Evolution, 2025, **79(6)**, 962–972
https://doi.org/10.1093/evolut/qpaf032
Advance access publication 16 February 2025
Original Article

Kin-discriminating partner choice promotes the evolution of helping

Thomas W. Scott[1,2], Geoff Wild[3], Andy Gardner[1]

[1]School of Biology, University of St Andrews, Dyers Brae, St Andrews, United Kingdom
[2]Present address: Department of Biology, University of Oxford, Oxford, United Kingdom
[3]Department of Mathematics, Western University, London, Ontario, Canada
Corresponding author. Email: thomas.scott@biology.ox.ac.uk

Abstract

Kin selection theory predicts that individuals should evolve to help relatives, either by helping indiscriminately in a population where they do not move very far from their relatives, or by discriminating kin and conditionally helping them. It has been argued that, because kin discrimination enables individuals to reduce how helpful they are with some social partners as well increase how helpful they are with others, this could lead to an increase or a decrease in the overall level of helping. Specifically, it was argued that kin discrimination would increase the overall level of helping if the function relating the optimal level of help and genetic relatedness is convex, but kin discrimination would decrease the overall level of helping if the function relating the optimal level of help and genetic relatedness is concave. However, this prediction was based on a model in which individuals were not able to choose their social partners but only adjust how helpful they should be toward those social partners they have been allocated. Here, we perform a mathematical analysis showing that being able to choose social partners increases the overall level of helping. Consequently, if kin discriminators are allowed to choose whom they help, kin discrimination is more likely to increase the overall level of helping than previously anticipated. We obtained these results in two complementary theoretical settings: one more general, which makes few demographic assumptions, and the other more specific and concrete, which assumes a patch-structured population with complete dispersal.

Keywords: social evolution, altruism, inclusive fitness, kin recognition, veil of ignorance, cooperation

Introduction

Kin selection theory predicts that—at all levels of biology, from bacteria to humans—high genetic relatedness between social partners favors the evolution of helping (Hamilton, 1963, 1964; Maynard Smith, 1964; West et al., 2007). Genetic relatedness promotes helping because, by increasing a relative's reproduction, a helpful individual is still passing its genes to the next generation, just indirectly. In viscous populations, where individuals do not disperse very far from their relatives, social partners tend to be closely related, such that individuals may evolve to help whomever they meet; this is called "indiscriminate helping" (Hamilton, 1964). Alternatively, in populations that may be well-mixed, individuals may evolve to identify their relatives using either genetic or environmental cues of relatedness, and then help them on a conditional basis; this is called "kin discrimination" (Green et al., 2024).

It is unclear whether overall levels of helping are expected to be higher or lower in indiscriminately helping versus kin-discriminating populations. Faria and Gardner (2020) pointed out that kin discriminators will evolve to help less when encountering nonrelatives as well as helping more when encountering close relatives, whereas indiscriminate helpers will exhibit an intermediate extent of helping with everyone. This difference between kin discriminators and non-discriminators implies there exists a geometric rule for determining which of these types displays the greater overall level of helping. If the optimal level of helping increases linearly with relatedness, indiscriminate helpers and kin discriminators are predicted to invest the same overall amount in helping. However, if the optimal level of helping is an accelerating (i.e., convex) function of relatedness, kin discriminators will disproportionately increase their help when encountering close relatives, resulting in more helping overall than compared with indiscriminate helpers. Conversely, if the optimal level of helping is a decelerating (i.e., concave) function of relatedness, kin discriminators will disproportionately decrease their help when encountering nonrelatives, resulting in less helping overall than compared with indiscriminate helpers.

Overall, Faria and Gardner (2020) argued that kin discrimination sometimes increases and sometimes decreases the overall level of helping. However, Faria and Gardner's analysis did not account for an important characteristic of kin discrimination behavior: while they assumed that kin discriminators can use relatedness to inform the degree to which they help a given partner, they neglected the possibility that discriminators might also use relatedness to inform partner choice (Scott, 2024; Scott et al., 2022, 2023). Partner choice may allow kin discriminators to reliably pick out close relatives, whom they will help intensively, increasing the overall level of helping in the population more often, and to a greater degree, than anticipated by Faria and Gardner, (2020).

Received October 30, 2024; revisions received January 16, 2025; accepted February 12, 2025

Associate Editor: Bram Kuijper; Handling Editor: Tim Connallon

© The Author(s) 2025. Published by Oxford University Press on behalf of The Society for the Study of Evolution (SSE).
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

We investigate this possible helpfulness-promoting effect of partner choice by analyzing the evolution of helping to see when kin discrimination results in a higher overall level of helping in the population than does indiscriminate helping. First, we generalize Faria and Gardner's (2020) analysis, allowing kin discriminators to engage in kin-discriminating partner choice as well as—or instead of—kin-discriminating helping. This reveals that kin-discriminating partner choice tends to promote helping, such that the overall level of help is higher under kin discrimination not only for linearity and convexity but also—in contrast to Faria and Gardner's prediction—in some concavity scenarios. Second, we illustrate our general results by developing a concrete model of pairwise social interactions in a subdivided population, in which kin discriminators have multiple social encounters before deciding with whom to help. We discuss the challenge that these results pose for the application of "veil of ignorance" thinking to social evolution.

General analysis

We consider a population in which every individual enacts help toward one social partner. Each individual obtains its social partner in one of two ways. First, the individual samples s individuals from its neighborhood. Then, in the "no kin-discriminating partner choice" scenario, a random individual from this sample is assigned as the individual's social partner; we denote the individual's relatedness to this social partner by R, and the expectation of this random variable by $\mathrm{E}(R)$. Here, $\mathrm{E}(R)$ is a population-level statistic; we make no assumption that the distribution of relatedness is the same for all patches. If $\mathrm{E}(R) > 0$, then individuals in the neighborhood are likely to be relatives. In the "kin-discriminating partner choice" scenario, the most highly related individual in the sample is assigned as the individual's social partner; we denote the individual's relatedness to this social partner by $R_{\max}$, and the expectation of this random variable by $\mathrm{E}(R_{\max})$. On these assumptions, in the "no kin-discriminating partner choice" scenario, each individual is simply assigned a social partner. But in the "kin-discriminating partner choice" scenario, every individual has some choice as to who their social partner will be, with their choice being informed by their degree of relatedness to each potential social partner.

In Appendix A, we justify our assumption that individuals capable of kin-discriminating partner choice will choose the closest relative amongst the s possible social partners (rather than, for instance, the least related individual), by showing that this maximizes their inclusive fitness. For greater s, individuals capable of kin-discriminating partner choice have a larger sample of potential social partners to choose from, meaning they are more likely to acquire a close relative, such that kin-discriminating partner choice is more effective. Therefore, s captures the efficacy of kin recognition, i.e., the capacity to perfectly identify close relatives (Reeve, 1989). We make no assumptions about how individuals recognize their kin—this could be through a genetic cue, such as a shared genetically encoded odor, or an environmental cue, such as growing up in the same nest (Green et al., 2024).

How much help an individual enacts toward their social partner is also determined in one of two ways: in the "no kin-discriminating helping" scenario, individuals do not adjust their level of helping according to the relatedness of their social partner, and in the "kin-discriminating helping" scenario, individuals may adjust their level of helping according to the relatedness of their social partner. In the former scenario, individuals effectively commit to a level of helping before they have found a partner; in the latter scenario, the commitment comes after the partnership is set.

Accordingly, we consider four different scenarios by combining how partnerships and helping are determined in different ways: no kin-discriminating helping and no kin-discriminating partner choice ("H-C-"); kin-discriminating helping and no kin-discriminating partner choice ("H+C-"); no kin-discriminating helping and kin-discriminating partner choice ("H-C+"); and kin-discriminating helping and kin-discriminating partner choice ("H+C+"). The first scenario (H-C-) corresponds to indiscriminate helping, and the latter three scenarios (H+C-, H-C+, and H+C+) correspond to different types of kin discrimination.

In the "no kin-discriminating helping" scenarios (H-C- and H-C+), all individuals in the population are expected to exhibit the same level of helping; specifically, that which is optimized according to the expected relatedness of social partners. We define the level of helping as $z_{\text{H-C-}} = Z(\mathrm{E}(R))$ in the "no kin-discriminating partner choice" scenario (H-C-) and $z_{\text{H-C+}} = Z(\mathrm{E}(R_{\max}))$ in the "kin-discriminating partner choice" scenario (H-C+). In the "kin discriminating helping" scenarios (H+C- and H+C+), individuals across the population are expected to vary in their levels of helping; specifically, they are expected to exhibit a level of helping which is optimized according to their specific relatedness to social partners. Accordingly, we define the population-average level of helping as $z_{\text{H+C-}} = \mathrm{E}(Z(R))$ in the "no kin-discriminating partner choice" scenario (H+C-) and $z_{\text{H+C+}} = \mathrm{E}(Z(R_{\max}))$ in the "kin-discriminating partner choice" scenario (H+C+). We assume that Z is an increasing function of its argument (relatedness).

