9-2020

Pathways to Social Evolution and Their Evolutionary Feedbacks

Yimen G. Araya-Ajoy
Norwegian University of Science and Technology, Norway

David F. Westneat
University of Kentucky, david.westneat@uky.edu

Jonathan Wright
Norwegian University of Science and Technology, Norway

Repository Citation

Araya-Ajoy, Yimen G.; Westneat, David F.; and Wright, Jonathan, "Pathways to Social Evolution and Their Evolutionary Feedbacks" (2020). *Biology Faculty Publications*. 224.
https://uknowledge.uky.edu/biology_facpub/224

Pathways to Social Evolution and Their Evolutionary Feedbacks

Digital Object Identifier (DOI)
https://doi.org/10.1111/evo.14054

Notes/Citation Information
Published in *Evolution*, v. 74, issue 9.

doi:10.1111/evo.14054

SSE SOCIETY for the STUDY of EVOLUTION
EVOLUTION INTERNATIONAL JOURNAL OF ORGANIC EVOLUTION

Pathways to social evolution and their evolutionary feedbacks

Yimen G. Araya-Ajoy,[1,2] **David F. Westneat,**[3] **and Jonathan Wright**[1]

[1] *Centre for Biodiversity Dynamics (CBD), Department of Biology, Norwegian University of Science and Technology (NTNU), Trondheim N-7491, Norway*

[2] *E-mail: yimencr@gmail.com*

[3] *Department of Biology, 101 Morgan Building, University of Kentucky, Lexington, KY 40506-0225, USA*

Received November 2, 2019
Accepted June 27, 2020

In the context of social evolution, the ecological drivers of selection are the phenotypes of other individuals. The social environment can thus evolve, potentially changing the adaptive value for different social strategies. Different branches of evolutionary biology have traditionally focused on different aspects of these feedbacks. Here, we synthesize behavioral ecology theory concerning evolutionarily stable strategies when fitness is frequency dependent with quantitative genetic models providing statistical descriptions of evolutionary responses to social selection. Using path analyses, we review how social interactions influence the strength of selection and how social responsiveness, social impact, and non-random social assortment affect responses to social selection. We then detail how the frequency-dependent nature of social interactions fits into this framework and how it imposes selection on traits mediating social responsiveness, social impact, and social assortment, further affecting evolutionary dynamics. Throughout, we discuss the parameters in quantitative genetics models of social evolution from a behavioral ecology perspective and identify their statistical counterparts in empirical studies. This integration of behavioral ecology and quantitative genetic perspectives should lead to greater clarity in the generation of hypotheses and more focused empirical research regarding evolutionary pathways and feedbacks inherent in specific social interactions.

KEY WORDS: **Frequency dependence, game theory, genetic relatedness, indirect genetic effects, kin selection, quantitative genetics, social behavior.**

Competition, cooperation, communication, and sexual reproduction provide the ecological context for powerful evolutionary forces (West-Eberhard 1979; Krebs and Davies 1991; Westneat and Fox 2010). Social evolution is concerned with the evolution of phenotypes mediating these interactions and the impact they have on the fitness of individuals other than the actor (Frank 1998). A distinguishing feature of social traits is that the ecological environment in which such traits operate includes the phenotypes of other individuals. Thus, the social part of the environment is underpinned by genes and can evolve in response to selection (Kirkpatrick and Lande 1989; Wolf et al. 1998; West-Eberhard 1979). Furthermore, the optimal phenotype for an individual depends in part on the genetic and phenotypic characteristics of the individuals with whom it interacts (Maynard-Smith and Price 1973; McNamara and Weissing 2010). This implies that social evolution creates a feedback between ecological and evolutionary processes, because the evolving phenotypes are also the ecological drivers of selection.

Different areas of evolutionary biology have used different approaches to study evolution in a social context. Behavioral ecology has explored the economics of social evolution using cost-benefit analyses (see Westneat and Fox 2010). One of the biggest contributions was providing an adaptive explanation to seemingly maladaptive "altruistic behaviors" through the process of kin selection (Hamilton 1964). In addition, game theory (e.g., Maynard-Smith and Price 1973) extended adaptive explanations to a broader array of social contexts, allowing predictions for evolutionary stable strategies (ESS) in frequency-dependent scenarios where the fitness effects of a social trait depend on the traits of social partners (Maynard Smith 1982; McNamara and

Weissing 2010). Models developed by quantitative geneticists have also greatly contributed to the study of social behavior, allowing the quantification of how social interactions generate variation in fitness in the form of social selection gradients (Wolf et al. 1999; Okasha 2006; Queller 1992b), and how patterns of inheritance are affected by individuals plastically adjusting to their social environment in the form of indirect genetic effects (IGEs; Moore et al. 1997). Importantly, quantitative genetic models of social evolution provide an approach for the study of responses to selection when related and un-related individuals interact, and directly connect statistical parameters to their theoretical counterparts (Frank 1998; Wolf et al. 1999; McGlothlin et al. 2010, 2014).

Quantitative genetic and behavioral ecology approaches together provide a general framework for studying social evolution. While behavioral ecology theory provides explanations regarding the ecology of long-term selection resulting in evolutionarily stable social strategies (Krebs and Davies 1991; Westneat and Fox 2010), quantitative genetic models can be used to derive short-term statistical descriptions of evolutionary responses to separate selection events (Frank 1997; Wolf et al. 1999). Despite their complementary aims, game theory and quantitative genetic models of social evolution differ in how they deal with the dynamics of evolutionary changes in the social environment feeding back to alter selection (Heino et al. 1998; Lion 2018). Such feedback is a result of the frequency dependence inherent in game theory (McGill and Brown 2007) and are essential for understanding long-term evolution. However, the role of these feedbacks is not made explicit when predicting short-term responses to selection on an event-by-event basis using quantitative genetic models (Heino et al. 1998; Lion 2018). This difference may have resulted in the frequency-dependent processes causing evolutionary feedbacks to be partly overlooked in social evolution studies based on the indirect genetic effects framework (Westneat 2012). A more explicit exploration of the connections between short-term predictions of the responses to selection with long-term expectations concerning evolutionarily stable strategies is needed to better understand the feedbacks caused by social evolution.

In this paper, we detail how behavioral ecology and quantitative genetics together make it possible to address four key features necessary to understand the evolutionary dynamics of social traits: (1) the optimal phenotype for an individual depends on the social environment; (2) the social environment has genes and may evolve; (3) how individuals choose, impact, or respond to their social environments can affect selection; and (4) how individuals choose, impact, or respond to their social environment can also evolve, further affecting the evolutionary dynamics of social traits. We first describe the sources of selection and the evolutionary feedback caused by interacting social strategies. We then review some theoretical models describing how social responsiveness, social impact, and non-random social assortment can each affect responses to social selection (Queller 1992b; Moore et al. 1997; Wolf et al. 1999; McGlothlin et al. 2010). Building upon this framework, we then detail how interactive fitness effects of the traits of social partners fit into this framework and create selection on the ability to choose, impact, or respond to social environments, further altering the evolutionary dynamics of social traits.

Understanding the Ecology of Selection in a Social Context

Quantitative genetic models of social evolution define the statistical parameters that need to be quantified in empirical studies of social evolution (Table 1). The Price equation describes how the phenotypic change in the mean trait of a population occurs through processes generating a covariance between phenotype and fitness (Robertson 1966; Price 1970). The role of social interactions in the evolutionary process similarly involves the covariance between phenotype and fitness, and it includes how the fitness of an individual is affected by its own phenotype as well as by the phenotype(s) of the social partner(s). Queller (1985, 1992b) described in regression form how social interactions may affect fitness, and the explicit link to the multiple regression approach of Lande and Arnold (1983) was then pointed out by Frank (1997). Following the notation of more recent treatments of the topic (e.g., Wolf et al. 1999), a statistical description of how social interactions generate variation in fitness is:

$$\boldsymbol{w} = \alpha + \beta_N \mathbf{z} + \beta_S \mathbf{z}' + \beta_I \mathbf{z}\mathbf{z}' + \boldsymbol{\varepsilon}, \qquad (1)$$

here, $\boldsymbol{w}$ represents the relative fitness of focal individuals and z is the phenotypic trait of those individuals. α is a constant representing the intercept in a statistical model and ε represents all the un-modeled factors affecting the fitness of individuals. β_N is the “non-social” selection gradient relating the relative fitness ($\boldsymbol{w}$) of individuals to their phenotype z (Wolf et al. 1999). We denote the relevant trait of social partner(s) as z' and this trait can be the same or different to the trait of the focal individual under study. Importantly, z' can represent the expression of a trait by a social partner in one instance, the mean social trait over time of a specific partner, or the mean social trait value for a social group (Heisler and Damuth 1987; Svensson and Connallon 2019). The term β_S represents the effect of the phenotype (z') of the social partner(s) or group on the fitness of the focal individuals ($\boldsymbol{w}$) while controlling for the effect of their own phenotype. The effect of one social partner is referred to as the social selection gradient in a neighbor-modulated approach of selection (Frank 1998; Wolf et al. 1999; Okasha 2006; Queller 1992a). The neighbor-modulated approach

