the initial
version of the manuscript, and all authors contributed to later versions of the manuscript.

320

References

322 Colgan, P. (1979). Is a super-territory strategy stable? *Am Nat* 144: 604-605.

Foster, K.R., Ratnieks, F. & Wenseleers, T. (2000). Spite in social insects. *Trends Ecol Evol* 15:
324 469–70.

Foster, K.R., Wenseleers, T. & Ratnieks, F. (2001). Spite: Hamilton's unproven theory. *Annales*
326 *Zoologici Fennici* 38:229–38.

Frank, S.A. (1998). *Foundations of Social Evolution*. Princeton University Press, Princeton, NJ.

328 Gardner, A. & West, S. A. (2004a). Spite among Siblings. *Science* 305: 1413–1414.

Gardner, A. & West, S. A. (2004b) Spite and the scale of competition. *J Evol Biol* 17: 1195–1203.

330 Gardner A. & West S.A. (2006). Spite. *Curr. Biol.* 16:R662–R664.

Gardner, A., West, S. A. & Buckling, A. (2004). Bacteriocins, spite and virulence. *Proc Biol Sci*
332 271: 1529–1535.

Gardner, A., Hardy, I. C. W., Taylor, P. D. & West, S. A. (2007). Spiteful soldiers and sex ratio
334 conflict in polyembryonic parasitoid wasps. *Am. Nat.* 169: 519–533.

- 336 Grafen, A. (1985). A geometric view of relatedness. In: *Oxford Surveys in Evolutionary Biology*
2: 28–90.
- Griffin A.S. & West S.A. (2002). Kin selection: fact and fiction. *Trends Ecol Evol* 17:15–21.
- 338 Griffin A.S., West, S.A. & Buckling, A. (2004). Cooperation and competition in pathogenic
bacteria. *Nature* 430:1024–7.
- 340 Hamilton, W. D. (1964). The genetical evolution of social behaviour. I and II. *Journal of*
Theoretical Biology 7: 1–52.
- 342 Hamilton, W. D. (1970). Selfish and spiteful behaviour in an evolutionary model. *Nature* 228:
1218–1220.
- 344 Hamilton, W.D. (1971). Selection of selfish and altruistic behaviour in some extreme models. In:
Man and Beast: Comparative Social Behavior {eds. Eisenberg J.F. & Dillon W.S.}. Smithsonian
346 Press, Washington DC.
- Keller, L., Milinski, M., Frischknecht, M., Perrin, N., Richner, H. & Tripet, F. (1994). Spiteful
348 animals still to be discovered. *Trends Ecol Evol* 9:103–3.
- Knowlton, N. & Parker, G. A. (1979). An evolutionarily stable strategy approach to
350 indiscriminate spite. *Nature* 279: 419–421.
- Krupp, D. B. (2013). How to distinguish altruism from spite (and why we should bother). *Journal*
352 *of Evolutionary Biology* 26: 2746–2749.
- Lehmann, L., Bargum, K. & Reuter, M. (2006). An evolutionary analysis of the relationship
354 between spite and altruism. *J Evol Biol* 19: 1507–1516.

- Murray, M.G. (1989). Environmental constraints on fighting in flightless male fig wasps. *Animal*
356 *Behaviour* 38:186–93.
- Parker, G. A. & Knowlton, N. (1980). The evolution of territory size—some ESS models.
358 *Journal of Theoretical Biology* 84: 445–476.
- Queller, D.C. (1994). Genetic relatedness in viscous populations. *Evol Ecol* 8:70–73.
- 360 Reinhold, K. (2003). Influence of male relatedness on lethal combat in fig wasps: a theoretical
analysis. *Proc. Roy. Soc. Lond. B* 270:1171–1175.
- 362 Rousset, F. (2004). *Genetic Structure and Selection in Subdivided Populations*. Princeton
University Press, Princeton NJ.
- 364 Rousset, F. & Billiard, S. (2000). A theoretical basis for measures of kin selection in subdivided
populations: finite populations and localized dispersal. *Journal of Evolutionary Biology* 13: 814–
366 825.
- Smead, R. & Forber, P. (2012). The evolutionary dynamics of spite in finite populations.
368 *Evolution* 67: 698–707.
- Taylor, P. D. (1992). Altruism in viscous populations—an inclusive fitness model. *Evol Ecol*. 6:
370 352–356.
- Taylor, P.D. (2010). Birth-death symmetry in the evolution of a social trait. *J Evol Biol*. 23: 2569-
372 2578.
- Taylor, P. D. & Frank, S. A. (1996). How to make a kin selection model. *Journal of Theoretical*
374 *Biology* 180: 27–37.