Not only relatedness but also the costs and benefits of helping will modulate the level of helping that is favored in different contexts. In making our comparison between indiscriminate versus kin-discriminating helping we assume that all else is equal—that is, the costs and benefits of helping are the same in these two scenarios. Moreover, we assume that kin discrimination gives individuals information only about relatedness and not about the costs and benefits of helping insofar as these also vary across different contexts. We make no assumptions about the scale of competition—that is, we are agnostic as to whether competition for resources occurs locally amongst relatives, or globally amongst nonrelatives (Frank, 1998; Queller, 1994; Taylor, 1992b, a). Specifically, our framework takes the relationship between optimized helping and relatedness (Z) as a given, and we examine the consequences of this for whether kin discrimination increases or decreases the overall level of helping. The scale of competition would affect the shape that Z takes; but, given that Z takes a particular shape, the scale of competition will not have any further effect. All details are summarized in Table 1. In Appendix B, we obtain the following results.

If the optimal level of helping (Z) is a linear function of its argument (relatedness), kin-discriminating helping neither increases nor decreases the overall level of helping. Mathematically this is because the linearity of Z leads to $z_{\text{H-C-}} = z_{\text{H+C-}}$ and $z_{\text{H-C+}} = z_{\text{H+C+}}$ (Figure 1b). Additionally, kin-discriminating partner choice increases the overall level of helping. Mathematically this is because, since Z is an

Table 1. Relatedness and overall level of helping in each of the four scenarios.

		Kin-discriminating partner choice?			
		No		Yes	
		Relatedness	Help	Relatedness	Help
Kin-discriminating helping?	No	$\mathrm{E}(R)$ (partner not known when deciding level of help)	$z_{\mathrm{H-C-}} = Z(\mathrm{E}(R))$	$\mathrm{E}(R_{\max})$ (partner not known when deciding level of help)	$z_{\mathrm{H-C+}} = Z(\mathrm{E}(R_{\max}))$
	Yes	R (partner known when deciding level of help)	$z_{\mathrm{H+C-}} = \mathrm{E}(Z(R))$	$R_{\max}$ (partner known when deciding level of help)	$z_{\mathrm{H+C+}} = \mathrm{E}(Z(R_{\max}))$

increasing function of its argument, $z_{\mathrm{H-C-}}, z_{\mathrm{H+C-}} \leq z_{\mathrm{H-C+}}, z_{\mathrm{H+C+}}$ (Figure 1b).

If the optimal level of helping (Z) is a concave (decelerating) function of its argument (relatedness), we find that kin-discriminating helping decreases the overall level of helping. This is because, in the case of kin-discriminating helping, the amount of help is averaged across the different relatedness contexts, where the average is reflected by the dashed lines in Figure 1. Consequently, the concavity of Z leads mathematically to $z_{\mathrm{H-C-}} \geq z_{\mathrm{H+C-}}$ and $z_{\mathrm{H-C+}} \geq z_{\mathrm{H+C+}}$ (Faria & Gardner, 2020; Jensen, 1906) (Figure 1a). Additionally, kin-discriminating partner choice increases the overall level of helping. Mathematically this is because, since Z is an increasing function of its argument, $z_{\mathrm{H-C+}} \geq z_{\mathrm{H-C-}}$ and $z_{\mathrm{H+C+}} \geq z_{\mathrm{H+C-}}$ (Figure 1a). We also obtain the following subsidiary results. First, kin-discriminating partner choice without kin-discriminating helping increases the overall level of helping relative to kin-discriminating helping without kin-discriminating partner choice ($z_{\mathrm{H-C+}} \geq z_{\mathrm{H+C-}}$, as implied by $z_{\mathrm{H-C+}} \geq z_{\mathrm{H-C-}}$ and $z_{\mathrm{H-C-}} \geq z_{\mathrm{H+C-}}$). Second, kin-discriminating partner choice with kin-discriminating helping may or may not increase the overall level of helping relative to indiscriminate helping ($z_{\mathrm{H-C-}} \geq z_{\mathrm{H+C+}}$ and $z_{\mathrm{H-C-}} \leq z_{\mathrm{H+C+}}$ are both possible). That is, it could be that the helpfulness-boosting effect of kin-discriminating partner choice outweighs the helpfulness-inhibiting effect of kin-discriminating helping, or alternatively the reverse could be true.

If the optimal level of helping (Z) is a convex (accelerating) function of its argument (relatedness), we find that kin-discriminating helping increases the overall level of helping. This is because, in the case of kin-discriminating helping, the amount of help is averaged across the different relatedness contexts, where the average is reflected by the dashed lines in Figure 1. Consequently, the convexity of Z leads mathematically to $z_{\mathrm{H-C-}} \leq z_{\mathrm{H+C-}}$ and $z_{\mathrm{H-C+}} \leq z_{\mathrm{H+C+}}$ (Faria & Gardner, 2020; Jensen, 1906) (Figure 1c). Additionally, kin-discriminating partner choice also increases the overall level of helping. Mathematically this is because, since Z is an increasing function of its argument, $z_{\mathrm{H-C+}} \geq z_{\mathrm{H-C-}}$ and $z_{\mathrm{H+C+}} \geq z_{\mathrm{H+C-}}$ (Figure 1c). We also obtain the following subsidiary results. First, kin-discriminating partner choice with kin-discriminating helping increases the overall level of helping relative to indiscriminate helping ($z_{\mathrm{H+C+}} \geq z_{\mathrm{H-C-}}$, as implied by $z_{\mathrm{H-C+}} \leq z_{\mathrm{H+C+}}$ and $z_{\mathrm{H-C+}} \geq z_{\mathrm{H-C-}}$). Second, kin-discriminating partner choice without kin-discriminating helping may or may not increase the overall level of helping relative to kin-discriminating helping without kin-discriminating partner choice ($z_{\mathrm{H+C-}} \geq z_{\mathrm{H-C+}}$ and $z_{\mathrm{H+C-}} \leq z_{\mathrm{H-C+}}$ are both possible). That is, it could be that the helpfulness-boosting effect of kin-discriminating partner choice exceeds the helpfulness-boosting effect of kin-discriminating helping, or alternatively the reverse could be true.

To sum up, although kin-discriminating helping per se might or might not boost overall levels of helping (with this depending on whether Z is convex or concave with respect to relatedness), kin-discriminating partner choice per se is expected to do so. If the two capacities always co-occur, then "kin discrimination" might or might not serve to boost overall levels of helping ($z_{\mathrm{H+C+}} \geq z_{\mathrm{H-C-}}$ and $z_{\mathrm{H+C+}} \leq z_{\mathrm{H-C-}}$ are both possible), but the dividing line does not coincide with the concavity / convexity distinction, *contra* Faria and Gardner (2020) (Figure 2). If Z is linear or convex (accelerating), then we expect "kin discrimination" to boost the overall level of helping ($z_{\mathrm{H+C+}} \geq z_{\mathrm{H-C-}}$), but if Z is concave (decelerating) then "kin discrimination" has the potential to increase or decrease the overall level of helping ($z_{\mathrm{H+C+}} \geq z_{\mathrm{H-C-}}$ and $z_{\mathrm{H+C+}} \leq z_{\mathrm{H-C-}}$ are both possible). We provide a graphical interpretation of these results in Figures 1 and 2.

The reason why kin-discriminating partner choice per se consistently increases the overall level of helping is that it allows individuals to pick out closer relatives, who elicit a higher level of helping. The reason why kin-discriminating helping per se does not necessarily increase the overall level of helping is that it simply allows individuals to fine-tune their level of helping to the relatedness of each social partner, but there is no guarantee that this fine tuning will result in greater helping overall (Faria & Gardner, 2020; Jensen, 1906).

In Appendix C, we obtain a further result, which is that, as partner choice becomes more effective (i.e., as s is increased), kin-discriminating helping has less effect on the overall level of helping, with no effect in the limit of perfect partner choice (Figure 2). This result arises because, with more effective partner choice, there is less relatedness variation amongst social partners, meaning there is no phenotypic consequence of being able to adjust one's level of help toward a given social partner. The implication of this result is that, if the capacities of kin-discriminating helping and kin-discriminating partner choice coincide, kin discrimination will increase the overall level of helping if there is sufficient partner choice (i.e., if s is sufficiently high).