Table 1. **Notation for the parameters used in the social evolution models and their biological interpretation, plus references to previous use and possible alternative notation in the literature.**

Symbol	Meaning
$\mathbf{w}$, $\mathbf{w}'$	Relative fitness of focal individual ($\mathbf{w}$) and social partner ($\mathbf{w}'$).
$\mathbf{z}$, $\mathbf{z}'$	Phenotypic trait of focal individual ($\mathbf{z}$) and social partner ($\mathbf{z}'$).
$\mathbf{a}$, $\mathbf{a}'$	Breeding values of focal individual ($\mathbf{a}$) and social partner phenotypes ($\mathbf{a}'$).
$\mathbf{e}$, $\mathbf{e}'$	Environmental effects on the phenotype of a focal individual ($\mathbf{e}$) and its social partners ($\mathbf{e}'$).
$\mathbf{zz}'$	Aggregate phenotype from the interaction between $\mathbf{z}$ and $\mathbf{z}'$ (P_xP_y in Queller 1992b).
$C_{ZZ'}$	Covariance between traits of interacting individuals (C_{ij} in Wolf et al. 1999).
C_{AZ}	Covariance between breeding values and focal trait values - i.e. additive genetic variance (McGlothlin et al. 2010) and $Cov(G_x,\ P_x)$ in Queller (1992b).
$C_{AZ'}$	Covariance between breeding values of the focal individual and trait values of social partners (McGlothlin et al. 2010; and $Cov(G_x,\ P_y)$ in Queller 1992b).
$C_{AZZ'}$	Covariance between breeding values of focal individual and the aggregate property $zz\prime$ ($Cov(G_x,\ P_xP_y)$ in Queller 1992b).
$\psi_{z\prime z}$	Social responsiveness of focal individuals to their social environment (ψ in Moore et al. 1997; and $\beta_{G_x,\ P_Y}$ in Queller 1992b).
$\psi_{zz\prime}$	Social impact of focal individuals on its social environment ($\beta_{P'P}$ in Queller 2011).
r	Genetic relatedness between the focal individual and its social environment (Queller 1992a).
β_N	Non-social selection gradient (Wolf et al. 1999; see β_C in Heisler & Damuth 1987; $\beta_{W_x,\ P_x}$ in Queller 1992b; and $\beta_{wz.y}$ in Frank 1997).
β_S	Social selection gradient (Wolf et al. 1999; see β_C in Heisler & Damuth 1987; $\beta_{W_Y,\ P_y}$ in Queller 1992b; and $\beta_{wy.z}$ in Frank 1997).
β_I	Interaction coefficient (Wolf et al. 1999; see $\beta_{W_{XY},\ P_X P_y}$in Queller 1992b).
$\Delta\bar{z}$	Phenotypic change from one generation to the next (Lande 1979).
s	Selection differential (Lande 1979).

can be extended to study the effect of a group of individuals by fitting the mean phenotype of a social group excluding the focal individual (Wolf et al. 1999; Bijma and Wade 2008), while including the focal individual in the estimation of the mean group phenotype results in the group selection coefficient in a contextual analysis of selection (Heisler and Damuth 1987; Goodnight et al. 1992). The term β_I describes how the relationship between trait and fitness depends upon the social environment in which it is expressed (Queller 1985, 1992b, 2011). The interactive fitness effect coefficient is estimated as a statistical interaction using the product of the interacting traits ($\mathbf{zz}'$). We can visualize the fitness consequences of social interactions as a fitness surface (Fig. 1), where the interactive fitness effects coefficient (β_I) causes the surface to be warped.

The term $\beta_I\mathbf{zz}'$ describes the circumstance in game theoretic models where fitness of one strategy depends on the strategy of the opponent, or its social group (Queller 1984, 1992b). Studying how the fitness of a phenotype depends on mean phenotype of the population ($\bar{\mathbf{z}}'$) is analogous to what game theorists term "playing the field" (p. 23, Maynard Smith 1982), where individuals are assumed to always interact with the entire population or with a sample of the population with representative probability. This also represents a scenario of frequency-dependent selection, where the fitness of a phenotype depends on its frequency in the population (Heino et al. 1998). There are two key empirical considerations to study frequency-dependent selection. (1) It is not possible to study how frequency-dependent selection affects the evolutionary trajectory of a social trait by focusing on single episodes of selection in isolation. It requires analyzing together several selection events (e.g., breeding seasons) and quantifying how the relationship between phenotypes and fitness is affected by changes in the mean phenotype of the population. (2) Mean centring for each episode of selection should be avoided, because it results in the mean phenotype within each episode of selection always equalling zero, causing this type of frequency-dependent selection to be completely overlooked. Importantly, when there are interactive fitness effects and they are overlooked, they will be subsumed into the direct and social selection gradient estimates. Therefore, unexplained fluctuations in estimates of the direct and social selection gradients for single episodes of selection may arise due to un-modeled interactive fitness effects. Including the interaction coefficient in an analysis of selection defines the direct and social selection gradients as the relationships between the focal and social phenotype when the mean focal and social phenotype values are zero. If the mean focal and social phenotype across all the selection events are centered around zero (and not per selection event), then the direct and social selection gradients can be interpreted as the effect of the focal and social

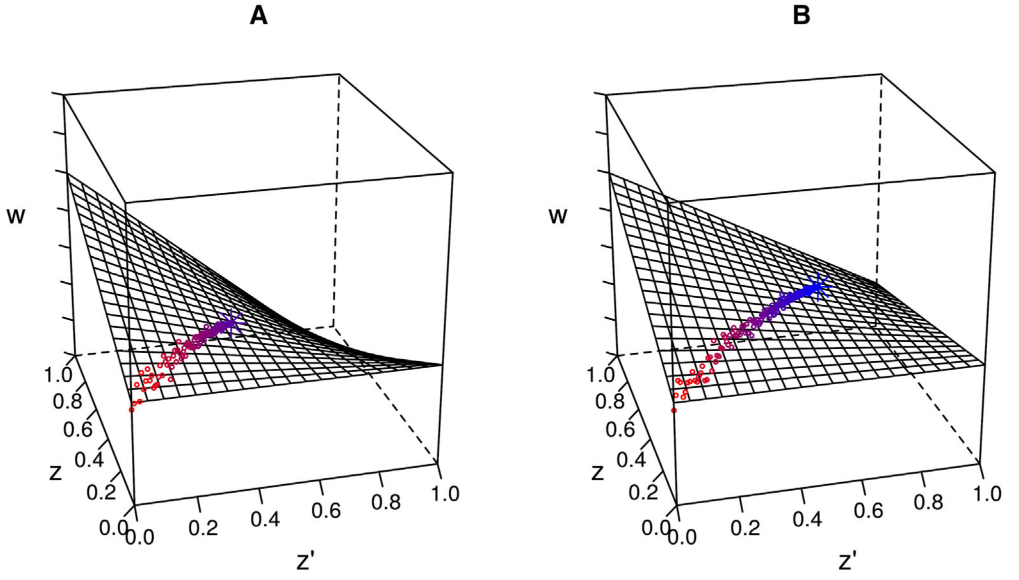

Figure 1. Fitness surfaces representing how the fitness of the focal phenotype is a function of its social environment. The curvature is determined by the interactive fitnesss coeffcient (β_I) .The circles depict the evolutionary dynamics of aggressiveness in two simulated scenarios of negative frequency dependence, where individuals interact at random and assuming perfect inheritance as in standard game theory. Aggressiveness increases fitness when the average aggressiveness in the population is close to zero (β_N= 0.01) in both scenarios, but the benefits of being aggressive decrease as the population average aggressivness increases (i.e., as in the hawk-dove game). In scenario (A), β_I equals 0.02 and in scenario (B), β_I equals 0.015. The circles depict the evolutionary trajectory toward the ESS, blue represents less agresssive while red represents more aggressive population average phenotype. The evolutionary equilibrium is achieved when the selection differential equals zero. The greater β_I in (A), represented as a more warped fitness surface, thus results in a lower ESS level of mean aggressivenss in the population (*z*') – see Appendix S2.

phenotype on the fitness of the focal individual in the average social environment over the study period.

Despite its importance in game theory models, the role of the interactive fitness coefficient (β_I) in determining responses to selection has not been fully appreciated (Westneat 2012), potentially because of the common statistical practice of mean centering the phenotype in each episode of selection. Theoretical studies of social evolution have highlighted the importance of interactive fitness effects in the context of synergistic effects (Queller 1985, 1992b, 2011), and Westneat (2012) has advocated for its integration into the indirect genetic effects (IGE) framework explicitly. However, interactive fitness effects have been largely neglected in theoretical IGE models (see appendix A1 in Wolf et al. 1999). This is also the case empirically. For instance, social selection gradients have been estimated in different taxa using both neighbor-modulated approach and multi-level selection analyses (Stevens et al. 1995; Eldakar et al. 2010; Formica et al. 2011; Fisher et al. 2017; Santostefano et al. 2019), but to our knowledge, there are few studies that have estimated interactive fitness effects (e.g., Brodie and Agrawal 2001; Santostefano et al. 2019). The lack of empirical estimates and their omission in some theoretical treatments of social evolution is hampering our ability to understand, describe, and ultimately predict the evolutionary trajectories of social traits.

