Title: Assortment and the analysis of natural selection on social traits

Running title: Phenotypic assortment and social selection

Word count: 4586

Authors: Grant C. McDonald[1], Damien R. Farine[1,3,4], Kevin R. Foster[1] & Jay M. Biernaskie[2]

[1]Department of Zoology; University of Oxford; South Parks Road; Oxford, UK OX13PS

[2]Department of Plant Sciences; University of Oxford; South Parks Road; Oxford UK OX13RB

[3]Department of Collective Behaviour, Max Planck Institute for Ornithology, Universitätsstraße 10, 78457 Konstanz, Germany.

[4]Chair of Biodiversity and Collective Behaviour, Department of Biology, University of Konstanz, Universitätsstraße 10, 78457 Konstanz, Germany.

Author for correspondence: Jay Biernaskie (email: jay.biernaskie@plants.ox.ac.uk)

Abstract. A central problem in evolutionary biology is to determine whether and how social interactions contribute to natural selection. This can be addressed with selection analyses that relate individual fitness to individual and social phenotypes. One such approach, known as social selection analysis, leads to the intuitive result that fitness effects from social partners will contribute to selection only if there is a correlation between the traits of individuals and their social partners (non-random phenotypic assortment). However, there have been inconsistencies in the application of this approach that center around the measurement of phenotypic assortment. Here, we use data analysis and simulations to resolve these inconsistencies, showing that: (i) not all measures of assortment are suitable for social selection analysis; and (ii) the interpretation of assortment, and how to detect non-random assortment, will depend on the scale at which it is measured. We discuss links between social selection and kin selection theory, and we provide a practical guide for the social selection approach.

Introduction

Plant and animal biologists have increasingly measured natural selection using multilevel selection analyses (Stevens et al. 1995; Tsuji 1995; Aspi et al. 2003; Donohue 2004; Weinig et al. 2007; Eldakar et al. 2010; Formica et al. 2011a; Laiolo and Obeso 2012; Campobello et al. 2015; Farine and Sheldon 2015a). These studies have provided important insights by quantifying how an individual's fitness can depend on the traits of conspecific neighbours or

2

social partners. In sociobiology, such traits are "social traits" that may evolve, in part, due to the helpful or harmful effects from others (or, equivalently, due to an individual's help or harm onto others; Hamilton, 1964). Selection analyses can help to understand the adaptive function of social traits and have the potential to bring together empirical and theoretical approaches in social evolution (Queller 1992; Frank 1997; Bijma and Wade 2008; Frank 2013; Goodnight 2013; McGlothlin et al. 2014).

There are two main approaches for measuring multilevel (or social) selection. The first, known as contextual analysis (Heisler and Damuth 1987; Goodnight et al. 1992), partitions individual fitness into those effects from a focal individual's own trait value and those from a collective trait of its whole group (e.g., the mean phenotype in the group, including the focal individual). The second approach, known as social selection analysis (Wolf et al. 1999), is similar but considers the "group" trait to be the mean phenotype of social partners, excluding the focal individual (see also Nunney 1985a,b; Okasha 2004). This is particularly useful when interactions occur in complex social networks, where group structure is not clearly defined (Formica et al. 2011a; McDonald et al. 2013), and when the traits of social partners directly cause fitness effects on a focal individual, whereas a whole-group trait does not (Okasha 2004, 2006). Moreover, the social selection approach leads to the simple result that fitness effects from social partners contribute to selection only when there is a phenotypic correlation between the traits of individuals and their social partners—that is, only when there is non-random phenotypic assortment (Wolf et al. 1999; McGlothlin et al. 2010, 2014). This is similar to the intuitive approach of kin selection, or inclusive fitness theory (Hamilton 1964, 1970), where kin-selected effects in social evolution occur only when there is non-random genetic assortment (relatedness).

3

Applications of the social selection approach in natural populations have only now started in earnest, with a study of body size in fungus beetles (Formica et al. 2011a) and a study of territory acquisition in songbirds (Farine and Sheldon 2015a). However, despite the potential benefits of this approach, there are already differences in how it is being applied. The key issue is that current applications use different measures of phenotypic assortment. In fact, there are many different ways of measuring assortment, particularly in the context of social network analysis (Leung and Chau 2007; Croft et al. 2008; Newman 2010; Farine 2014), and it is not clear which measure is relevant for social selection analysis. Moreover, assortment has been measured at different scales—across all individuals in a population and within local groups—and the strength of assortment can seem to depend on the scale at which it is measured (e.g., group size; Formica et al. 2011a). It has not been clear why this occurs or what it means for interpreting social selection.