Note that if we relax the assumption that Z is increasing and, indeed, say that it is constant with respect to its argument (relatedness), we find that there is no impact of kin-discriminating helping or kin-discriminating partner choice on the overall level of helping ($z_{\mathrm{H-C-}} = z_{\mathrm{H+C-}} = z_{\mathrm{H-C+}} = z_{\mathrm{H+C+}}$). In this case, we would still expect individuals to choose high-relatedness social partners, as it benefits one's inclusive fitness to provide a standard unit of help to a closer relative than to a more distant relative (Appendix A). This would apply to empirical scenarios where individuals distribute resources of

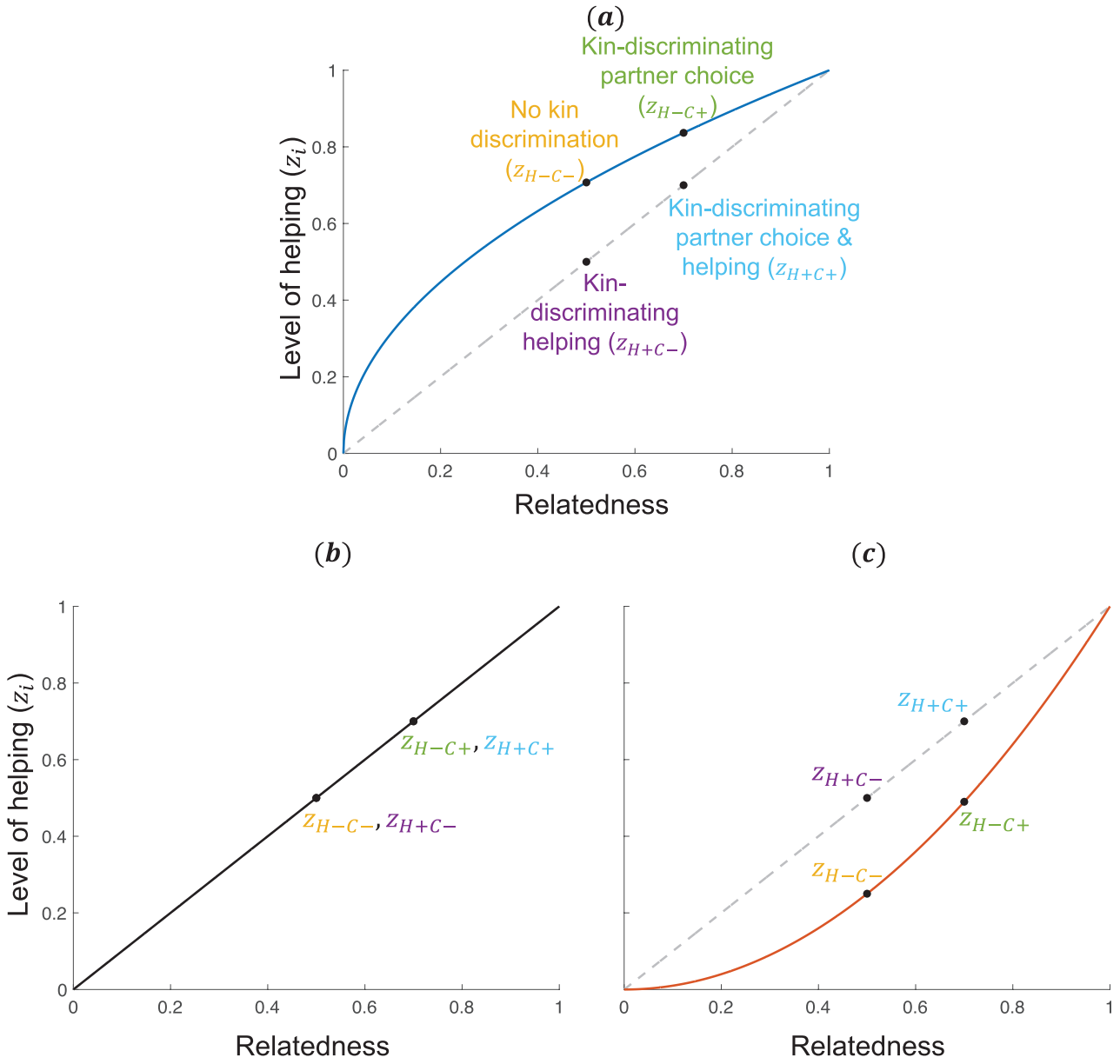


Figure 1. The effect of kin-discriminating helping and kin-discriminating partner choice on the overall level of helping. The black dots show the overall level of helping in four scenarios: no kin discrimination (z_{H-C-}); kin-discriminating helping (z_{H+C-}); kin-discriminating partner choice (z_{H-C+}); kin-discriminating partner choice & helping (z_{H+C+}). The solid lines in panels a–c show different relationships between the optimal level of helping and relatedness: concave (*a*); linear (*b*); convex (*c*). The gray dashed lines represent the average level of help across different relatedness contexts. Kin-discriminating partner choice increases relatedness, shifting the dots rightwards along the solid lines, resulting in an increase in the overall level of helping ($z_{H-C+} \geq z_{H-C-}$ & $z_{H+C+} \geq z_{H+C-}$). Kin-discriminating helping pulls the overall level of helping away from the solid lines and onto the secants (gray dashed lines), which may increase (*c*; $z_H \geq z_{H-C-}$ & $z_{H+C+} \geq z_{H-C+}$), decrease (*a*; $z_H \leq z_{H-C-}$ & $z_{H+C+} \leq z_{H-C+}$) or leave unchanged (*b*; $z_{H+C-} = z_{H-C-}$ & $z_{H+C+} = z_{H-C+}$) the overall level of helping.

fixed size to other individuals. These resources could be nests, territories, or even the actor's own body as a food offering as in instances of matriphagy. This contrasts with the more general scenario, in which individuals can decide how much help to give, as well as to whom it is given.

Illustration

We illustrate the results of our analysis in more concrete terms, by developing a model of pairwise social interactions in a subdivided population, where kin discriminators are allowed to have multiple social encounters before deciding who to help. We assume a patch-structured population with $n \geq 2$ hermaphrodite diploid breeders in each patch. Each breeder produces the same large number of offspring before dying, and these offspring interact socially within their patch, in a way that affects their survival to adulthood. All surviving individuals disperse away to seek reproductive opportunities elsewhere, mating monogamously with unrelated mating partners en route and arriving at a new patch (global competition). Following dispersal, n individuals are chosen at random on each patch to be breeders and all other individuals perish, returning the population to the beginning of the lifecycle.

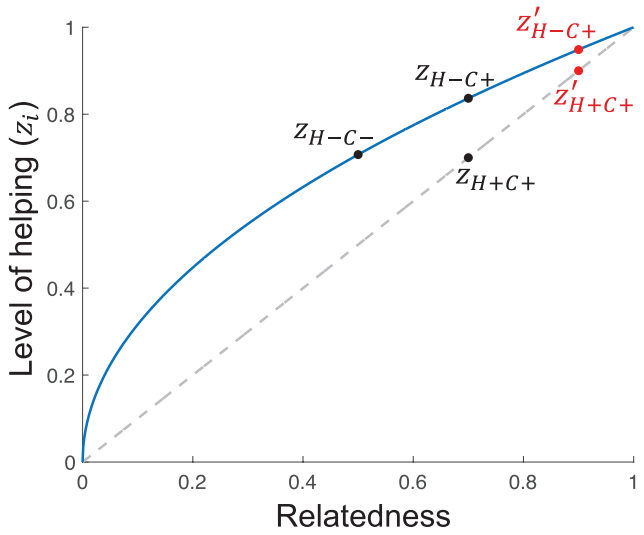


Figure 2. Kin discrimination increases the overall level of helping with sufficient partner choice, but need not in general. The black dots show the overall level of helping when there is: no kin discrimination (z_{H-C-}); kin-discriminating partner choice (z_{H-C+}); kin-discriminating partner choice & helping (z_{H+C+}). The red dots show the impact of further increasing partner choice, s (z'_{H-C+} & z'_{H+C+}). The solid line represents the relationship between the optimal level of helping and relatedness, which we assume here is concave. The gray dashed line represents the average level of help across different relatedness contexts. Kin-discriminating partner choice increases relatedness, which shifts the dots rightwards along the solid lines, increasing the overall level of helping; kin-discriminating helping pulls the overall level of helping away from the solid line and onto the secant (gray dashed line), decreasing the overall level of helping. There are two key results: (1) if the capacities for kin-discriminating helping and partner choice coincide, this may decrease the overall level of helping ($z_{H+C+} < z_{H-C-}$) or increase it ($z'_{H+C+} > z'_{H-C-}$), relative to indiscriminate helping; (2) for increased kin-discriminating partner choice, *s*, the influence of kin-discriminating helping on the overall level of helping, mediated by the curvature of the solid line, is reduced, as is evidenced by z'_{H-C+} & z'_{H+C+} being closer to each other than z_{H-C+} & z_{H+C+}.

These assumptions mean that, in every patch, each juvenile individual is related to a proportion $1/n$ of her patch-mates by one-half (siblings) and a proportion $(n\text{-}1)/n$ by zero (nonrelatives) (Michod & Hamilton, 1980). We assume that, in the lifecycle stage where social interactions occur, each individual obtains a 'recipient' by choosing the most-related individual from a sample of s of her patch mates, chosen randomly with replacement (Scott, 2024; Scott et al., 2022, 2023). Consequently, each individual has a probability $p = 1 - \left(\frac{n-1}{n}\right)^s$ of obtaining a sibling rather than a nonrelative recipient. This yields $p = 1/n$ for $s = 1$ (no kin-discriminating partner choice) and $p \to 1$ as $s \to \infty$ (perfect kin-discriminating partner choice).

Individuals then have the opportunity to help their recipients. A focal individual invests z into helping, at a personal survival cost $C(z)$, to give a survival benefit $B(z)$ to her recipient. We assume the following functional forms: $C(z) = \left(\frac{\beta}{1+\beta}\right) z^{\left(\frac{1+\beta}{\beta}\right)}$ and $B(z) = z$, where $\beta > 0$. We choose these functional forms because they conveniently have the focal individual's inclusive fitness $-C(z) + B(z)\, r$ being maximized at

$$z^* = r^\beta. \tag{1}$$

The β parameter determines the curvature of the function $Z(r) = r^\beta$ relating the optimal level of helping and relatedness. For $0 < \beta < 1$ the function is concave, for $\beta = 1$ the function is linear and for $\beta > 1$ the function is convex.