Interacting fitness effects, as described in Equation 1, will determine the strength of selection acting on phenotypes. We can describe the magnitude and strength of selection in the trajectory of a population toward the ESS by focusing on the selection differentials (**s**) across selection events (Robertson 1966; Price 1970; Lande and Arnold 1983). The vector of such selection differentials (**s**) across time can be described as:

$$\mathbf{s} = \left(\beta_N + \beta_I \bar{\mathbf{z}}'\right)P + (\beta_S + \beta_I \bar{\mathbf{z}})C_{ZZ'}, \tag{2}$$

where the first term is a function of the non-social selection gradient (β_N) and the phenotypic variance (P) of trait z, with β_I describing how the relationship between phenotype and relative fitness depends upon the mean social environment ($\bar{z}'$). The second term $(\beta_S + \beta_I \bar{z})C_{ZZ'}$ represents the strength of social selection. This term shows that the net force of social selection in a particular selection event depends upon the social selection gradient (β_S), the interactive fitness effect (β_I) and the mean focal phenotype ($\bar{z}$), all moderated by the covariance between the phenotypes of interacting individuals ($C_{ZZ'}$). Similar formulations of this equation have been presented elsewhere, for instance in Queller (1992b) and in the appendix of Wolf et al. (1999). Here, we present a different formulation to highlight that the relationship between the focal trait and fitness is modulated by the interactive fitness effects (see Appendix S1 for details on the derivation), thereby explicitly linking these equations with game theory.

When there are interactive fitness effects, even in the absence of a covariance between the traits of interacting individuals, the net force of selection is a function of the mean social environment. We use the explicit definition of frequency dependence in game theory models for continuous traits, usually described as $W(z,\ \bar{z}')$, where the fitness function W relating the phenotype z to fitness $\boldsymbol{w}$ is a function of the mean social environment ($\bar{z}'$) (McGill and Brown 2007). Consider the classic hawk-dove game example (Maynard Smith and Price 1973) in which two fixed types interact at random. We can extend this model to a quantitative trait defined as aggressiveness. In this simple scenario, β_I captures the negative frequency dependence of such a game (Queller 1984) and how the adaptive value of playing "hawk" or "dove" depends upon the mean aggressiveness in the social environment. If we know both β_N and β_I, we can then predict how populations will arrive at the ESS (see the colored circles in Fig. 1). Measuring β_I thus enables predicting the equilibrium mean aggressiveness (ESS) and the evolutionary responses after stochastic perturbations shifting the mean phenotype away from the ESS (see Appendix S2). Interactive fitness effects imply that the strength of selection varies as the mean population phenotype changes, which is consistent with the idea that the rate of evolution slows as the population approaches the ESS (Fig. 1; Fig. S1). Interactive fitness effects are therefore a key component of social evolution models and thus need to be considered when empirically studying the evolutionary dynamics of social traits.

Predicting Evolutionary Responses in a Social Context

Phenotypic responses to selection occur via various pathways causing covariation between genes and fitness (Frank 1997; Queller 1992b). Following standard quantitative genetic assumptions, we can describe the additive genetic effects on a phenotypic trait using breeding values (Falconer and Mackay 1996). We can alter Equation 2 to predict the changes in the mean phenotype across generations ($\Delta\bar{z}$) by including how the breeding values in the focal individual co-vary with its phenotype (i.e., the additive genetic variance, C_{AZ}), and how the breeding values in the focal individual co-vary with the phenotype of its social environment ($C_{AZ'}$). This last term describes the genetic underpinning of the covariance between the phenotypes of interacting individuals ($C_{ZZ'}$), which results in a genetically explicit version of Equation 2:

$$\Delta\bar{\mathbf{z}} = \left(\beta_N + \beta_I\bar{\mathbf{z}}'\right)C_{AZ} + (\beta_S + \beta_I\bar{\mathbf{z}})C_{AZ'}. \tag{3}$$

Phenotypic change across generations is a function of additive genetic variation in the trait (C_{AZ}), the non-social selection gradient and how the interactive fitness effects modulate the adaptive value of the phenotype ($\beta_N + \beta_I\bar{\mathbf{z}}'$). Evolutionary change will also be a function of the social selection gradient, the mean focal trait, and any interactive fitness effects ($\beta_S + \beta_I\bar{\mathbf{z}}$), if the genes of the focal individual and the phenotype of its social partner co-vary ($C_{AZ'} \neq 0$). Versions of this equation, but without the interactive fitness effects have been presented elsewhere (e.g., McGlothlin et al. 2010). Equation 3 as we present it here explicitly describes evolutionary changes across generations, combining the frequency dependence described in game theory (Queller 1984, 1985; Westneat 2012) with the genetic underpinnings of the covariance between the phenotypes of interacting individuals (Frank 1997; Wolf et al. 1999; McGlothlin et al. 2010).

The Social Evolution Path Diagram

How social interactions affect inheritance and selection can be studied using path analyses. This approach is general, as the different processes that affect phenotypic evolution can be described via different formulations of the Price equation (Price 1970; Frank 1997; Queller 2017), which can then be modified into specific multiple regressions (Lande and Arnold 1983) and then combined with path analytical models (Wright 1934; Frank 1997; Morrissey 2014) to be used in empirical studies. The use of path analyses to understand selection has a long history in evolutionary genetics (Wright 1934; Kingsolver and Schemske 1991; Scheiner et al. 2000) and the path analytical representation has been used in several treatments of social evolution (McGlothlin et al. 2010; Queller 2011; Hadfield and Thomson 2017), although not necessarily in a formal statistical manner (but see Frank 1997). We use path analyses as both a conceptual and statistical tool to illustrate and link social evolution models with the key parameters that empiricists need to estimate in evolutionary studies of social behavior. The social path diagram focuses on the causal pathways linking genes or breeding values $\boldsymbol{a}$ to relative fitness $\boldsymbol{w}$ (Fig. 2). Rearrangement of Equation 3 reveals how the selection gradients define three routes by which the genes of the focal individual can affects its fitness:

$$\Delta\bar{\mathbf{z}} = C_{AZ}\beta_N + C_{AZ'}\beta_S + \left(\bar{\mathbf{z}}'C_{AZ} + \bar{\mathbf{z}}C_{AZ'}\right)\beta_I. \tag{4}$$

Wolf et al. (1999) provide a non-genetic version of this equation in their appendix, and in Appendix S1 of this paper we provide details on the formal derivation. We now discuss in turn the mechanisms underpinning each of these routes in Figure 2.

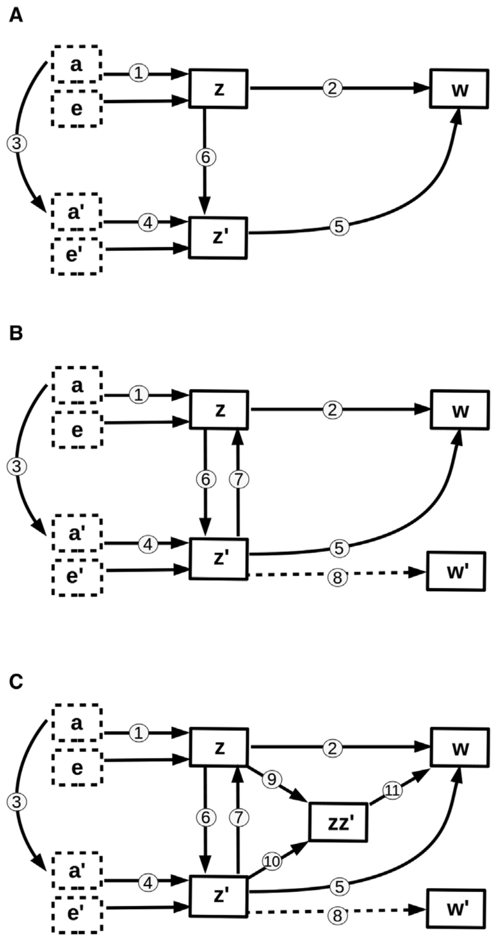

Figure 2. Path diagrams of social evolution. Symbols in solid line boxes are measured variables, symbols in dashed line boxes represent unmeasured (statistically inferred) variables. Numbers in circles are paths (ρ) representing the relationships between variables. (A) Indirect routes to fitness. ρ_1 reflects the genetic underpinning of the focal individual's trait, represented as the effect of the breeding values a on trait z with e representing unmeasured environmental effects. ρ_2 is the relationship between the trait of the focal individual *z* and its relative fitness w (β_N). ρ_3 represents a measure of genetic relatedness (*r*) between interacting individuals. ρ_4 reflects the genetic underpinning of the trait *z*′ of social partner(s), represented as the effect of genes a′ on trait z′ with e′ again representing unmeasured environmental effects. ρ_5 is the relationship between the trait of the social partner z′ and the focal individual's fitness ω, which can be interpreted as the social selection gradient (β_S). ρ_6 is the social impact ($\psi_{zz'}$) of the phenotype of individual z on the phenotype z′ of the social partner(s). (B) Social responsiveness and the evolution of the social environment. ρ_7 represents the social responsiveness ($\psi_{z'z}$) of the phenotype of the focal individual to the phenotype of its social partner. ρ_9 represents the direct selective pressures (β'_N) on the social partner(s) phenotype and w′ the fitness of the social partner(s). (C) Interacting fitness effects. ρ_9 and ρ_{10} represent the influences of the focal and social partner(s) traits on the aggregate character zz′. ρ_{11} represents interactive fitness effects (β_I) mediated by the aggregate character (zz′).