- Taylor, P. D., Irwin, A. J. & Day, T. (2001). Inclusive fitness in finite deme-structured and
376 stepping-stone populations. *Selection* 1:153–164.
- Tullock, G. (1979). On the adaptive significance of territoriality: comment. *Am. Nat.* 113:772-
378 775.
- Verner, J. (1977). On the adaptive significance of territoriality. *Am. Nat.* 111: 769–775.
- 380 Vickery, W. L., Brown, J. S. & FitzGerald, G. J. (2003). Spite: altruism’s evil twin. *Oikos* 102:
413–416.
- 382 West, S. A. & Gardner, A. (2010). Altruism, spite, and greenbeards. *Science* 327: 1341–1344.
- West, S.A., Griffin, A. & Gardner A. (2007). Social semantics: altruism, cooperation, mutualism,
384 strong reciprocity and group selection. *J Evol Biol.* 20:415–32.
- West, S.A., Murray, M.G., Machado, C.A., Griffin, A.S. & Herre, E.A. (2001). Testing
386 Hamilton's rule with competition between relatives. *Nature* 409:510–513.

388 **Appendix: Territory-size model**

Deriving the fitness function

390 Here, we derive an expression for the fitness of a focal actor with a mutant territory size strategy,
based on the models of Knowlton and Parker (1979; Parker and Knowlton 1980). We consider a
392 population that is structured into d demes of n individuals competing for territories, where each
deme has A units of available territory. The focal actor’s strategy, x , represents a continuous
394 number of territory units that it attempts to gain ($x > 0$). The average strategy of the actor’s deme

mates is y , and the average strategy in all other demes is z .

396 We first calculate the expected offspring production (expected fecundity, F) for the focal
actor, an individual in the actor's deme, and an individual in another deme. These expected
398 values depend on: (1) the probability of an individual acquiring a territory (assuming that
available spaces are acquired completely randomly); (2) the cost associated with the individual's
400 strategy (assuming fecundity declines linearly with increasing territory size strategy; $f(x) = 1 -$
 ax , where $0 < a < 1$). For the focal actor, there are A/y spaces available in the deme, and we use
402 the simplifying assumption that a mutant individual has priority to claim the territory units
denoted by its strategy (Knowlton and Parker 1979). Therefore, the focal actor has a $1/n$
404 probability of acquiring a territory, and its expected fecundity is

$$406 \quad F_x = \frac{1}{n} \frac{A}{y} f(x). \quad (\text{A1})$$

408 The space available for others in the patch depends on whether or not the focal actor claims a
territory. The actor gains access to the patch with probability A/ny , and in this case $(A - x)/y$
410 spaces remain; otherwise, A/y spaces are available. The expected fecundity for one of the $n - 1$
deme mates of the focal actor is therefore

$$412 \quad F_y = \frac{1}{n-1} \left(\frac{A}{ny} \frac{A-x}{y} f(y) + \left(1 - \frac{A}{ny} \right) \frac{A}{y} f(y) \right). \quad (\text{A2})$$