If there is neither kin-discriminating partner choice nor kin-discriminating helping (H-C-) then a focal individual's relatedness valuation of her recipient is $r = E(R) = 1/(2n)$, and making this substitution into Equation (1) yields a level of helping

$$z_{H-C-} = \left(\tfrac{1}{2n}\right)^\beta. \tag{2}$$

If there is no kin-discriminating partner choice but there is kin-discriminating helping (H+C-) then a focal individual's relatedness valuation of her recipient is $r = R$, i.e., with probability $p = 1/n$ her recipient is a sibling and hence related by $R = ½$, such that from Equation (1) she helps by an amount $z^* = (1/2)^\beta$, and with probability $1\text{-}p = (n\text{-}1)/n$ her recipient is a non-relative and hence related by $R = 0$, such that from Equation (1) she helps by an amount $z^* = 0$. Taking an average over all actors, this yields an overall level of helping

$$z_{H+C-} = \tfrac{1}{n}\left(\tfrac{1}{2}\right)^\beta. \tag{3}$$

If there is kin-discriminating partner choice but no kin-discriminating helping (H-C+) then a focal individual's relatedness valuation of her recipient is $r = E(R_{max})$, i.e., with probability $p = 1 - \left(\frac{n-1}{n}\right)^s$ her recipient is a sibling and hence related by $R_{max} = ½$ and with probability $1 - p = \left(\frac{n-1}{n}\right)^s$ her recipient is a non-relative and hence related by $R_{max} = 0$, which yields an expected relatedness of $E(R_{max}) = \frac{1}{2}\left(1 - \left(\frac{n-1}{n}\right)^s\right)$. Making this substitution into Equation (1) yields an overall level of helping

$$z_{H-C+} = \left(\tfrac{1}{2}\left(1 - \left(\tfrac{n-1}{n}\right)^s\right)\right)^\beta. \tag{4}$$

Finally, if there is both kin-discriminating partner choice and kin-discriminating helping (H+C+) then a focal individual's relatedness valuation of her recipient is $r = R_{max}$, i.e., with probability $p = 1 - \left(\frac{n-1}{n}\right)^s$ her recipient is a sibling and hence related by $R_{max} = ½$, which from Equation (1) elicits a level of helping $z^* = (1/2)^\beta$, and with probability $1 - p = \left(\frac{n-1}{n}\right)^s$ her recipient is a non-relative, which from Equation (1) elicits a level of helping $z^* = 0$. Taking an average over all actors, this yields an overall level of helping

$$z_{H+C+} = \left(1 - \left(\tfrac{n-1}{n}\right)^s\right)\left(\tfrac{1}{2}\right)^\beta. \tag{5}$$

Comparing our z values reveals that: if $0 < \beta < 1$ (i.e., concavity) then we have $z_{H\text{-}C\text{-}} \geq z_{H+C\text{-}}$, $z_{H\text{-}C+} \geq z_{H+C+}$, $z_{H\text{-}C+} \geq z_{H\text{-}C\text{-}}$ and $z_{H+C+} \geq z_{H+C\text{-}}$; if $\beta = 1$, (i.e., linearity) then we have $z_{H\text{-}C\text{-}} = z_{H+C\text{-}} \leq z_{H\text{-}C+} = z_{H+C+}$; and if $\beta > 1$ (i.e., convexity) then we have $z_{H\text{-}C\text{-}} \leq z_{H+C\text{-}}$, $z_{H\text{-}C+} \leq z_{H+C+}$, $z_{H\text{-}C+} \geq z_{H\text{-}C\text{-}}$, $z_{H+C+} \geq z_{H+C\text{-}}$. We also find that $z_{H+C+} \to z_{H\text{-}C+}$ as $s \to \infty$. That is, this illustrative treatment exactly recovers the results of our general analysis. In Appendix D, we undertake a more formal "neighbor-modulated fitness" analysis of the illustrative model and obtain identical results to the more intuitive inclusive fitness analysis presented here (Scott & Wild, 2023; Taylor et al., 2007).

Discussion

Kin discrimination can involve one or both of the following: (1) kin-discriminating partner choice, in which individuals choose who to help; (2) kin-discriminating helping, in which individuals choose how much to help a given partner. Previous theory developed by Faria and Gardner (2020) on

the consequences of kin discrimination for the overall level of helping has considered (2) but not (1). We found that kin-discriminating partner choice allows individuals to pick out closer relatives, who elicit a higher level of helping, increasing the overall level of helping relative to when individuals help indiscriminately (Figure 1). In contrast, kin-discriminating helping allows individuals to fine-tune their level of helping to the relatedness of each social partner, and this fine tuning may increase (Figure 1c), decrease (Figure 1a), or leave unchanged (Figure 1b) the overall level of helping relative to when individuals help indiscriminately (Faria & Gardner, 2020). Therefore, if the capacities for kin-discriminating partner choice and helping coincide, then kin discrimination might or might not serve to boost the overall level of helping, but will do so if there is sufficient kin-discriminating partner choice (Figure 2).

The "veil of ignorance" is an idea that has been imported to evolutionary biology from moral philosophy (Frank, 2003; Harsanyi, 1955; Okasha, 2012; Rawls J, 1971). It has been argued that a veil of ignorance can reduce conflict within biological groups (Queller & Strassmann, 2013). This within-group conflict reduction could lead to an increase in the overall level of helping. Specifically, if individuals know how related they are to their social partners they have the opportunity to restrict help to nonrelatives, but if individuals are ignorant with respect to the relatedness of their social partners they cannot identify their nonrelatives and so do not have the opportunity to restrict help. Consequently, a veil of ignorance may be expected to cause ignorant individuals (indiscriminate helpers) to help more than knowledgeable individuals (kin discriminators). For example, in social insects where the queen mates multiple times, colonies comprise different classes of relative (patrilines) and individuals usually can't tell which nest-mates are more closely related (Keller, 1997; Queller & Strassmann, 2002). It has been assumed that this ignorance promotes cooperation within the colony (Queller & Strassmann, 2013).

This veil-of-ignorance thinking in regard to the overall level of helping in biological populations has recently been challenged by Faria and Gardner (2020). They pointed out that, although ignorant individuals do not have the opportunity afforded to informed individuals of identifying and restricting help toward nonrelatives, they also lack the ability to identify and upregulate help toward relatives. Consequently, a veil of ignorance will only increase the overall level of helping if the informed downregulation of help toward nonrelatives is greater than the informed upregulation of help toward relatives (Faria & Gardner, 2020). Faria and Gardner (2020) showed that this will only be the case if the optimal level of helping is a concave (i.e., decelerating) function of relatedness. Consequently, for the veil-of-ignorance argument to be convincing, we would also need an account of why helping should increase concavely with relatedness; no such account has yet been given.

Our work provides a further challenge to veil-of-ignorance thinking in regard to the overall level of helping in biological populations. In nature, individuals who are informed as to their relatedness to social partners (kin discriminators) are often able to choose their social partners, unlike ignorant individuals (indiscriminate helpers). Consequently, informed individuals choose close relatives, with whom they help intensively, tending to increase the overall level of helping relative to ignorant individuals, in opposition to veil-of-ignorance thinking. The reason for this discrepancy is that, in nature but not moral philosophy, lifting the veil of ignorance does not just reveal to individuals their social partners; it also reveals the identities of other potential social partners, whom the individuals may choose to help instead. In other words, in nature, kin discriminators and indiscriminate helpers differ from each other not only with regard to their ignorance over relatedness, but also with regard to whether they can choose their social partners. This extra biological detail (partner choice) means that veil-of-ignorance thinking in moral philosophy is not always directly analogous to veil-of-ignorance thinking in evolutionary biology, so may not tend to work in the same way. A broader point here is that care should be taken when applying theories developed in the social sciences or philosophy to evolutionary biology, as biological complexities may weaken the analogy and lead to opposite predictions; formal modeling can be useful here (Scott & West, 2019). We note that veil-of-ignorance thinking may still be important in biology, even if the overall helping investment is not typically increased. For instance, veils of ignorance over relatedness may reduce within-group conflict or prevent social groups from fragmenting into smaller kin groups (Queller & Strassmann, 2013; Rautiala & Gardner, 2023).

A large theoretical literature since Hamilton (1964) has examined the conditions under which different types of helping and kin discrimination behaviors are evolutionarily stable (i.e., the evolutionary *causes* of kin discrimination) (Bourke, 2011; Frank, 1998; Green et al., 2024; Reeve, 1989; West et al., 2021). Our focus in the present paper is different. Rather than examining the evolutionary stability of kin recognition, we have investigated its impact upon the overall level of helping in the population (i.e., the evolutionary *consequences* of kin discrimination). Future work might consider the consequences of other types of recognition on the overall level of helping. This could include recognition of distant rather than close kin (Kennedy et al., 2021); recognition not of kinship per se, but of something else, like a reputation for being cooperative (Barclay, 2016); specific mechanisms of kin recognition such as phenotype matching, which may be susceptible to errors or biases in estimating relatedness (Krupp & Taylor, 2013; Lehmann & Perrin, 2002).