ROUTE 1: DIRECT ROUTE TO FITNESS (β_N)

The first term in Equation 4, $C_{AZ}\beta_N$, represents the direct route to fitness (Fig. 2A) providing a causal pathway linking breeding values (***a***) to relative fitness (***w***). Path 1 links breeding values with the phenotype and path 2 links phenotype and fitness via the nonsocial selection gradient (β_N). For tractability in the path analysis, we can assume that the breeding values are scaled by their variance. In this way, path 1 is equal to the additive genetic variance of the focal trait and can be estimated using pedigree information (Wilson et al. 2010). The evolutionary change caused by this route can then be estimated by multiplying its paths ($p_1 p_2$).

ROUTE 2: INDIRECT ROUTES THROUGH THE SOCIAL SELECTION GRADIENT (β_S)

The second term in Equation 4 ($C_{AZ'}\beta_S$) describes the effect of non-random distributions of genotypes in the social environment and the social selection gradient (Frank 1997; Wolf et al. 1999; McGlothlin et al. 2010). Whenever there is a covariance between the genes of the focal individual and the phenotype of its social partner(s) ($C_{AZ'}$), a nonzero social selection gradient (β_S) will affect the evolutionary trajectory of a social trait. Relatedness and indirect genetic effects (IGEs) may generate non-random distributions of genes in the social environment, and their evolutionary consequences have been described elsewhere (Queller 1992a,b; Wolf et al. 1998; McGlothlin et al. 2010). Here, we review their effects on responses to selection with the aim of discussing the evolution of behavioral mechanisms underpinning these processes. We classify these behavioral mechanisms as "social assortment," "social impact," and "social responsiveness" (for equivalent concepts, see: Edelaar and Bolnick 2019).

Relatedness and social assortment

When interacting individuals are non-randomly assorted and the fitness of the focal individual is affected by the phenotype of its social partner, social selection can affect the evolutionary trajectory of a social trait (Queller 1992a; Frank 1997; McGlothlin et al. 2014). In this scenario, the covariance between the genes in the focal individual and the phenotype of its social partner(s) ($C_{AZ'}$ in equation 4) can be caused by non-random social assortment, resulting in covariance between the breeding values of the interacting partners ($C_{AA'}$). The genetic covariance between the traits of interacting individuals can be expressed as the regression coefficient ***r*** (Michod and Hamilton 1980; Queller 1992a), which can be understood as a measure of genetic relatedness (Grafen 1985). The importance of genetic relatedness in the evolution of social behavior is captured in Hamilton's rule (Hamilton 1964a,b) and its neighbour-modulated version (Frank 1997; McGlothlin et al. 2014; Queller 1992a) and, as such, plays a key role in the evolution of cooperation (Grafen 1985). In general, the coefficient ***r*** (path 3, Fig. 2A) can represent all the behavioural and demographic mechanisms that determine the genetic similarity between an individual and its social environment. The parameter ***r*** can thus be a function of dispersal behaviour determining population viscosity, genetically determined habitat choice where social interactions tend to be more likely between individuals of a similar genotype, or of kin discrimination where individuals preferentially interact with kin (Hamilton 1964). In addition, ***r*** can also represent the non-random assortment caused by mate choice (Andersson 1994). Therefore, despite the fact that we refer to this parameter as ***r*** in order to be consistent with its reference to kin selection, it can also involve other process that directly or indirectly result in a non-random distribution of genotypes in social interactions (McGlothlin et al. 2010). In combination with the additive genetic variation of the trait in the social partner(s) (path 4) and the social selection gradient (path 5), relatedness between interacting individuals (path 3) defines an indirect route between the focal genes and fitness. Its effect on evolutionary change can be estimated by the product of these paths ($p_3 p_4 p_5$).

Social impact

We use the term "social impact" to refer to situations where the focal individual's trait impacts the phenotype(s) of its social partner(s) (path 6, Fig. 2) and it creates an alternative social route when the social partner(s) phenotype in turn influences the focal's fitness (path 5, Fig. 2). In this scenario, the covariance between the genes underpinning the focal trait and the trait of the social partner ($C_{AZ'}$ in equation 4) is a function of the average social impact. This process has also been referred to as "kith selection" by Queller (2011). Social impact plays an analogous role to other forms of environmental modification, such as niche construction (Odling-Smee et al. 2003; Saltz and Nuzhdin 2014), but in this case individuals affect the phenotype(s) of interacting conspecifics. This route may represent any phenotype used to manipulate a social partner. For example, z could represent a sexual ornament in a male that induces a female to differentially increase her care (z') in their joint offspring (see Haaland et al. 2017). Trait z will increase a male's fitness indirectly through its effects on his female partner. This route is depicted by paths 1, 5, and 6 in Figure 2, and its effect on evolutionary change can be estimated as their product ($p_1 p_5 p_6$). When there is social impact, the non-social selection gradient reflects the effect of the focal phenotype on fitness independent of its indirect effect on fitness through the phenotype of the social partner. Thus, the overall relationship between the phenotype and fitness comes from the direct route plus the indirect routes ($p_1 p_2 + p_1 p_5 p_6$). The path analysis clarifies that social impact will change the pattern of selection through its indirect effect on fitness through the phenotype of the social partner.

Social responsiveness

We refer to "social responsiveness" in situations where the focal individual adjusts its phenotype to that of the social partner (Van Cleve and Akçay 2014; Dingemanse and Araya-Ajoy 2015). This is a form of phenotypic plasticity that changes the pattern of inheritance of the focal trait (Moore et al. 1997). How the pattern of inheritance can be changed in this way by social interactions has been studied using indirect genetic effects models (IGEs: Griffing 1967; Moore et al. 1997) and their evolutionary consequences have been well described in uni- and multivariate scenarios (Wolf et al. 1998; McGlothlin et al. 2010). One of the key insights of indirect genetic effect models (Moore et al. 1997) and early maternal effect models (Kirkpatrick and Lande 1989) is that even if there is no (direct) genetic variance in the focal trait, changes in the social environment may still result in a change in the mean focal phenotype. Focusing on the paths of Figure 2, if there are genes underpinning the expression of the social partner(s) trait (p_4), and there is also selection (p_8) on the social partner(s) trait (z'), then social responsiveness (p_7) will lead to a change in the mean phenotype of the focal trait. As an example, consider a simplified hypothetical two-trait scenario concerning parental provisioning behavior (z) of altricial offspring, where parents plastically adjust to the begging behavior of their offspring (z') (see Smiseth et al. 2008). If the mean begging behavior of the offspring increases from one generation to the next ($\Delta\bar{z}'$), perhaps because offspring that begged more gained a selective advantage, then the mean level of provisioning behavior in the next generation will also change due to the plastic response of parents, even in the absence of a change in the gene frequencies for parental provisioning behavior ($\boldsymbol{a}$). This dynamic will have consequences in terms of subsequent selective events on parental provisioning z and offspring begging z' before any type of ESS compromise is reached (see Smiseth et al. 2008).