Here, we address these issues using network simulations and data from Farine & Sheldon (2015a,b) and Formica et al. (2011a,b). To establish the measure of phenotypic assortment that correctly translates social effects into total phenotypic selection, we first outline the theory behind the social selection approach. We then use simulations and data to (i) ask how currently used measures of phenotypic assortment compare to the definition derived from theory; and (ii) explain how the interpretation of phenotypic assortment, and how to detect non-random assortment, will depend on the scale at which it is measured. We use our results to further address the links between social selection and kin selection theory and to provide a practical guide for social selection analyses.

The social selection approach

4

In this section, we outline the theory behind the social selection approach, as developed by Wolf et al. (1999; see also Okasha 2004; McGlothlin et al., 2010, 2014). We follow the convention in evolutionary quantitative genetics of defining an individual's fitness w as the number of offspring produced over a selective episode (W) relative to the population average ($\bar{W}_{pop}$; hence $w = W / \bar{W}_{pop}$; Lande and Arnold 1983). Social selection partitions fitness into those effects due the phenotypes of focal individuals, z, and those due to the mean phenotype of focal individuals' social partners, z'. In practice, the mean phenotype of social partners may be weighted by the frequency of interactions between different social partners, particularly in applications to animal social networks (Formica et al., 2011a; Farine and Sheldon, 2015a). More generally, social selection analysis is appropriate whenever a focal plant or animal's fitness may be influenced by the potentially correlated traits of its neighbours or social partners. The other main multilevel selection approach, contextual analysis, may be more appropriate when social groups are highly cohesive and have the potential to produce emergent, whole-group traits that affect the fitness of focal individuals in the group (Okasha 2004, 2006).

Fitness effects in social selection analysis are derived by least-squares multiple regression, following the multivariate selection approach of Lande & Arnold (1983). The relevant regression model predicts individual fitness as a function of the paired phenotypic variables (z, z') from all focal individuals in the population:

$$w = 1 + \beta_{wz|z'}(z - \bar{z}) + \beta_{wz'|z}(z' - \bar{z}') + \varepsilon \ , \qquad (1)$$

5

where $\bar{w}_{\text{pop}} = 1$ is the mean fitness of all focal individuals in the population; $\bar{z}$ is the mean

phenotype of those individuals; $\bar{z}'$ is the mean of all focal individuals' social environments

(i.e. the mean "group" trait across all focal individuals in the population); and ε is the residual

error. Fitness effects are measured by partial regression coefficients, where $\beta_{wz_iz'} \equiv \beta_N$ is the

independent effect of a focal individual's phenotype on its own fitness (non-social selection

gradient), and $\beta_{wz'_iz} \equiv \beta_S$ is the independent effect of the mean phenotype of focal

individuals' social partners on focal individual fitness (social selection gradient).

The key step in the social selection approach is to combine the fitness effects from

Equation 1 with a measure of phenotypic assortment. This follows from writing the within-

generation change in mean phenotype over an episode of selection (the selection differential),

using Price's (1970) equation, as $S = \text{cov}(w, z)$ or:

$$S = \beta_N \text{cov}(z, z) + \beta_S \text{cov}(z', z)$$
$$= (\beta_N + \beta_S \beta_{z'z}) \text{var}(z) \tag{2}$$

where, in the bottom line, $\text{var}(z)$ is the phenotypic variance, taken over all focal individuals,

and the quantity in parentheses is equal to the slope of a regression of individual fitness on

individual phenotype (the total selection gradient, β_{wz}). Specifically, the total selection

gradient is the sum of a non-social component, β_N, and a social component, $\beta_S \beta_{z'z}$, where $\beta_{z'z}$

$\equiv \beta_A$ is the slope of a regression of focal individuals' mean social partner phenotype on the

phenotype of focal individuals (herein the "assortment coefficient"; further details in the

6

Appendix). Crucially, this is the coefficient that correctly weights the social selection gradient so that the components of selection sum up to the total ($\beta_N + \beta_S\beta_A = \beta_{wz}$).