Future work could also extend our analysis in the following ways. First, it could generalize to scenarios in which the costs and benefits of helping are different in indiscriminate and kin-discriminating helping scenarios. An example of this is if there are synergistic effects of multiple social partners engaging in helping and kin discriminating enables one to identify high-relatedness situations in which this is most likely to occur. Second, it could generalize to scenarios in which the function relating relatedness and the optimal level of helping (Z) is not uniformly convex/concave/linear, but, for example, is sigmoidal or a step function. Third, it could examine how demographic processes (migration rate, group size, etc.) affect the scale of competition and consequently the shape of the function relating relatedness and the optimal level of helping (Z). Fourth, it could examine "harming" rather than "helping" behaviors, specifically how the ability to choose poorly related individuals to harm affects the overall level of harming in the population compared to indiscriminate harming (Gardner & West, 2004, 2010; Krupp & Maciejewski, 2022; Krupp & Taylor, 2015; Lehmann et al., 2009).

Our work emphasizes the importance of kin-discriminating partner choice in promoting helpfulness. This relates to recent

theoretical work showing that partner choice is often required for kin discrimination to evolve in the first place (Scott, 2024; Scott et al., 2022, 2023). Putting these two strands of research together, we might therefore expect that, when we do observe kin discrimination in nature, partner choice is likely to also be present (as a pre-requisite), and therefore kin discrimination is likely to be promoting the level of helpfulness. It would be very useful to test these theoretical predictions empirically. Kin-discriminating partner choice can be empirically distinguished from kin-discriminating helping. Cornwallis et al., (2009) undertook a meta-analysis of kin discrimination in cooperatively breeding vertebrates and found only a weak effect of relatedness in determining the amount of help provided (kin discriminating helping), which they attributed to confounding factors like variation in physical condition influencing the cost of helping. However, Cornwallis et al. (2009) found a strong effect of relatedness in determining the probability of helping (kin-discriminating partner choice). Kin-discriminating partner choice arises in cooperatively breeding vertebrates as a consequence of their mobility, in which individuals move around and encounter multiple potential social partners within their group or larger social network before deciding who to help (Green et al., 2024; Scott et al., 2022).

Data availability

No new data were generated or analyzed in support of this research.

Author contributions

TWS, GW, and AG contributed to conceptualization; TWS, GW, and AG contributed to mathematical modeling; TWS wrote the first draft of the manuscript; TWS, GW, and AG contributed to editing the manuscript.

Funding

For funding, we thank the ERC (834164 & 771387), NERC (NE/V011537/1), and the Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-06626).

Conflict of interest

The authors have no conflicts of interest to declare.

Acknowledgments

We thank D. B. Krupp, B. Kuijper and one anonymous reviewer for comments on the manuscript.

References

Barclay, P. (2016). Biological markets and the effects of partner choice on cooperation and friendship. *Current Opinion in Psychology*, *7*, 33–38. https://doi.org/10.1016/j.copsyc.2015.07.012

Bourke, A. F. G. (2011). *Principles of Social Evolution*. Oxford University Press.

Brown, S. P., & Taylor, P. D. (2010). Joint evolution of multiple social traits: A kin selection analysis. *Proceedings Biological Sciences*, *277*(1680), 415–422. https://doi.org/10.1098/rspb.2009.1480

Cornwallis, C. K., West, S. A., & Griffin, A. S. (2009). Routes to indirect fitness in cooperatively breeding vertebrates: Kin discrimination and limited dispersal. *Journal of Evolutionary Biology*, *22*(12), 2445–2457. https://doi.org/10.1111/j.1420-9101.2009.01853.x

Faria, G. S., & Gardner, A. (2020). Does kin discrimination promote cooperation? *Biology Letters*, *16*(3), 20190742. https://doi.org/10.1098/rsbl.2019.0742

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

Frank, S. A. (2003). Repression of competition and the evolution of cooperation. *Evolution (N Y)*, *57*, 693–705.

Gardner, A., & West, S. A. (2004). Spite and the scale of competition. *Journal of Evolutionary Biology*, *17*(6), 1195–1203. https://doi.org/10.1111/j.1420-9101.2004.00775.x

Gardner, A., & West, S. A. (2010). Greenbeards. *Evolution; International Journal of Organic Evolution*, *64*(1), 25–38. https://doi.org/10.1111/j.1558-5646.2009.00842.x

Green, J. P., Biernaskie, J. M., Mee, M. C., & Leedale, A. E. (2024). The evolution of kin discrimination across the tree of life. *Annual Review of Ecology, Evolution, and Systematics*, *55*(1), 347–367. https://doi.org/10.1146/annurev-ecolsys-102221-051057

Hamilton, W. D. (1963). The evolution of altruistic behavior. *Am Nat*, *97*(896), 354–356. https://doi.org/10.1086/497114.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. I & II. *Journal of Theoretical Biology*, *7*(1), 1–16. https://doi.org/10.1016/0022-5193(64)90038-4

Harsanyi, J. C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy*, *63*(4), 309–321. https://doi.org/10.1086/257678

Jensen, J. L. W. V. (1906). Sur les fonctions convexes et les inégalités entre les valeurs moyennes. *Acta Mathematica*, *30*(0), 175–193. https://doi.org/10.1007/bf02418571

Keller, L. (1997). Indiscriminate altruism: Unduly nice parents and siblings. *Trends in Ecology and Evolution*, *12*(3), 99–103. https://doi.org/10.1016/s0169-5347(96)10065-3

Kennedy, P., Sumner, S., Botha, P., Welton, N. J., Higginson, A. D., & Radford, A. N. (2021). Diminishing returns drive altruists to help extended family. *Nature Ecology and Evolution*, *5*(4), 468–479. https://doi.org/10.1038/s41559-020-01382-z

Krupp, D. B., & Maciejewski, W. (2022). The evolution of extraordinary self-sacrifice. *Scientific Reports*, *12*(1), 90. https://doi.org/10.1038/s41598-021-04192-w

Krupp, D. B., & Taylor, P. D. (2013). Enhanced kin recognition through population estimation. *American Naturalist*, *181*(5), 707–714. https://doi.org/10.1086/670029

Krupp, D. B., & Taylor, P. D. (2015). Social evolution in the shadow of asymmetrical relatedness. *Proceedings Biological Sciences*, *282*(1807), 20150142. https://doi.org/10.1098/rspb.2015.0142

Lehmann, L., Feldman, M. W., & Rousset, F. (2009). On the evolution of harming and recognition in finite panmictic and infinite structured populations. *Evolution; International Journal of Organic Evolution*, *63*(11), 2896–2913. https://doi.org/10.1111/j.1558-5646.2009.00778.x

Lehmann, L., & Perrin, N. (2002). Altruism, dispersal, and phenotype-matching kin recognition. *American Naturalist*, *159*(5), 451–468. https://doi.org/10.1086/339458

Lehmann, L., & Rousset, F. (2010). How life history and demography promote or inhibit the evolution of helping behaviours. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences*, *365*(1553), 2599–2617. https://doi.org/10.1098/rstb.2010.0138

Maynard Smith, J. (1964). Group selection and kin selection. *Nature*, *201*, 1145–1147.

Michod, R. E., & Hamilton, W. D. (1980). Coefficients of relatedness in sociobiology. *Nature*, *288*(5792), 694–697. https://doi.org/10.1038/288694a0

Okasha, S. (2012). Social justice, genomic justice and the veil of ignorance: Harsanyi meets Mendel. *Economics and Philosophy*, *28*(1), 43–71. https://doi.org/10.1017/s0266267112000119

Queller, D. C. (1994). Genetic relatedness in viscous populations. *Evolutionary Ecology*, *8*(1), 70–73. https://doi.org/10.1007/bf01237667

Queller, D. C., & Strassmann, J. E. (2002). The many selves of social insects. *Science (New York, N.Y.)*, *296*(5566), 311–313. https://doi.org/10.1126/science.1070671

Queller, D. C., & Strassmann, J. E. (2013). The veil of ignorance can favour biological cooperation. *Biology Letters*, *9*(6), 20130365. https://doi.org/10.1098/rsbl.2013.0365

Rautiala, P., & Gardner, A. (2023). The geometry of evolutionary conflict. *Proceedings of the Royal Society B: Biological Sciences*, *290*(1992), 20222423. https://doi.org/10.1098/rspb.2022.2423

Rawls J. (1971). *A theory of justice*. Harvard University Press., Cambridge, MA.

Reeve, H. K. (1989). The evolution of conspecific acceptance thresholds. *American Naturalist*, *133*(3), 407–435. https://doi.org/10.1086/284926

Rousset, F. (2004). Genetic structure and selection In *Subdivided Populations (MPB-40)*. Princeton University Press.