Impact, responsiveness and IGE models

A key parameter in indirect genetic effect (IGE) models is the coefficient psi (ψ; Moore et al. 1997) representing the effect of the phenotype of social partner(s) on the phenotype of a focal individual (Van Cleve and Akçay 2014; Dingemanse and Araya-Ajoy 2015). When viewed from a behavioral ecology perspective, this coefficient is in fact underpinned by two linked but different processes that we have defined (above) as "social impact" and "social responsiveness." Social responsiveness is related to social impact, but when we refer to responsiveness, it is the phenotype of the social partner(s) that influences the phenotype of the focal individual. In contrast, when we refer to impact, it is the phenotype of the focal individual that affects the expression of the social partner(s) phenotype. To clarify our argument, we start by describing the phenotypic expression of the phenotype of the focal individual z and its social partner z', focusing on the coefficient ψ (Moore et al. 1997) but using the different subscripts to denote directionality and causality:

$$z = c + \psi_{z'z} z' + \boldsymbol{a} + \boldsymbol{e}, \tag{5a}$$

$$z' = c' + \psi_{zz'} z + \boldsymbol{a}' + \boldsymbol{e}', \tag{5b}$$

where $\boldsymbol{a}$ and $\boldsymbol{e}$ are the additive genetic effects and environmental variables affecting trait z of the focal individuals. As with z' (above), we denote with a prime the symbols for terms referring to the phenotype of the social partner(s) that are analogous to the parameters describing the phenotype of the focal individual. In Equation 5b, c' is the mean phenotype of the social environment, $\boldsymbol{a}'$ is the genetic underpinning (breeding value), and $\boldsymbol{e}'$ represents environmental effects on trait z'. $\psi_{Z'Z}$ in Equation 5a represents the "social responsiveness" of the phenotype of the focal individual to the phenotype of its interacting partner(s). In contrast, $\psi_{ZZ'}$ in Equation 5b represents "social impact" of the phenotype of the focal individuals z on the phenotype of its interacting partner(s) z'. These effects can be non-reciprocal (one trait affects the other trait, but not *vice versa*) versus reciprocal situations where the two traits affect each other (Moore et al. 1997). When there are non-reciprocal effects and the focal and social traits are different, it may be possible to distinguish between impact and responsiveness (e.g., a male's ornament versus the female preference, Andersson 1994; or parent provisioning and offspring begging, Smiseth et al. 2008). However, in a scenario with reciprocal effects within the same trait, it may difficult to distinguish between impact and responsiveness empirically, even if they are theoretically distinct (e.g., escalation in aggressive contests; Enquist and Leimar 1983).

When the focal and the social environment phenotype are the same (e.g., level of aggressiveness), then social impact and social responsiveness can be seen as two sides of the same coin. Therefore, p_6 and p_7 in Figure 2 may be better represented as a single double headed arrow (see Moore et al. 1997). It is also important to note that in a scenario where traits are the same, then p_2 and p_8 both reflect the relationship between the phenotype and fitness, but p_2 is a partial regression coefficient measuring only the direct effects of the trait on fitness, while p_8 subsumes both the direct and indirect effects. Furthermore, when the trait in the focal individual and the trait in the social environment are the same or there are reciprocal responses between different traits, there will be feedback affecting the genetic underpinning of the traits further affecting the additive genetic variance of the focal trait and thus its evolutionary trajectory even if the social selection gradient is zero (Moore et al. 1997; McGlothlin and Brodie 2009; McGlothlin et al. 2010). Importantly, in both scenarios of reciprocal and non-reciprocal responses, the evolutionary consequences of IGEs will be determined by the evolutionary compromise at equilibrium (the ESS) between responsiveness and impact at the population level. However, differentiating between social impact and social responsiveness, and recognizing that different traits can underlie these two processes, is essential for understanding how selection may shape the average covariance between phenotypes and their social environment (Queller 2011; Edelaar and Bolnick 2019).

ROUTE 3: THE INTERACTION EFFECT (β_I)

Interactions in fitness effects between the focal and social partner(s) traits (β_I) are a major driving force of social evolution. The product of the phenotypic values of interacting individuals can be treated as an aggregate character resulting from the social interaction(s) experienced by the focal individual (Salt 1979; Heisler and Damuth 1987). We can focus on the change in mean phenotype in one selection event using an alternative formulation of Equation 4 provided by Queller (1992b) to incorporate interacting fitness effects in terms of covariances between breeding values and the traits:

$$\Delta\bar{\mathbf{z}} = C_{AZ}\,\beta_N + C_{AZ'}\beta_S + \beta_I C_{AZZ'}, \tag{6}$$

here, $C_{AZZ'}$ represents the covariance between the breeding values of the focal trait and the aggregate character zz'. This shows that the covariance between the aggregate property $\mathbf{zz}'$ and the focal phenotype breeding values mediate the contribution of interactive fitness effects on the responses to selection. The aggregate trait $\mathbf{zz}'$ can be estimated as the product of focal and social phenotypes, and can therefore be treated as another route by which the genes of the focal individual can impact its fitness (Queller 1985, 1992a,b).

Using the formula for estimating the covariance between a vector and a product (Bohrnstedt and Goldberger 1969; Brown and Alexander 1991), we can show the underpinnings of the covariance between breeding values and the aggregate property $\mathbf{zz}'$. This clarifies how Equation 6 relates to Equation 4, along with some important underlying assumptions of this framework. We can express the covariance between the breeding values of the focal trait and the aggregate trait as:

$$C_{AZZ'} = \bar{z}C_{AZ'} + \bar{z}'C_{AZ} + E\big[(\mathbf{a} - \bar{a})(\mathbf{z} - \bar{z})\big(\mathbf{z}' - \bar{z}'\big)\big]. \tag{7}$$

The term $(\bar{z}C_{AZ'} + \bar{z}'C_{AZ})$ describes the effect of the mean population focal and social phenotypes on the covariances between the breeding values and the aggregate property zz'. This term is also in Equation 4, showing how the effects in the responses to selection of the mean focal and social phenotypes are mediated by the interactive fitness effects $\beta_I(\bar{z}C_{AZ'} + \bar{z}'C_{AZ})$. One might expect that the covariance $C_{AZZ'}$ is nonzero if there is a genetic basis for the focal trait ($C_{AZ} \neq 0$) or if there is covariance between the breeding values of the focal phenotype and the trait of the social partner ($C_{AZ'} \neq 0$). However, Equation 7 shows that this will not be the case if the means of the focal and social phenotypes are set to zero. This again highlights certain undesirable consequences of mean centering traits for each selection event in scenarios where there is frequency dependence.

The second term in equation 7 is $E[(\mathbf{a} - \bar{a})(\mathbf{z} - \bar{z})(\mathbf{z}' - \bar{z}')]$, where E refers to the expectation of the product inside the brackets, describes how any non-linear relationship between the three variables will affect the covariance $C_{AZZ'}$. This term will be zero if variables conform to multivariate normality (Bohrnstedt and Goldberger 1969). There are certain types of non-random social assortments that will affect the skewness and kurtosis of the joint distribution of breeding values and social and focal phenotypes, which would lead to violations of multivariate normality and thus affect $C_{AZZ'}$. For example, this may happen when individuals with higher values of the focal trait are associated with more related individuals as compared with individuals with lower trait values, or when individuals with higher trait values have a disproportionally stronger impact on the phenotype of their social partner(s) (see Appendix S3 for an example using simulations). These types of non-random distributions may be common in natural populations, but their relative importance in social evolution remains an empirical question.

The aggregate property $\mathbf{zz}'$ creates an alternative route between the genes of the focal phenotype and fitness, which resembles other indirect routes in social selection, but is mediated by the aggregate property $\mathbf{zz}'$ and its effects on fitness (β_I). Importantly, the links between the aggregate property $\mathbf{zz}'$ and the genes of the focal individual may fluctuate over time as the mean values of the social traits change, and thus the effect of β_I on the response to selection will also change. In this case, the covariance between the character $\mathbf{zz}'$ and the breeding values of the focal individuals ($\boldsymbol{a}$) may be better represented as a vector $\mathbf{C_{AZZ'}}$ of values that will vary across time as the focal phenotype and the social environment evolve. Furthermore, β_I will not only affect the response to selection directly, but will also affect the patterns of non-random social assortment of genes in the social environment, because it will impose selection on the traits that affect how individuals chose, impact or respond to their social environment.

The Evolution of Social Impact, Responsiveness and Non-Random Social Assortment

A central question in social evolution is how non-random social assortment, social impact and social responsiveness evolve (Akçay and Van Cleve 2012; Dingemanse and Araya-Ajoy 2015). We emphasize that interactive social fitness effects (β_I) impose selection on other traits that determine how individuals impact, respond and "choose" their social environments. Many of the traits mediating these processes coevolve in context of cooperation (e.g., communication) and competition (e.g., manipulation versus resistance). Game theorists have addressed the ecological conditions favoring the evolution of strategies responsive to the social environment (i.e., phenotypically plastic) leading to conditional ESSs (Tomkins and Hazel 2007) versus unresponsive strategies leading to fixed ESSs. Behavioral ecologists have also shown that non-random social assortment (i.e., "choice" of social partners) represents an evolvable element under selection that is important for the evolution of altruism (Fletcher and Doebeli 2006, 2009). Furthermore, empirical work has confirmed that these processes can change in response to selection. For instance social responsiveness of male flies to potential mates ($\psi_{z'z}$) has been shown to evolve under artificial selection (Chenoweth et al. 2010), and the responsiveness of female crickets to their mate's song varies among populations experiencing different ecological characteristics (Bailey and Zuk 2012). Thus, while the mean relatedness and the mean levels of social impact and responsiveness may be stable in the short term, they are expected to change over time in response to consistent selection pressures on the traits that determine their values. This will in turn further affect any evolutionary response to social and non-social selection (Kazancıoğlu et al. 2012).