414

Finally, for an individual in another deme in the population, there are A/z spaces available, and so
 416 the expected fecundity for one of these individuals is

$$418 \quad F_z = \frac{1}{n} \frac{A}{z} f(z) . \quad (A3)$$

420 We next calculate the focal actor's fitness, $W(x, y, z)$, which is the number of its offspring
 that survive to compete for a territory in the next generation. This can be partitioned into two
 422 terms, the first term accounting for offspring that compete on the focal actor's natal deme (those
 that did not disperse, with probability $1-m$, and those that dispersed but landed on the natal deme,
 424 with probability m/d) and the second term accounting for offspring that disperse with probability
 m to compete in the $d - 1$ non-natal demes:

$$426 \quad W = \frac{\left(1-m + \frac{m}{d}\right)F_x}{(1-m) F_x + (n-1) (1-m) F_y + \frac{1}{d} (mF_x + (n-1) m F_y) + \frac{d-1}{d} n m F_z} + \frac{\frac{d-1}{d} m F_x}{(1-m) n F_z + \frac{1}{d} (mF_x + (n-1) m F_y) + \frac{d-1}{d} n m F_z} , \quad (A4)$$

428 where the denominator of the first and second terms account for, respectively, all offspring
 430 competing in the focal actor's natal deme and all offspring competing in any other deme in the
 population. Equation A4 is the fitness function used to calculate the inclusive fitness effect in
 432 Equation 7 of the main text. To express the focal individual's fitness in terms of x , y , and z_p (the
 average territory size strategy in the entire population, including the focal individual), we

434 substituted $(x + (n - 1)y - dnz_p)/(n - nd)$ for z in Equation A4. This gives the fitness function used
to calculate the inclusive fitness effect in Equation 9 of the main text.

436 **Deriving probabilities of genetic identity**

Next, we derive probabilities of genetic identity by descent in a finite deme-structured
438 population, following the approach of Taylor et al. (2000). In particular, we needed the
probability of identity between the focal actor and a randomly selected deme mate (Q_1), between
440 the actor and a randomly selected individual in another deme (Q_2), and between the actor and a
randomly selected individual in the entire population (including itself), defined as

442

$$\bar{Q} = \frac{1}{d} \left(\frac{1}{n} + \frac{n-1}{n} Q_1 \right) + \frac{d-1}{d} Q_2 . \quad (\text{A5})$$

444

The remaining probabilities of identity are given by the following recursive equations:

446

$$Q_1 = \left((1 - m)^2 \left(\frac{1}{n} + \frac{n-1}{n} Q_1 \right) + (1 - (1 - m)^2) Q_2^p \right) (1 - u)^2 \quad (\text{A6})$$

448

$$Q_2 = \left((1 - m)^2 Q_2 + (1 - (1 - m)^2) Q_2^p \right) (1 - u)^2 , \quad (\text{A7})$$

450

where u is the “contrived mutation rate” from Taylor et al. (2000). We solved Equations B1-B3
452 simultaneously and evaluated the solution in the limit of a low mutation rate ($u \rightarrow 0$), giving:

454 $Q_1 = 1 - 2dnu$ (A8)

$$Q_2 = 1 + \left(\frac{2d(m-1)^2}{(m-2)m} - 2dn \right) u \quad (\text{A9})$$

456 $\bar{Q} = 1 + \frac{2(d(1 - (m-2)m(n-1)) - 1)}{(m-2)m} u$. (A10)