Scott, T. W. (2024). Crozier's paradox and kin recognition: Insights from simplified models. *Journal of Theoretical Biology*, *581*, 111735. https://doi.org/10.1016/j.jtbi.2024.111735

Scott, T. W., Grafen, A., & West, S. A. (2022). Multiple social encounters can eliminate Crozier's paradox and stabilise genetic kin recognition. *Nature Communications*, *13*(1), 3902. https://doi.org/10.1038/s41467-022-31545-4

Scott, T. W., Grafen, A., & West, S. A. (2023). Host–parasite coevolution and the stability of genetic kin recognition. *Proceedings of the National Academy of Sciences*, *120*(30), e2220761120. https://doi.org/10.1073/pnas.2220761120

Scott, T. W., & West, S. A. (2019). Adaptation is maintained by the parliament of genes. *Nature Communications*, *10*(1), 5163. https://doi.org/10.1038/s41467-019-13169-3

Scott, T. W., & Wild, G. (2023). How to make an inclusive-fitness model. *Proceedings of the Royal Society B: Biological Sciences*, *290*(2008), 20231310. https://doi.org/10.1098/rspb.2023.1310

Taylor, P. D. (1992a). Altruism in viscous populations-an inclusive fitness model. *Evolutionary Ecology*, *6*(4), 352–356. https://doi.org/10.1007/bf02270971

Taylor, P. D. (1992b). Inclusive fitness in a homogeneous environment. *Proc R Soc Lond B Biol Sci*, *249*, 299–302.

Taylor, P. D., & Frank, S. A. (1996). How to make a kin selection model. *Journal of Theoretical Biology*, *180*(1), 27–37. https://doi.org/10.1006/jtbi.1996.0075

Taylor, P. D., Wild, G., & Gardner, A. (2007). Direct fitness or inclusive fitness: How shall we model kin selection? *Journal of Evolutionary Biology*, *20*(1), 301–309. https://doi.org/10.1111/j.1420-9101.2006.01196.x

West, S. A., Cooper, G. A., Ghoul, M. B., & Griffin, A. S. (2021). Ten recent insights for our understanding of cooperation. *Nature Ecology and Evolution*, *5*(4), 419–430. https://doi.org/10.1038/s41559-020-01384-x

West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, *20*(2), 415–432. https://doi.org/10.1111/j.1420-9101.2006.01258.x

Appendix A

We show in this appendix that, when choosing a partner to help, kin-discriminating partner choice causes individuals to choose partners with higher (rather than lower) relatedness. We assume that an individual has a choice of social partners with varying relatedness (r), and then enacts an amount of helping that is optimal according to the relatedness of the partner they have chosen. The inclusive fitness of the individual (actor) can be written as

$$H = \alpha - C\left(z^*\left(r\right)\right) + rB\left(z^*\left(r\right)\right), \tag{A1}$$

where: α is baseline fitness; $C(z)$ is the cost of helping by an amount z; $B(z)$ is the benefit of receiving an amount of help z; r is the relatedness of actor and recipient; $z^*(r)$ is the optimal level of help as a function of relatedness.

Then, if the individual can choose to effectively increase the relatedness of the social partner to whom she directs her help, she would experience a marginal change in inclusive fitness given by:

$$\frac{dH}{dr} = -C'\left(z^*\left(r\right)\right)\frac{dz^*}{dr} + rB'\left(z^*\left(r\right)\right)\frac{dz^*}{dr} + B\left(z^*\left(r\right)\right) \tag{A2}$$

$$= \frac{dz^*}{dr}\left(-C'\left(z^*\left(r\right)\right) + rB'\left(z^*\left(r\right)\right)\right) + B\left(z^*\left(r\right)\right).$$

By virtue of $z^*(r)$ being the optimal level of help, it satisfies the condition $-C'\left(z^*\left(r\right)\right) + rB'\left(z^*\left(r\right)\right) = 0$. This means that we can simplify Equation A2 to obtain

$$\frac{dH}{dr} = \underbrace{\frac{dz^*}{dr}\left\{-C'\left(z^*\left(r\right)\right) + rB'\left(z^*\left(r\right)\right)\right\}}_{=0} + B\left(z^*\left(r\right)\right) \tag{A3}$$

$$= B\left(z^*\left(r\right)\right),$$

which is positive, so the individual's inclusive fitness is always increased by choosing a social partner of higher relatedness. This means that, if an individual is allowed to encounter more than one individual before choosing her social partner (kin-discriminating partner choice), she will choose the closest relative.

Appendix B

In this appendix, we obtain the results presented in the main text general analysis. An actor samples s individuals from her neighborhood. Let R_i denote the relatedness between the actor and the ith individual sampled, let $\bar{R}$ denote the mean of the sample, and let R_{max} be the largest value sampled. Obviously, $\mathrm{E}[R_{max} - \bar{R}] \geq 0$, where E is an expectation taken with respect to the joint distribution of $\{R_1, \ldots, R_s\}$. It follows that the expected value of R_{max}, $\mathrm{E}[R_{max}]$, is never less than the expected value of the sample mean, $\mathrm{E}[\bar{R}]$. If $\{R_1, \ldots R_s\}$ is an iid sample from a distribution with expectation μ_r, then we can go further and say that $\mathrm{E}[R_{max}]$ is never less than μ_r, because $\mathrm{E}[\bar{R}]$ is an unbiased estimator of μ_r in this case.

Let Z be an increasing function of a single variable. Then we have $\mathrm{E}[Z(R_{max}) - Z(\bar{R})] \geq 0$, which we can re-write as $z_{H+C+} \geq z_{H+C-}$ (in keeping with the main text). We also have $Z(\mathrm{E}[R_{max}]) - Z(\mathrm{E}[\bar{R}]) \geq 0$, which we can re-write as $z_{H-C+} \geq z_{H-C-}$ (in keeping with the main text). If Z is convex, then Jensen's, (1906) inequality says $Z(\mathrm{E}[R_{max}]) \leq \mathrm{E}[Z(R_{max})]$ and $Z(\mathrm{E}[\bar{R}]) \leq \mathrm{E}[Z(\bar{R})]$. The inequality also assures us that the secant line we construct using a convex function (e.g., dashed line in Figure 1c) lies above the graph of that function (e.g., solid curve in Figure 1c). It follows that, for increasing and convex Z, we have

$$Z\left(\mathrm{E}\left[\bar{R}\right]\right) \leq \mathrm{E}\left[Z\left(\bar{R}\right)\right], \; Z\left(\mathrm{E}\left[R_{max}\right]\right) \leq \mathrm{E}\left[Z\left(R_{max}\right)\right]. \tag{B1}$$

Thinking of Z as the optimal level of helping, we re-write the previous line as, $z_{H-C-} \leq z_{H+C-}$, $z_{H-C+} \leq z_{H+C+}$ (in keeping with the main text). When kin discrimination occurs, we evaluate Z before the expectation (the actor decides on a level of help after sampling); when it does not occur, we evaluate Z after the expectation (the actor decides on a level of help before sampling). When kin-discriminating partner choice occurs, we evaluate at R_{max} and when it does not occur we evaluate at $\bar{R}$.

If Z is concave then $Z(\mathrm{E}[R_{\max}]) \geq \mathrm{E}[Z(R_{\max})]$ and $Z(\mathrm{E}[\bar{R}]) \geq \mathrm{E}[Z(\bar{R})]$, and we have

$$E\left[Z\left(\bar{R}\right)\right] \leq Z\left(\mathrm{E}\left[\bar{R}\right]\right), \; E\left[Z\left(R_{max}\right)\right] \leq Z\left(\mathrm{E}\left[R_{max}\right]\right). \tag{B2}$$

This can be re-written as $z_{\text{H+C-}} \leq z_{\text{H-C-}}$, $z_{\text{H+C+}} \leq z_{\text{H-C+}}$ (in keeping with the main text).

If Z is linear, then $Z(\mathrm{E}[R_{\max}]) = \mathrm{E}[Z(R_{\max})]$ and $Z(\mathrm{E}[\bar{R}]) = \mathrm{E}[Z(\bar{R})]$, and we have

$$E\left[Z\left(\bar{R}\right)\right] = Z\left(\mathrm{E}\left[\bar{R}\right]\right) \leq Z\left(\mathrm{E}\left[R_{max}\right]\right) = E\left[Z\left(R_{max}\right)\right], \tag{B3}$$

or simply, $z_{\text{H-C-}} = z_{\text{H+C-}} \leq z_{\text{H-C+}} = z_{\text{H+C+}}$ as in the main text.

Appendix C

In this appendix we show that, as partner choice becomes more perfect, kin-discriminating helping has less effect on the overall level of helping, with no effect in the limit of perfect partner choice.

We collect a sample of s independent and identically distributed observations. Let $F(x)$ be the cumulative distribution function associated with the probability distribution from which a single observation has been taken. We will assume that there is a number r_1 such that $F(r_1) = 1$ and $F(r) < 1$ for any $r < r_1$. In other words, we will assume that there is a maximum attainable value, r_1, that any given observation may achieve.