The evolution of individual impact upon, responsiveness to, and choice of social environments requires genetic variation in traits underlying these processes. Thus far we have assumed that $\psi_{zz'}$ represents the mean impact of a focal's trait on its social environment, but this value can vary genetically among individuals (Saltz and Nuzhdin 2014). Therefore, we now represent it in bold as a vector of values ($\boldsymbol{\psi}_{zz'}$). Hence, the impact of a given focal trait value for an individual might deviate from the population mean impact due to differences in another heritable trait. For example, the impact of a male's tail length may vary depending on his inherited ability to perform a courtship display involving the tail. We can also think about this as an interaction between traits of the same individual, where the impact of a trait (tail; z) depends on another trait (its ability to display; $\boldsymbol{\psi}_{zz'}$). In a similar fashion, social responsiveness ($\boldsymbol{\psi}_{z'z}$) can vary among individuals due to differences in cognitive ability to perceive variation in their social environment (Chenoweth et al. 2010). The traits that mediate social assortment and determine the mean genetic relatedness of each focal individual to the rest of their social group ($\boldsymbol{r}$) can also be conceptualized as a trait that can evolve (Gardner et al. 2007). For instance, individual variation in the ability to discriminate between individuals of different genotypes may evolve under circumstances where altruism could be directed to an array of recipients that vary in relatedness (Hepper 1986).

Assuming that these traits exhibit genetic variation, then the evolutionary change in traits affecting relatedness, social impact, and social responsiveness can be described in terms of their covariances with fitness: $Cov(\mathrm{r, w})$; $Cov(\psi_{zz'}, \mathrm{w})$; and $Cov(\psi_{z'z}, \mathrm{w})$. Several mechanisms could generate a covariance between fitness and social assortment, social impact and responsiveness. However, the aim of this section is to highlight that evolutionary change in these associated characters is expected when there is a covariance between the aggregate character $\mathbf{zz}'$ and individual genetic relatedness ($\boldsymbol{r}$), social impact ($\boldsymbol{\psi}_{zz'}$), or social responsiveness ($\boldsymbol{\psi}_{z'z}$), combined with an effect of the aggregate character $\mathbf{zz}'$ on the fitness of the focal individual ($\beta_I \neq 0$). The expected effect of variation in social assortment, social impact and responsiveness on fitness through the aggregate property $\mathbf{zz}'$ can be estimated using the formula to calculate the covariance between a vector and a product (Bohrnstedt and Goldberger 1969; Brown and Alexander 1991). For instance, the covariance between fitness and relatedness ($Cov[\boldsymbol{r}, \boldsymbol{w}]$) that is caused by the interactive fitness effects (β_I) can be described as:

$$\begin{aligned}\Delta r \;=\; & Cov\,[r, w] = \beta_I C_{RZZ'} \\ & = \beta_I \left(\left[\bar{z} C_{RZ'} + \bar{z}' C_{RZ} \right] + E\left[(\mathbf{r} \;-\; \bar{r})\,(\mathbf{z} \;-\; \bar{z})\left(\mathbf{z} \;-\; \bar{z}'\right) \right] \right). \qquad (8)\end{aligned}$$

We can infer from Equation 8 that the phenotypic distribution caused by variation in relatedness will result in a covariance between the aggregate trait and individual relatedness ($C_{RZZ'}$), and if the aggregate property zz' affects fitness (β_I) then relatedness and fitness will co-vary, imposing selection on any mechanism affecting the relatedness between interacting individuals.

Feedbacks in Social Evolution

Evolutionary feedbacks can arise because interactive fitness effects can change the adaptive value of phenotypes as the social environment evolves, and because it may impose selection on the traits determining levels of impact, responsiveness and relatedness. This in turn will shape any subsequent short-term phenotypic responses to social selection on other social traits (Kazancıoğlu et al. 2012). In a recursive equation, we can describe the different social processes that may affect the mean phenotype in the subsequent generation $\bar{z}_{t+1}$. For simplicity, in Equations 9a and 9b we assume that there is only genetic variance in the focal phenotype, and in the traits influencing social impact, responsiveness and relatedness, and also that their respective genetic variance is equal to one. In scenarios where the focal and social traits are different and there are no reciprocal responses, the expected mean phenotype in generation t+1 can be described as:

$$\begin{aligned}&\bar{z}_{t+1} = \bar{z}_t + \beta_N + \beta_S\left(r_t + \psi_{zz'\,t}\right) + \beta_I\left(\bar{z}_t' + \bar{z}_t r_t + \bar{z}_t \psi_{zz'\,t}\right) \\ &+\bar{e}_{t+1}, \qquad (9a)\end{aligned}$$

where the mean phenotype in the next generation ($\bar{z}_{t+1}$) depends upon the direct selection gradient (β_N), while the contribution of the social selection gradient is modulated by relatedness and social impact $\beta_S(r_t + \psi_{zz'\,t})$. In turn, impact, relatedness and the mean focal and social phenotypes will modulate the effect of the interactive fitness effects coefficient ($\beta_I[\bar{z}_t' + \bar{z}_t r_t + \bar{z}_t \psi_{zz'\,t}]$). In Equation 9a, $\bar{e}_{t+1}$ reflects the mean effect of the social environment on the expression of the phenotype at time step t+1:

$$\bar{e}_{t+1} = \left(\bar{z}_t' + \beta_N'\right)\left(\psi_{z'zt} + \beta_{\psi_{z'z}}\right). \qquad (9b)$$

The effect of the social environment on the expression of the focal phenotype at time t+1 depends on the mean social environment at time t and the selection experienced by the social environment ($\bar{z}_t' + \bar{\beta}_N'$), mediated by social responsiveness and selection on social responsiveness ($\psi_{z'zt} + \beta_{\psi_{z'z}}$). The effect of social responsiveness is shaped by the selection imposed by interactive fitness effects on the social traits at time t, but its effects on the phenotype only occur at time t+1 as this is mediated by phenotypic plasticity. Equations 9a and 9b therefore describe the potential feedbacks that can arise when the optimal phenotype for an individual depends upon the phenotype of its social environment. The fates of the social traits may also be determined by their correlation and feedbacks with other traits (McGlothlin et al. 2010; Hadfield and Thomson 2017). However, the univariate treatment shown here illustrates the potential for these types of evolutionary feedbacks in social evolution.

As an example of these feedback, consider the evolution of helping-at-the-nest in cooperatively breeding birds. In a hypothetical scenario, let $\boldsymbol{z}$ represent the tendency to help in the focal individual. Helping has a fitness cost to the focal individual ($\beta_N = -0.1$), but it is beneficial when a social partner helps the focal individual ($\beta_S = 0.7$). When interactions are at random (i.e., $r = 0$), the best strategy is not to help but to associate with social partners that help you. The black lines in Figure 3A show that under this scenario the tendency to help in the population will decrease over time. However, if individuals with the tendency to help associate with each other ($r = 0.05$), for instance by means of delayed dispersal causing population viscosity, then individuals with the same genetic tendency to help will interact with each other. Under this scenario, individuals with the helping tendency will reap the benefits of being helped and the helping tendency will increase over time (Fig. 3A, red lines). The effect of genetic relatedness ($\boldsymbol{r}$) within breeding groups has been demonstrated to be crucial in determining the evolution of helping (see Cornwallis et al. 2009; Wright et al. 2010). Helping behavior will

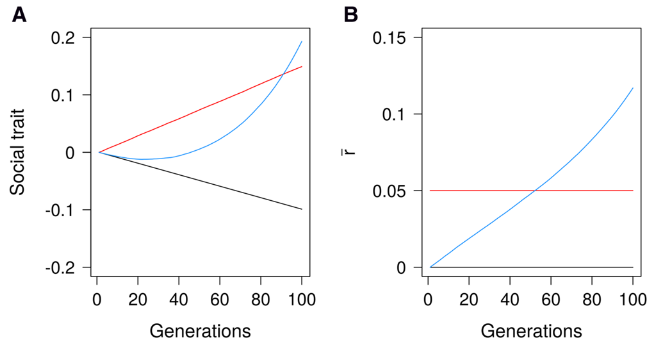

Figure 3. **Numerical evaluation of evolutionary change scenarios using Equations 8 and 9 for the (A) social trait, and (B) the relatedness between interacting individuals. Black lines represent a scenario where individuals interact at random ($r = 0$) and individuals that express the behavior are at a direct fitness disadvantaged ($\beta_N = -0.1$). Red lines represent a scenario where individuals interact preferentially with related individuals ($\bar{r} = 0.05$) and the behavior is favored because of this combination of non-random genetic assortment and social selection effects ($\beta_S = 0.7$). Blue lines represent a scenario where, in addition to the negative direct fitness effects ($\beta_N = -0.1$) and the positive social selection effects ($\beta_S = 0.7$), there are interactive effects on fitness ($\beta_I = 0.01$) – see Figure 2. This causes the social trait and relatedness to both evolve, creating a feedback that speeds up the evolution of the social trait, even under conditions where interacting individuals were initially unrelated ($\bar{r}_1 = 0$).**

evolve when the benefits of being helped outweigh the costs of helping others (i.e., $r\beta_S > \beta_N$).