458 These are the probabilities of genetic identity used in Equations 7 and 9 of the main text.

460

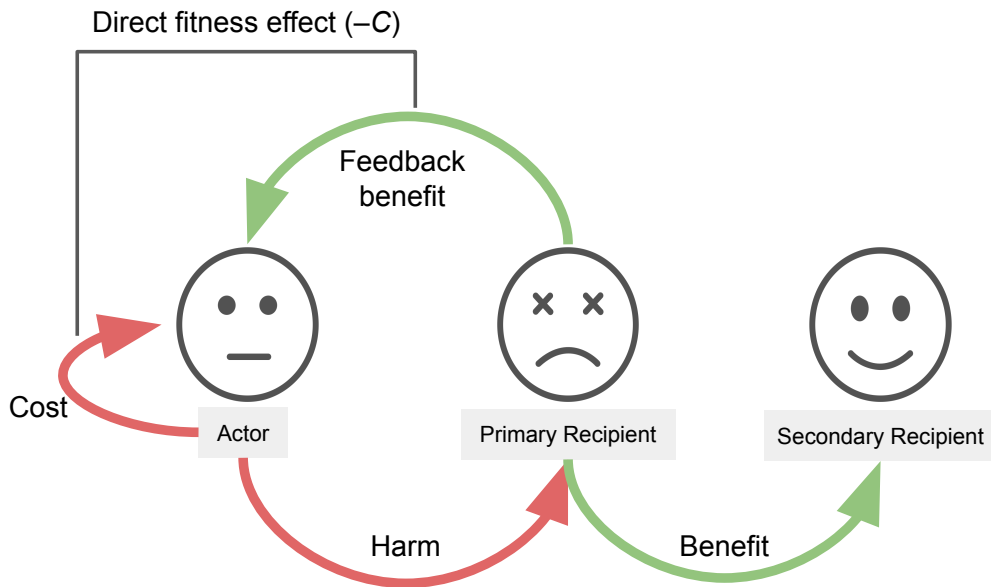


Figure 1. Partitioning the fitness effects of a harming trait. When a focal actor harms a primary recipient, this reduces competition and may therefore benefit the unharmed secondary recipients and the actor itself (“feedback benefit”). Some modelling approaches include the actor in the set of secondary recipients of the harming trait. However, the total direct fitness effect ($-C$ in Hamilton’s rule) includes the fecundity cost of expressing the harming trait plus the feedback benefit.

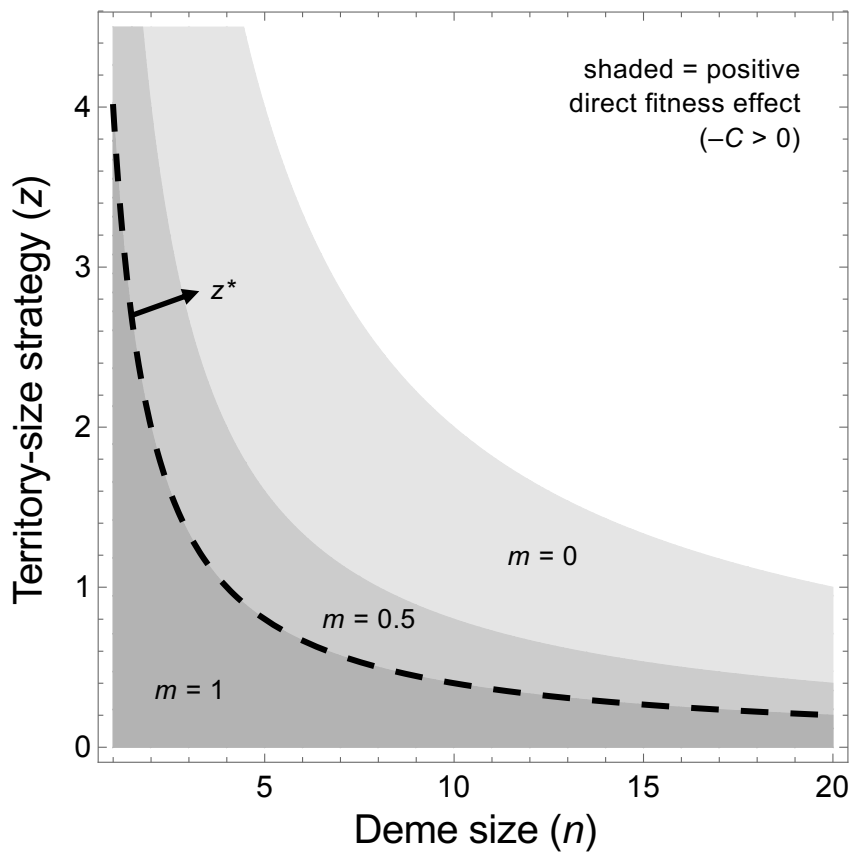


Figure 2. Territory size and direct fitness. Larger territory size is promoted by smaller population size (smaller dn) and reduced offspring migration from the deme (smaller m), both of which increase the direct benefit to an actor for harming its deme mates. However, reduced migration also increases the relatedness among deme mates, which inhibits larger territory size. Ultimately, the optimal territory size strategy (z^* , dashed line) is independent of migration rate and evolves as if the population were fully mixed ($m = 1$). Other parameters used: $d = 5$, $a = 0.05$.