Proposition 1. *Let R_{max} be the largest observation in the sample. The expectation of R_{max} can be made arbitrarily close to r_1 by choosing a sufficiently large sample size, s.*

Proof. The proof follows from combining two separate observations. We first observe that given any ε such that $0 < \varepsilon < r_1$, the quantity

$$0 \cdot P\left\{R_{\max} \leq r_1 - \frac{\varepsilon}{2}\right\} + \left(r_1 - \frac{\varepsilon}{2}\right) P\left\{R_{\max} > r_1 - \frac{\varepsilon}{2}\right\}$$

cannot overestimate the expectation of $R_{\max}$. In other words,

$$\left(r_1 - \tfrac{\varepsilon}{2}\right)\left(1 - F_s\left(r_1 - \tfrac{\varepsilon}{2}\right)\right) \leq E\left[R_{max}\right] \tag{C1}$$

where $F_s(r_1)$ is the cdf associated with $R_{\max}$. Second, we observe that the quantity $F_s\left(r_1 - \frac{\varepsilon}{2}\right) = \left[F\left(r_1 - \frac{\varepsilon}{2}\right)\right]^s$ tends to zero—in a pointwise fashion—in the limit of large s. Thus, there exists an integer s_1 such that

$$F_s\left(r_1 - \frac{\varepsilon}{2}\right) \leq \frac{\varepsilon}{2r_1} < \frac{\varepsilon}{2}\frac{1}{r_1 - \varepsilon}$$

for all $s > s_1$. It follows that

$$1 - F_s\left(r_1 - \tfrac{\varepsilon}{2}\right) > 1 - \tfrac{\varepsilon}{2}\tfrac{1}{r_1 - \varepsilon/2} \tag{C2}$$

whenever $s > s_1$. Combining inequalities (C1) and (C2), we conclude that for sufficiently large s,

$$\left(r_1 - \frac{\varepsilon}{2}\right)\left(1 - \frac{\varepsilon}{2}\frac{1}{r_1 - \varepsilon/2}\right) < \mathrm{E}\left[R_{\max}\right]$$

which simplifies to $0 < r_1 - \mathrm{E}[R_{max}] < \varepsilon$, as required.□

Corollary to Proposition 1. *For continuous Z, the quantity $Z(\mathrm{E}[R_{\max}])$ can be made arbitrarily close to $Z(r_1)$ by choosing a sufficiently large sample size, s.*

We suppose that $Z(r)$ is a strictly increasing function of r, with $Z(0) = 0$. We suppose further that Z is differentiable on $(0, r_1]$ with $|Z'(r)| \leq M$, where it is understood that $Z'(r_1)$ is calculated using a limit from below. [Alternatively, we could consider a smooth extension of Z that expands the domain to include some $r_2 > r_1$. Then we would assume that, for the extended function, $|Z(r)| \leq M$ for all r in an open interval containing r_1.] For later use, we note that for a sufficiently small $\varepsilon > 0$ we also have $\left|Z'\left(r_1 - \frac{\varepsilon}{M}\right)\right| < M$. Moreover, from Taylor's Remainder Theorem we have

$$\left|Z\left(r_1 - \tfrac{\varepsilon}{M}\right) - Z\left(r_1\right)\right| \leq M\tfrac{\varepsilon}{M} = \varepsilon. \tag{C3}$$

The inequality in (C3) follows from a technical argument that allows us to bound the error associated with approximating a function (namely, $Z(r_1 - \varepsilon/M)$) with its zeroth degree Taylor polynomial centered nearby (namely $Z(r_1)$). Given that Z is increasing the previous line (C3) simplifies to $Z(r_1) - Z\left(r_1 - \frac{\varepsilon}{M}\right) \leq \varepsilon$ or

$$Z\left(r_1\right) - \varepsilon \leq Z\left(r_1 - \tfrac{\varepsilon}{M}\right). \tag{C4}$$

Proposition 2. *If Z has the properties described above, then the expectation of $Z(R_{\max})$ can be made arbitrarily close to $Z(r_1)$ by choosing a sufficiently large sample size, s.*

Proof. We treat (small) $\varepsilon > 0$ as given and observe that, because Z is a strictly increasing function, the quantity

$$Z(0) \cdot P\left\{R_{\max} \leq r_1 - \frac{\varepsilon}{2M}\right\} + Z\left(r_1 - \frac{\varepsilon}{2M}\right) P\left\{R_{\max} > r_1 - \frac{\varepsilon}{2M}\right\}$$

cannot overestimate the expectation of $Z(R_{\max})$. Using the fact that $Z(0) = 0$, this observation becomes

$$Z\left(r_1 - \tfrac{\varepsilon}{2M}\right)\left(1 - F_s\left(r_1 - \tfrac{\varepsilon}{2M}\right)\right) \leq E\left[Z\left(R_{\max}\right)\right]. \tag{C5}$$

If we incorporate the information given in (C4) into (C5), we obtain

$$\left(Z\left(r_1\right) - \tfrac{\varepsilon}{2}\right)\left(1 - F_s\left(r_1 - \tfrac{\varepsilon}{2M}\right)\right) \leq E\left[Z\left(R_{\max}\right)\right] \tag{C6}$$

As we observed in the proof to Proposition 1, $F_s\left(r_1 - \frac{\varepsilon}{2M}\right)$ converges pointwise to 0 in the limit of large sample size s. Therefore we can find a sample size s that is large enough to guarantee

$$F_s\left(r_1 - \frac{\varepsilon}{2M}\right) < \frac{\varepsilon/2}{Z\left(r_1\right)} < \frac{\varepsilon/2}{Z\left(r_1\right) - \frac{\varepsilon}{2}},$$

or equivalently

$$1 - F_s\left(r_1 - \tfrac{\varepsilon}{M}\right) > 1 - \tfrac{\varepsilon/2}{Z(r_1) - \frac{\varepsilon}{2}}. \tag{C7}$$

Combining inequalities (C6) and (C7), we obtain

$$\left(Z\left(r_1\right) - \tfrac{\varepsilon}{2}\right)\left(1 - \tfrac{\varepsilon/2}{Z(r_1) - \frac{\varepsilon}{2}}\right) \leq E\left[Z\left(R_{\max}\right)\right], \tag{C8}$$

which simplifies to $Z(r_1) - \varepsilon \leq \mathrm{E}[Z(R_{\max})]$, as required.□

Appendix D

In this appendix, we analyze the illustrative model presented in the main text using a "neighbor-modulated fitness" rather than an "inclusive fitness" approach (Scott & Wild, 2023; Taylor et al., 2007). A focal individual (actor) invests x_{ij} into helping, where the i subscript indicates whether the focal individual is interacting with a sibling or a nonrelative ($i \in \{sib, non\}$) and the j subscript indicates whether individuals are capable of kin-discriminating helping or help indiscriminately ($j \in \{disc, indisc\}$). We assume that

kin-discriminating helpers ($j = disc$) can independently optimize their helping investments toward siblings ($x_{sib,disc}$) and nonrelatives ($x_{non,disc}$); in other words, they are modeled as two independent, separate traits, each with their own genetic basis.

Our lifecycle assumptions mean that each individual is chosen to be someone's recipient, in expectation, one time per generation. A focal individual therefore receives an expected overall helping investment of y_{ij}, where y_{ij} is the average helping investment amongst individuals who choose the focal individual as their recipient. We denote the population average helping investment by z_{ij}.

Consequently, the focal individual incurs a cost of helping $C\left(x_{ij}\right)$ and receives a benefit of helping $B\left(y_{ij}\right)$. We assume the following functional forms: $C\left(x_{ij}\right) = \left(\frac{\beta}{1+\beta}\right) x_{ij}^{\left(\frac{1+\beta}{\beta}\right)}$ & $B\left(y_{ij}\right) = y_{ij}$.

The focal individual's survival probability will be proportional to the following (recall that $1 - \left(\frac{n-1}{n}\right)^s$ gives the probability of helping, or equivalently being helped by, a sibling, and $\left(\frac{n-1}{n}\right)^s$ gives this probability for a non-sibling):

$$v_j = \left(1 - \left(\frac{n-1}{n}\right)^s\right)\left(1 - \left(\frac{\beta}{1+\beta}\right) x_{sib,j}^{\left(\frac{\beta}{1+\beta}\right)} + y_{sib,j}\right) + \left(\frac{n-1}{n}\right)^s \left(1 - \left(\frac{\beta}{1+\beta}\right) x_{non,j}^{\left(\frac{\beta}{1+\beta}\right)} + y_{non,j}\right). \quad \text{(D1)}$$

Substituting z_{ij} in for x_{ij} and y_{ij} in Equation D1, we can obtain the survival probability for a random individual drawn from the population. It is proportional to

$$\bar{v}_j = \left(1 - \left(\frac{n-1}{n}\right)^s\right)\left(1 - \left(\frac{\beta}{1+\beta}\right) z_{sib,j}^{\left(\frac{\beta}{1+\beta}\right)} + z_{sib,j}\right) + \left(\frac{n-1}{n}\right)^s \left(1 - \left(\frac{\beta}{1+\beta}\right) z_{non,j}^{\left(\frac{\beta}{1+\beta}\right)} + z_{non,j}\right). \quad \text{(D2)}$$

The focal individual's neighbor-modulated fitness, i.e., her number of offspring after one full iteration of the lifecycle, is then obtained as

$$w_j = \frac{v_j}{\bar{v}_j}. \quad \text{(D3)}$$

The equilibrium levels of helping enacted toward siblings ($z^*_{sib,j}$) and nonrelatives ($z^*_{non,j}$) will satisfy the following conditions:

$$\left.\frac{\mathrm{d}w}{\mathrm{d}x_{sib,j}}\right|_{x_{sib,j},y_{sib,j},z_{sib,j}=z^*_{sib,j}} = 0,$$

$$\left.\frac{\mathrm{d}w}{\mathrm{d}x_{non,j}}\right|_{x_{non,j},y_{non,j},z_{non,j}=z^*_{non,j}} = 0. \quad \text{(D4)}$$