Thus, Equation 2 shows how to correctly measure phenotypic assortment for the analysis of social selection. Researchers may typically want to know the components of the total selection gradient and will therefore want to measure the assortment coefficient, β_A. Alternatively, as in the top line of Equation 2, assortment can be measured as the "interactant covariance", cov(z', z), taken over all individuals in the population (Wolf *et al.*, 1999). Moreover, when the selection gradients in Equation 2 are calculated using variance-standardized phenotypes (i.e. standardized to a mean of 0 and a variance of 1; herein denoted z^* and z'^*, respectively), the corresponding assortment coefficient $\beta_{A*} \equiv \beta_{z'*z*}$ will equal both the interactant covariance (cov[z'^*, z^*]) and the Pearson correlation between focal individual and social partner phenotypes ($r_{z'*z*} \equiv$ cov[z'^*, z^*]/[$\sigma_{z*}\sigma_{z'*}$]). The important point about all of these measures is that, when used in the appropriate version of Equation 2, they provide the correct weighting of the social selection gradient—that is, they make the social selection approach work.

How (not) to measure phenotypic assortment

In this section, we ask whether currently used measures of phenotypic assortment match the definition derived from theory. In the first application, Formica et al. (2011a) analyzed variance-standardized trait values and measured assortment using the Pearson correlation between focal individual phenotypes and the (weighted) mean phenotypes of social partners. Hence, this is a correct measure of assortment for social selection analysis (equivalent to β_{A*}). However, Formica et al. (2011a) also calculated assortment at multiple scales, both

7

across all individuals in the population (as in the theory above) and within each individual group separately. We address the consequences of this group-level measure of assortment in the following section. In the second application, Farine & Sheldon (2015a) also analyzed standardized trait values but used a common network measure of assortment, known as Newman's assortativity (Newman 2002, 2003; herein "network assortment", denoted r_N). Although network assortment is also based on a Pearson correlation coefficient, we show in the Appendix that it does not, in general, match any version of assortment from Equation 2. Below, we use a simple example to explain why this is the case.

Network assortment does not generally match an assortment measure from social selection theory because the network measure considers a different set of paired variables, as illustrated in Figure 1. Whereas a social selection analysis considers the relationship between a focal individual's phenotype and the mean phenotype of its social partners (z and z', where each focal individual yields one row of data; Fig. 1b), network assortment considers the relationship between a focal individual's phenotype and the phenotype of each its social partners individually (we call these variables z_N and z'_N, where each focal individual can yield multiple rows of data; Fig. 1c). As a simple example, we consider the network of four individuals (A, B, C and D) in Figure 1a, assigning individuals A to D a trait value (z) of 1 to 4, respectively. Then, using the appropriate matrices in Fig. 1b and c (and using both non-standardized and variance-standardized trait values), we calculate the following measures of assortment: $\text{cov}(z', z) = -0.19$; $\beta_A = -0.15$; $\beta_{A*} = -0.23$; and $r_N = -0.26$ (note that r_N does not vary with standardization of trait values). Hence, in this example, network assortment is not equivalent to any measure of assortment from social selection theory.

8

To examine these differences more generally, we compared network assortment (r_N) and the assortment coefficient (β_A or β_{A*}) across a range of realistic social networks. Using previously published networks of behavioral interactions from 172 animal groups (Shizuka and McDonald 2015a,b), we randomly assigned a trait value to each individual, generating networks with random variation in the level of assortment, and repeated this 100 times for each network (see Supplementary Information for simulation details). Using these networks, we found: (i) when traits are unstandardized, r_N and β_A roughly correspond, with modest differences (Fig. 2a); (ii) when traits are variance-standardized, the differences are more extreme, with r_N consistently underestimating the strength of assortment relative to β_{A*} (Fig. 2b); and (iii) even the sign of r_N can differ from that of β_A and β_{A*}, implying that using network assortment can predict the wrong direction of social selection (top left and bottom right quadrants of Fig. 2a and b).

Thus, our results imply that using network assortment for social selection analysis will ultimately lead to an incorrect estimate of how fitness effects from social partners contribute to selection. As a concrete example of this, we reanalyzed data from Farine & Sheldon (2015a,b), using both network assortment and the regression measures of assortment from social selection theory. This shows that, as expected, network assortment consistently misestimates the social component of selection (particularly when using variance-standardized data; Table 1). Instead, we recommend using a regression of the mean phenotype of social partners (which can be weighted by association or interaction strengths) on focal individual phenotypes (β_A). This provides the correct weighting of the social selection gradient (β_S) and hence the correct contribution of social selection to the total selection gradient.