The order of events described above follows standard cooperative breeding theory, where the evolution of delayed dispersal and kin discrimination are important precursors to the evolution of helping behavior (Koenig and Dickinson 2004, 2016). The social phenotypes expressed within such cooperative breeding groups are, however, likely to show interactive effects ($\beta_I > 0$). For instance, if helping also causes an increase in the recipient's subsequent ability to help (e.g., helped young grow faster and are able to more quickly provide greater help themselves as a result), then the non-additive benefits from helping will select for mechanisms that allow individuals to preferentially associate even more with related individuals that also have the helping tendency. Therefore, converse to the standard theory, the fitness benefits of being helped could feedback and influence the evolution of greater natal philopatry or ability to discriminate kin, which would further increase relatedness within groups (blue line, Fig. 3B) and further favor the evolution of helping (blue line, Fig. 3A).

Consider also an example from sexual selection with extravagant male ornaments and female mate choice. Let z be the length of the tail in males of a bird species and z' the female preference for long tails. Having an ornament may come at a direct selection cost (β_N = -0.1), perhaps because it decreases a male's ability to escape from predators, which males with less attractive ornaments do not pay. This would normally lead to shorter tails across time (black lines, Fig. 3A). However, when males with ornaments encounter females with the preference, ornamented males will have higher fitness than the males without. The adaptive value of the ornament in each episode of selection is thus dependent upon the interaction fitness coefficient and the mean preference in the population ($\beta_I \bar{\mathbf{z}}'$). If there is variation in the processes by which males encounter females, these interactive fitness effects will favor mechanisms (***r***) where long-tailed males encounter females with the preference for long tails. This will lead to the evolution of strategies, such as the use of leks for sexual display by long-tailed males that could be located in places frequented by females with the preference (i.e., "hot-spots"; Beehler and Foster 1988), and this will in turn select for even longer-tailed males even when they were initially selected against (blue lines, Fig. 3A). Similarly, if there are behavioral displays that amplify the impact ($\psi_{ZZ'}$) of an ornament (Bogaardt and Johnstone 2016), for instance a dancing display, this will further benefit males that have long tails and perform the dancing display, further favoring the evolution of even longer tails. Our model thus links previously separated ideas in sexual selection and highlights the evolution of sequences of traits due to feedbacks.

We acknowledge that in many instances the evolution of social assortment, impact, and responsiveness are more complicated than we have depicted. The evolutionary dynamics of impact, responsiveness and non-random social assortment are especially complicated in scenarios involving evolutionary conflicts of interest. Specifically, the co-evolution of impact and responsiveness are intertwined in any game theoretical scenario of conditional ESSs and in the additional complexity of behavioral dynamics in any manipulation versus resistance or response versus counter-response (McNamara et al. 1999). This includes social interactions such as biparental care, collective vigilance, agonistic interactions, kin competition, and parent-offspring conflict (Westneat and Fox 2010). Despite all this complexity, it is only by decomposing and understanding the anatomy of different episodes of social selection that we will be able to unravel the types of behavioral and evolutionary dynamics resulting from the interplay between the evolutionary interests of interacting individuals, and how they might end in specific evolutionarily stable equilibria.

Conclusions

In this paper, we have integrated key theoretical elements regarding social evolution and presented them using a set of path diagrams intended to stimulate empirical measurement of key pathways. We describe the links between long-term predictions of evolutionary stability by game theory models and short-term quantitative genetics models providing statistical descriptions of single episodes of selection. This synthesis highlights that interactive effects on fitness of focal and social partner traits can generate evolutionary feedbacks between social traits, the social environment and the processes allowing individuals to "choose," impact, and/or respond to their social environment. We explicitly discuss these feedbacks in terms of pathways that are constructed based upon sets of multiple regression coefficients, which allows most of the parameters outlined here to be suitable for estimation in empirical studies. It is likely that many, if not all, the paths to social evolution in Figure 2C will interact to determine the selective landscape in a multi-trait context, further affecting the responses to selection (McGlothlin et al. 2010; Queller 2011; Van Cleve and Akçay 2014). Measuring these paths will not be easy. However, by isolating the distinct processes affecting the expression and adaptive value of social traits during any specific episode of selection, we may better understand the ecological conditions determining eventual evolutionary outcomes resulting from social interactions. We hope that this paper encourages empiricists to estimate interactive fitness effects and thus better connect empirical studies with the different fields of theory studying social evolution.

ACKNOWLEDGMENTS

We thank Pim Edelaar, Niels Dingemanse, Jeremy Van Cleve, Francesca Santostefano, María Moirón, and Joel McGlothlin for very useful comments on the manuscript and for insightful discussions about the topic. This work was also partly supported by the U.S. National Science Foundation and the Research Council of Norway through its Centres of Excellence funding scheme (SFF-III 223257/F50) and the research grant spatdif (274930).

AUTHOR CONTRIBUTIONS

YAA led the process and all authors contributed to the ideas and co-wrote the paper.

DATA ARCHIVING

The code to generate the data that supports the findings of this study is available on https://github.com/YimenAraya-Ajoy/SocialEvolutionPathways.

LITERATURE CITED

Akçay, E., and J. Van Cleve. 2012. Behavioral responses in structured populations pave the way to group optimality. Am. Nat. 179:257–269.

Andersson, M. B. 1994. Sexual selection. Princeton Univ. Press. Monographs in Behavior and Ecology, Princeton, NJ.

Bailey, N. W., and M. Zuk. 2012. Socially flexible female choice differs among populations of the Pacific field cricket: geographical variation in the interaction coefficient psi (Ψ). Proc. Biol. Sci. 279: 3589–3596.

Beehler, B. M., and M. S. Foster. 1988. Hotshots, hotspots, and female preference in the organization of lek mating systems. Am. Nat. 131:203–219.

Bijma, P., and M. J. Wade. 2008. The joint effects of kin, multilevel selection and indirect genetic effects on response to genetic selection. J. Evol. Biol. 21:1175–1188.

Bogaardt, L., and R. A. Johnstone. 2016. Amplifiers and the origin of animal signals. Proc. R. Soc. B Biol. Sci. 283:1–6.

Bohrnstedt, G. W., and A. S. Goldberger. 1969. On the exact covariance of products of random variables. J. Am. Stat. Assoc. 64:1439–1442.

Brodie, E. D., and A. F. Agrawal. 2001. Maternal effects and the evolution of aposematic signals. Proc. Natl. Acad. Sci. USA 98:7884 LP-7887.

Brown, D., and N. Alexander. 1991. The analysis of the variance and covariance of products. Biometrics 47:429–444.

Chenoweth, S. F., H. D. Rundle, and M. W. Blows. 2010. Experimental evidence for the evolution of indirect genetic effects: changes in the interaction effect coefficient, psi (Psi), due to sexual selection. Evolution 64:1849–1856.

Cornwallis, C. K., S. A. West and A. S Griffin. 2009. Routes to indirect fitness in cooperatively breeding vertebrates: kin discrimination and limited dispersal. J. Evol. Biol. 22:2445–2457.

Dingemanse, N. J., and Y. G. Araya-Ajoy. 2015. Interacting personalities: behavioural ecology meets quantitative genetics. Trends Ecol. Evol. 30:88–97.

Edelaar, P., and D. I. Bolnick. 2019. Appreciating the multiple processes increasing individual or population fitness. Trends Ecol. Evol. 34:435–446.

Eldakar, O. T., D. S. Wilson, M. J. Dlugos, and J. W. Pepper. 2010. The role of multilevel selection in the evolution of sexual conflict in the water strider aquarius remigis. Evolution 64:3183–3189.

Enquist, M., and O. Leimar. 1983. Evolution of fighting behaviour: Decision rules and assessment of relative strength. J. Theor. Biol. 102:387–410.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Prentice Hall, Harlow, England.

Fisher, D. N., S. Boutin, B. Dantzer, M. M. Humphries, J. E. Lane, and A. G. McAdam. 2017. Multilevel and sex-specific selection on competitive traits in North American red squirrels. Evolution 71:1841–1854.

Fletcher, J. A., and M. Doebeli. 2009. A simple and general explanation for the evolution of altruism. Proc. R. Soc. B Biol. Sci 276:13–19.

———. 2006. How altruism evolves: assortment and synergy. J. Evol. Biol. 19:1389–1393.

Formica, V. A., J. W. Mcglothlin, C. W. Wood, M. E. Augat, R. E. Butterfield, M. E. Barnard, and E. D. Brodie. 2011. Phenotypic assortment mediates the effect of social selection in a wild beetle population. Evolution 65:2771–2781.

Frank, S. A. 1998. Foundations of social evolution. Princeton Univ. Press, Princeton, NJ.

———. 1997. The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis. Evolution 51:1712–1729.

Gardner, A., S. A. West, and N. H. Barton. 2007. The relation between multilocus population genetics and social evolution theory. Am. Nat. 169:207–226.