These conditions can be written equivalently as follows using the chain rule, where each derivative and partial derivative below is evaluated at the equilibrium $x_{ij}, y_{ij}, z_{ij} = z^*_{ij}$:

$$\frac{\partial w}{\partial x_{sib,j}} + \frac{\partial w}{\partial y_{sib,j}} \frac{\mathrm{d}y_{sib,j}}{\mathrm{d}x_{sib,j}} = 0,$$

$$\frac{\partial w}{\partial x_{non,j}} + \frac{\partial w}{\partial y_{non,j}} \frac{\mathrm{d}y_{non,j}}{\mathrm{d}x_{non,j}} = 0. \quad \text{(D5)}$$

(Brown & Taylor, 2010; Frank, 1998; Lehmann & Rousset, 2010; Rousset, 2004; Taylor & Frank, 1996). To evaluate these conditions, we first note that $\frac{\mathrm{d}y_{sib,j}}{\mathrm{d}x_{sib,j}}$ and $\frac{\mathrm{d}y_{non,j}}{\mathrm{d}x_{non,j}}$ give the marginal change in partner trait value with a change in actor trait value, equivalent to the coefficient of relatedness between the actor and its social partner (r_{ij}) (Frank, 1998; Rousset, 2004; Taylor & Frank, 1996). We can therefore evaluate $\frac{\mathrm{d}y_{sib}}{\mathrm{d}x_{sib}} = r_{sib,j}$ and $\frac{\mathrm{d}y_{non}}{\mathrm{d}x_{non}} = r_{non,j}$. We obtain

$$\frac{\partial w}{\partial x_{sib,j}} + \frac{\partial w}{\partial y_{sib,j}} r_{sib,j} = 0,$$

$$\frac{\partial w}{\partial x_{non,j}} + \frac{\partial w}{\partial y_{non,j}} r_{non,j} = 0. \quad \text{(D6)}$$

Next, we can obtain the partial derivatives $\frac{\partial w}{\partial x_{sib,j}}$, $\frac{\partial w}{\partial y_{sib,j}}$, $\frac{\partial w}{\partial x_{non,j}}$ & $\frac{\partial w}{\partial y_{non,j}}$ by differentiating Equation D3 with respect to $x_{sib,j}$, $y_{sib,j}$, $x_{non,j}$ & $y_{non,j}$ respectively. Evaluating everything at the equilibrium $x_{ij}, y_{ij}, z_{ij} = z^*_{ij}$, we obtain the equilibrium levels of helping toward siblings and nonrelatives, respectively, as

$$z^*_{sib,j} = {r_{sib,j}}^{\beta},$$

$$z^*_{non,j} = {r_{non,j}}^{\beta}. \quad \text{(D7)}$$

Therefore, the β parameter solely determines whether the optimal level of helping increases linearly ($\beta = 1$), convexly ($\beta > 1$) or concavely ($\beta < 1$) with relatedness.

In the "kin-discriminating helping" scenario ($j = disc$), individuals know how related they are to each of their social partners, and can adjust their level of help accordingly, such that $r_{sib,disc} = \frac{1}{2}$ and $r_{non,disc} = 0$. Substituting these into Equation D7, we obtain the equilibrium levels of helping toward siblings and nonrelatives in the "kin-discriminating helping" scenario as

$$z^*_{sib,disc} = \left(\frac{1}{2}\right)^{\beta},$$

$$z^*_{non,disc} = 0. \quad \text{(D8)}$$

Helping nonrelatives is disfavored ($z^*_{non,disc} = 0$) because it provides no indirect or direct fitness benefits to the actor.

The average level of helping in the "kin-discriminating helping" scenario, $\bar{z}^*_{disc}$, is calculated by summing the weighted levels of helping toward siblings and nonrelatives, where the weights are probabilities of interaction (i.e., by taking $z^*_{sib,disc}\left(1 - \left(\frac{n-1}{n}\right)^s\right) + z^*_{non,disc}\left(\frac{n-1}{n}\right)^s$). We obtain

$$\bar{z}^*_{disc} = \left(\tfrac{1}{2}\right)^{\beta}\left(1 - \left(\tfrac{n-1}{n}\right)^s\right). \quad \text{(D9)}$$

In the "no kin-discriminating helping" scenario ($j = indisc$), in which individuals help indiscriminately, individuals must commit to a given level of help before they have observed whether their social partner is a sibling or nonrelative. This level of help will be optimized with respect to the average relatedness between social partners, given by $r_{sib,indisc} = r_{non,indisc} = \frac{1}{2}\left(1 - \left(\frac{n-1}{n}\right)^s\right)$. Substituting these into Equation D7, we obtain the equilibrium level of helping (toward siblings, nonrelatives, and on average) in the "no kin-discriminating helping" scenario as

$$\bar{z}^*_{indisc} = z^*_{sib,indisc} = z^*_{non,indisc} = \left(\tfrac{1}{2}\left(1 - \left(\tfrac{n-1}{n}\right)^s\right)\right)^{\beta} \quad \text{(D10)}$$

Note that, as s tends to infinity, $\bar{z}^*_{indisc}$ (Equation D10) and $\bar{z}^*_{disc}$ (Equation D9) converge on $z^*_{sib,disc}$ (Equation D8). That is, with perfect partner choice, kin discriminators pick out relatives with certainty, meaning the equilibrium level of helping becomes the optimal level of helping toward relatives, and this is true irrespective of whether individuals are capable of kin-discriminating helping.

First, we ask when kin-discriminating partner choice per se increases the overall level of helping (i.e., holding fixed

the level of kin-discriminating helping). Kin-discriminating partner choice is encapsulated by the s parameter, with $s = 1$ implying no kin-discriminating partner choice, and greater s implying greater kin-discriminating partner choice. Using Equations D9 and D10, we can observe that $\bar{z}^*_{disc}$ and $\bar{z}^*_{indisc}$ are increasing functions of s. Therefore, kin-discriminating partner choice per se increases the overall level of helping in the population, and this is true irrespective of whether the population exhibits kin-discriminating helping.

Second, we ask when kin-discriminating helping per se increases the overall level of helping (i.e., holding fixed the level of kin-discriminating partner choice). We examine this by comparing the values of $\bar{z}^*_{disc}$ (Equation D9) and $\bar{z}^*_{indisc}$ (Equation D10). We recover Faria and Gardner's result, finding that: $\bar{z}^*_{disc} > \bar{z}^*_{indisc}$ whenever $\beta > 1$; $\bar{z}^*_{disc} < \bar{z}^*_{indisc}$ whenever $\beta < 1$; $\bar{z}^*_{disc} = \bar{z}^*_{indisc}$ whenever $\beta = 1$. This implies that kin-discriminating helping: increases the overall level of helping when the optimal level of help increases convexly with relatedness ($\beta > 1$); decreases the overall level of helping when the optimal level of help increases concavely with relatedness ($\beta < 1$); neither increases nor decreases the overall level of helping when the optimal level of help increases linearly with relatedness ($\beta = 1$).

Finally, we ask, when the capacities for kin-discriminating helping and kin-discriminating partner choice coincide, when does this increase the overall level of helping? In other words, when is a population exhibiting both kin-discriminating helping and kin-discriminating partner choice more helpful than a population exhibiting neither? We examine this by evaluating $\bar{z}^*_{disc} > \bar{z}^*_{indisc}|_{s=1}$ using Equations D9 and D10. We obtain the following condition for when kin discrimination increases the overall level of helping:

$$\left(\tfrac{1}{n}\right)^{\beta} < 1 - \left(\tfrac{n-1}{n}\right)^{s}. \tag{D11}$$

We can see that this condition is more likely to be satisfied for higher β and s. That is, kin discrimination is more likely to increase the overall level of helping when there is more partner choice and when the optimal level of help increases more convexly with relatedness. We can also see that this condition is not always satisfied. Specifically, kin-discriminating partner choice has a helpfulness-boosting effect, which increases with s. Kin-discriminating helping may have a helpfulness-boosting effect, which increases as β increases above 1, or a helpfulness-inhibiting effect, which increases as β decreases below 1. Kin discrimination only increases the overall level of helping if there is a net helpfulness-boosting effect.

We also highlight that, although kin-discriminating helping and kin-discriminating partner choice may both affect whether the overall level of helping is increased or decreased by kin discrimination, when these capacities coincide, the effect of kin-discriminating partner choice dominates. To see this, note that, as kin-discriminating partner choice (s) increases from 1 to infinity, the curvature of returns from helping (β) goes from fully determining whether Equation D11 is satisfied ($\beta > 1$) or not ($\beta \leq 1$), to having no influence over it at all. This result makes sense because, for increased kin-discriminating partner choice, the relatedness between kin discriminators and their social partners becomes less variable (i.e., closer relatives are picked out more reliably). Consequently, for increased kin-discriminating partner choice, kin discriminators invest less variably in helping across social interactions (i.e., investment is high more reliably). Therefore, for increased kin-discriminating partner choice, the kin-discriminating helping component of kin discrimination becomes less phenotypically visible, minimizing the effect of kin-discriminating helping on the overall level of helping.

Note that this neighbor-modulated analysis quantitatively recovers the results of the inclusive fitness analysis presented in the main text. Specifically note that, aside from the difference in notation, Equation 4 in the main text is equivalent to Equation D10 in this Appendix; Equation 5 in the main text is equivalent to Equation D9 in this Appendix.