9

How to interpret assortment at different scales

In this section, we explain how the interpretation of phenotypic assortment, and how to detect non-random assortment, can depend on the scale at which it is measured. Ultimately, evolution occurs at the population level, and so this is the level at which selection and assortment should typically be measured (as in the theory above). Yet it may also be of interest to measure selection at a more local scale—for example, to ask how local ecological variation influences selection (Arnqvist 1992; Kasumovic et al. 2008) and/or to examine selection occurring in small competitive arenas within a larger population, as in Formica et al.'s (2011a) study of body size in male beetles. However, this study also raises the question of how to interpret measures of assortment at different scales: when measured across all groups in the population, Formica et al. (2011a) found relatively weak assortment ($\beta_{A*} = -0.158$). In contrast, when measured within groups (i.e., using the [$z*$, $z'*$] data from each group separately), they found much stronger negative assortment, particularly in the smallest groups (see data in Fig. 3d). Formica et al. (2011a) speculated that, in these groups, large males had strong preferential associations with smaller males, causing social selection on body size. Here, we re-evaluate this hypothesis and the interpretation of phenotypic assortment in this example.

Our approach follows Farine & Sheldon (2015a) and other social network analyses, using null models to ask whether observed measures of phenotypic assortment deviate from expectations under completely random interactions (e.g., no preferential associations) (Croft et al. 2008, 2011; Farine and Whitehead 2015). It is already well known that, even when social interactions are completely random, there are mechanisms that can give rise to phenotypic covariance between the traits of social partners (Wolf et al. 1999). In particular,

10

there may be behavioral modification of social partners by a focal individual and/or "indirect genetic effects" (IGEs), where genes expressed in one individual have phenotypic effects on another (Moore et al. 1997). The aim of our null model is to show that, even in the absence of preferential associations, behavioral modification, or IGEs, non-zero measures of phenotypic assortment can still arise, simply owing to the measurement of assortment coefficients at a local scale (e.g., a small group or subset of a larger population).

We first derive a null model for an assortment coefficient measured within a group with completely random interactions, assuming that focal individuals interact randomly and equally with all other individuals in the group. In the Appendix, we show that the predicted assortment coefficient is $\beta_A = -1/(n-1)$, where n is the number of focal individuals (herein "group size"). Hence, even with completely random social interactions, the expected assortment coefficient will always be negative. This is necessarily true because, in this null model, the assortment coefficient is measured with respect to the mean phenotype of all focal individuals under consideration—in this case, the mean phenotype of the whole group ($\beta_A = [z' - \overline{z}]/[z - \overline{z}]$; see Appendix). For example, if a focal male beetle is larger than the mean of his whole group (including himself), then all other males in the group must necessarily be smaller than the mean and vice versa (implying a negative regression slope). In any large group (large n) this effect will be negligible because a single focal male will have a negligible contribution to the mean phenotype in his group; in contrast, it becomes increasingly important as group size declines and a focal male has an increasingly large contribution to the mean phenotype in his group.

To further explore the null model, we simulated random social networks (groups) and calculated the assortment coefficient using both unstandardized and variance-standardized

11

trait values (yielding β_A and β_{A*} respectively). We varied both the number of individuals in the group and the "density" of social interactions—the proportion of interacting pairs in the group out of the total possible number of pairs—and generated 1000 networks for each group size and density combination (see Supplementary Information for simulation details). Using these random groups, we found: (i) when traits are unstandardized, β_A matches the $-1/(n-1)$ prediction, on average, irrespective of interaction density (Fig. 3a); and (ii) when traits are standardized, β_{A*} depends on both group size and interaction density—in particular, when interaction density is high, negative assortment coefficients can arise across a wide range of group sizes (Figure 3b).