Goodnight, C. J., J. M. Schwartz, and L. Stevens. 1992. Contextual analysis of models of group selection, soft selection, hard selection, and the evolution of altruism. Am. Nat. 140:743–761.

Grafen, A. 1985. A geometric view of relatedness. P. In R. Dawkins and M. Ridley, eds. Oxford surveys in evolutionary biology. Oxford Univ. Press, Oxford, U.K.

Griffing, B. 1967. Selection in reference to biological groups i. individual and group selection applied to populations of unordered groups. Aust. J. Biol. Sci. 20:127–140.

Haaland, T., J. Wright, B. Kuijper and I. I Ratikainen. 2017. Differential allocation revisited: when should mate quality affect parental investment? Am. Nat. 190:534–546.

Hadfield, J. D., and C. E. Thomson. 2017. Interpreting selection when individuals interact. Methods Ecol. Evol. 8:688–699.

Hamilton, W. D. 1964. The genetical evolution of social behaviour. II. J. Theor. Biol. 7:1–16.

Heino, M., J. A. J. Metz, and V. Kaitala. 1998. The enigma of frequency-dependent selection. Trends Ecol. Evol. 13:367–370.

Heisler, I. L., and J. Damuth. 1987. A method for analyzing selection in hierarchically structured populations. Am. Nat. 130:582–602.

Hepper, P. G. 1986. Kin recognition: functions and mechanisms a review. Biol. Rev. 61:63–93.

Kazancıoğlu, E., H. Klug, and S. H. Alonzo. 2012. The evolution of social interactions changes predictions about interacting phenotypes. Evolution 66:2056–2064.

Kingsolver, J. G., and D. W. Schemske. 1991. Path analyses of selection. Trends Ecol. Evol. 6:276–280.

Kirkpatrick, M., and R. Lande. 1989. The evolution of maternal characters. Evolution 43:485–503.

Koenig, W., and Dickinson J. 2004. Ecology and evolution of cooperative breeding in birds. Cambridge Univ. Press, Cambridge, U.K.

Koenig, W. and Dickinson, J. 2016. Cooperative Breeding in Vertebrates: Studies of Ecology, Evolution, and Behavior. Cambridge University Press.

Krebs, J. R., and N. B. Davies. 1991. Behavioural ecology: an evolutionary approach. 4th ed. Blackwell, Oxford.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33:402–416.

Lande, R., and S. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

Lion, S. 2018. Theoretical approaches in evolutionary ecology: environmental feedback as a unifying perspective. Am. Nat. 191:21–44.

Maynard-Smith, J., and G. R. Price. 1973. The logic of animal conflict. Nature 246:15–18.

Maynard Smith, J. M. 1982. Evolution and the theory of games. Cambridge Univ. Press, Cambridge.

McGill, B. J., and J. S. Brown. 2007. Evolutionary game theory and adaptive dynamics of continuous traits. Annu. Rev. Ecol. Evol. Syst. 38:403–435.

McGlothlin, J. W., and E. D. Brodie. 2009. How to measure indirect genetic effects: the congruence of trait-based and variance-partitioning approaches. Evolution 63:1785–1795.

McGlothlin, J. W., A. J. Moore, J. B. Wolf, and E. D. Brodie. 2010. Interacting phenotypes and the evolutionary process. III. Social evolution. Evolution 64:2558–2574.

McGlothlin, J. W., J. B. Wolf, E. D. Brodie, and A. J. Moore. 2014. Quantitative genetic versions of Hamilton's rule with empirical applications. Philos. Trans. R. Soc. B Biol. Sci. 369:20130358.

McNamara, J. M., C. E. Gasson, and A. I. Houston. 1999. Incorporating rules for responding into evolutionary games. Nature 401:368.

McNamara, J. M., and F. J. Weissing. 2010. Evolutionary game theory. Pp. 88–106 In T. Székely, A. J. Moore, and J. Komdeur, eds. Social behaviour: genes, ecology and evolution. Cambridge Univ. Press, Cambridge, U.K.

Michod, R. E., and W. D. Hamilton. 1980. Coefficients of relatedness in sociobiology. Nature 288:694–697.

Moore, A., E. D. Brodie, and J. Wolf. 1997. Interacting phenotypes and the evolutionary process: I. Direct and indirect genetic effects of social interactions. Evolution 51:1352–1362.

Morrissey, M. B. 2014. Selection and evolution of causally covarying traits. Evolution 68:1748–1761.

Odling-Smee, F. J., K. N. Laland, and M. W. Feldman. 2003. Niche construction: the neglected process in evolution. Princeton Univ. Press, Princeton.

Okasha, S. 2006. Evolution and the levels of selection. Oxford Univ. Press, Oxford, U.K.

Price, G. R. 1970. Selection and covariance. Nature 227:520.

Queller, D. C. 1992a. A general model for kin selection. Evolution 46:376–380.

———. 2011. Expanded social fitness and Hamilton's rule for kin, kith, and kind. Proc. Natl. Acad. Sci. USA 108:10792–10799.

———. 2017. Fundamental theorems of evolution. Am. Nat. 189:345–353.

———. 1984. Kin selection and frequency dependence: a game theoretic approach. Biol. J. Linn. Soc. 23:133–143.

———. 1985. Kinship, reciprocity and synergism in the evolution of social behaviour. Nature 318:366–367.

———. 1992b. Quantitative genetics, inclusive fitness, and group selection. Am. Nat 139:540–558.

Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. Anim. Sci 8:95–108.

Salt, G. W. 1979. A comment on the use of the term emergent properties. Am. Nat 113:145–148.

Saltz, J. B., and S. V. Nuzhdin. 2014. Genetic variation in niche construction: implications for development and evolutionary genetics. Trends Ecol. Evol. 29:8–14.

Santostefano, F., D. Garant, P. Bergeron, P. Montiglio, and D. Réale. 2019. Social selection acts on behavior and body mass but does not contribute to the total selection differential in eastern chipmunks. Evolution 74:89–102.

Scheiner, Mitchell, and Callahan. 2000. Using path analysis to measure natural selection. J. Evol. Biol. 13:423–433.

Smiseth, P. T., J. Wright, and M. Kölliker. 2008. Parent-offspring conflict and co-adaptation: behavioural ecology meets quantitative genetics. Proc. Biol. Sci. 275:1823–1830.

Stevens, L., C. J. Goodnight, and S. Kalisz. 1995. Multilevel selection in natural populations of Impatiens capensis. Am. Nat. 145:513–526.

Svensson, E. I., and T. Connallon. 2019. How frequency-dependent selection affects population fitness, maladaptation and evolutionary rescue. Evol. Appl. 12:1243–1258.

Tomkins, J. L., and W. Hazel. 2007. The status of the conditional evolutionarily stable strategy. Trends Ecol. Evol. 22:522–528.

Van Cleve, J., and E. Akçay. 2014. Pathways to social evolution: reciprocity, relatedness, and synergy. Evolution 68:2245–2258.

West-Eberhard, M. J. 1979. Sexual selection, social competition, and evolution. Proc. Am. Phil. Soc. 123:222–234.

Westneat, D. F. 2012. Evolution in response to social selection: the importance of interactive effects of traits on fitness. Evolution 66:890–895.

Westneat, D. F., and C. W. Fox. 2010. Evolutionary behavioral ecology. Oxford Univ. Press, Oxford.

Wilson, A. J., D. Réale, M. N. Clements, M. B. Morrissey, E. Postma, C. A. Walling, L. E. B. Kruuk, and D. H. Nussey. 2010. An ecologist's guide to the animal model. J. Anim. Ecol. 79:13–26.

Wolf, J. B., E. D. Brodie, J. M. Cheverud, A. J. Moore, and M. J. Wade. 1998. Evolutionary consequences of indirect genetic effects. Trends Ecol. Evol. 13:64–69.

Wolf, J. B., E. D. Brodie III, and A. Moore. 1999. Interacting phenotypes and the evolutionary process. II. Selection resulting from social interactions. Am. Nat. 153:254–266.

Wright, J., P. G. McDonald, L. te Marvelde, A. J. N. Kazem, and C. M. Bishop. 2010. Helping effort increases with relatedness in bell miners, but "unrelated" helpers of both sexes still provide substantial care. Proc. Biol. Sci. 277:437–445.

Wright, S. 1934. The method of path coefficients. Ann. Math. Stat. 5:161–215.

Associate Editor: J. A. Moorad
Handling Editor: T. Chapman

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Appendix S1. Derivations for the equations of responses to selection in the presence of interactive fitnesss effects.
Appendix S2. Negative frequency dependence and the hawk-dove game
Appendix S3. Multivariate normality and evolutionary responses to selection.
Figure S1. Evolutionary dynamics of aggressiveness in two simulated scenarios of negative frequency dependence, where individuals interact at random and assuming perfect inheritance as in standard game theory.
Figure S2. Multivariate distributions of breeding values(a), and focal (z) and social (z') phenotypes.