Returning to the results from Formica et al. (2011a,b), we now explicitly test whether their observed assortment measures differ from random. Formica et al. (2011a) measured body size assortment in six groups that varied in the number of competing males. Figures 3c & d show that these assortment measures, expressed as either β_A or β_{A*}, are increasingly negative as group size declines, consistent with our null models. To formally test for non-random assortment, we generated randomized versions of the observed social groups using 'node-based' randomizations (Croft et al. 2008, 2011; Farine and Whitehead 2015). This process maintains the observed structure of the social network but randomly swaps the attributes of the nodes in the network (i.e. the patterns of interactions are held constant but the traits of focal individuals are randomized across individuals; full details in Supplementary Information, Figure S1.) We generated 1000 randomizations for each group, yielding distributions of β_A and β_{A*} to compare with the observed values. Figures 3c & d show that out of the six groups, only one observed measure of assortment was significantly lower than the null expectation (more extreme than 95% of the simulated distribution). Hence, there is

12

only little evidence for non-random social interactions. Instead, negative assortment is largely accounted for by variation in group size.

Thus, our results imply that measuring selection and phenotypic assortment at the group level can confuse the selective effects of group size and non-random associations, making interpretation difficult. This is because the observed assortment coefficients in small groups are likely to be strongly negative—even when social interactions are completely random—simply owing to small group size. Hence, whenever social selection is measured at a small scale, it is important is to evaluate the observed patterns against a suitable null model with random interactions. Taking the example of body size in male beetles, it can appear that large size was favored by social selection in small groups, driven by large males actively seeking out smaller males and avoiding the harmful effects from larger rivals (given $\beta_S < 0$; Formica et al. 2011a). Our reanalysis, showing little evidence for such non-random associations, instead implies that relatively large individuals simply had higher fitness when competing in small groups. Some may still see this benefit of large size in small groups, measured by $-1/[n-1]\beta_S > 0$ in the group-level analysis, as a form of social selection. However, a selection analysis in any large population of groups (where $-1/[n-1] \approx 0$) would capture the same group-size-dependent benefit as a part of non-social selection (β_N). For this reason, we suggest that most of what appears as social selection in the group-level analyses from Formica et al. (2011a) could instead be interpreted as a non-social benefit of large body size. Indeed, this interpretation reconciles with the population-level analysis, where larger body size was primarily favored by non-social selection (Formica et al. 2011a).

Links to kin selection theory

13

There are many similarities between the social selection approach and kin selection theory (recently reviewed by McGlothlin et al. (2014)). In particular, the total selection gradient from Equation 2 is analogous to the left-hand side of Hamilton's rule for the evolution of a social trait, $-C + BR > 0$ (Hamilton 1964). As in the social selection approach, Hamilton's rule is the sum of a non-social component ($-C$; direct effect of a focal individual's trait on its own fitness) and a social component that is weighted by an assortment coefficient (BR; indirect or kin-selected effect, where B is a fitness effect on [or from] social partners, and R is the genetic relatedness between social partners). Moreover, all components of Hamilton's rule can be measured by regression coefficients (Queller, 1992; Gardner et al. 2011). Hence, not surprisingly, our two main results above have analogies in kin selection theory: (i) relatedness is the slope of a regression of the average genetic value of social partners on focal individual genetic value (e.g., Hamilton 1972; Queller 1992; Frank 1997; Gardner et al. 2011)); and (ii) this regression measure of relatedness will be negative in any well-mixed population or group with completely random social interactions (e.g., Hamilton 1970, 1971; Grafen 1985; Pepper 2000; Gardner and West 2004; Bijma and Wade 2008).

A major difference between the approaches, however, is that social selection uses phenotypic predictors of fitness (z and z') whereas the strict version of Hamilton's rule uses genetic predictors only (alleles or breeding values; Gardner et al. 2011). It follows that phenotypic assortment includes, in addition to genetic relatedness, many possible sources of correlation between social partners, including: shared environmental conditions, reciprocity ("tit for tat" behavior), social partner manipulation, and IGEs (Wolf et al. 1999; McGlothlin et al. 2014). This is potentially problematic when it comes to classifying interactions according to their adaptive function, which is typically based on the genetic version of

14

Hamilton's rule (selfish: $-C > 0$, $B < 0$; mutually beneficial: $-C > 0$, $B > 0$; altruistic: $-C < 0$, $B > 0$; spiteful: $-C < 0$, $B < 0$) (Hamilton 1964; West et al. 2007). For example, reciprocated cooperation (tit-for-tat) among non-relatives is considered a mutually beneficial trait that is favored owing to a net direct fitness benefit to a focal cooperator ($-C > 0$, $RB = 0$; West et al. 2007). In a social selection analysis, however, this same net direct fitness benefit would be the sum of a negative non-social component of selection ($\beta_N < 0$) and a positive social component, measuring the fitness benefit coming back to the focal individual ($\beta_S\beta_A > 0$, given $\beta_S > 0$ and $\beta_A > 0$). Naively equating the terms of Hamilton's rule and social selection would, in this case, lead to an incorrect diagnosis that the helping behavior is altruistic, with a direct fitness cost (Wyatt et al. 2013). Hence, this example shows that the components of kin selection and social selection theory do not always match.

So why bother with the social selection approach? First, theoretical treatments of social selection are particularly useful for partitioning the various mechanisms by which social interactions contribute to social evolution, including the relative roles of relatedness and IGEs in kin selection (Moore et al. 1997; Wolf et al. 1999; McGlothlin et al. 2010; McGlothlin et al., 2014). Second, in empirical studies of social selection, field researchers do not usually have information about the underlying genetics, and so the phenotypic approach of social selection (from Eq. 2) can provide valuable insights. In particular, the approach can generate alternative hypotheses about the mechanisms by which social interactions contribute to the total selection for a trait (McGlothlin et al. 2014). Moreover, social selection and kin selection approaches agree on the key result that if assortment—phenotypic or genetic—is effectively random (β_A or $R \approx -1/[n-1]$), then natural selection will be driven primarily by non-social effects. Hence, given the correct methods for measuring phenotypic assortment

15

and assessing departures from random, social selection analyses can address the fundamental question of whether or not fitness effects from social partners contribute at all to the total selection on a trait.

Practical guide for social selection analyses

In sum, we argue that the social selection approach can be extremely useful and should have widespread applications in plant and animal biology. However, the approach has its limitations, and there have been inconsistencies in the measurement and interpretation of a key component, phenotypic assortment. Our work resolves these inconsistencies and leads to the following practical guide.

1. To correctly weight the social selection gradient and sum up the components of selection, it is critical to incorporate a suitable measure of phenotypic assortment. Typically, this will be a regression of focal individuals' social environments (the mean or weighted mean of social partner phenotypes) on focal individual phenotypes, measured across all individuals in the population (i.e., the level at which evolution occurs; as in Equation 2). Although network assortment (Newmann's assortativity) is useful for many questions in social network analysis (Krause et al. 2011; Farine and Whitehead 2015; Farine et al. 2015), we have shown that it is not a suitable measure of assortment for social selection analysis.

2. It is important to recognize that if assortment is measured at small scale (e.g., within a local group or some other subset of the population), completely random social interactions will generate non-trivial negative measures of assortment that reflect small group size only.

16

To test for non-random assortment in these cases, we advise randomization or permutation tests to compare empirical measures of assortment with a realistic null model with random interactions (see Farine and Whitehead 2015; Farine 2017). We also advise careful interpretation of selection components inferred from a small-scale social selection analysis. For example, if assortment and selection are measured within small subsets of a larger population, then the conclusions drawn from group-level analyses should be reconciled with those from a population-level social selection analysis.

3. When non-random phenotypic associations are detected, the cause of this assortment will usually be unknown. In this case, a social selection analysis should lead to alternative hypotheses about the possible environmental, genetic, or behavioral mechanisms that underlie phenotypic assortment (McGlothlin et al. 2014). Before knowing these mechanisms, care should be taken not to equate the terms of social selection and Hamilton's rule, which is typically used to classify the adaptive function of social traits. There has been much debate and subsequent clarification about the semantics of social adaptation (Foster et al. 2006a,b; West et al. 2007), and we hope that the social selection approach can proceed without adding further confusion.

Literature Cited

Arnqvist, G. 1992. Spatial variation in selective regimes: sexual selection in the Water strider, *Gerris odontogaster*. Evolution 46:914–929.

Aspi, J., A. Jäkäläniemi, J. Tuomi, and P. Siikamäki. 2003. Multilevel phenotypic selection on morphological characters in a metapopulation of Silene tatarica. Evol. Int. J. Org. Evol. 57:509–517.

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.

Campobello, D., J. F. Hare, and M. Sarà. 2015. Social phenotype extended to communities: Expanded multilevel social selection analysis reveals fitness consequences of interspecific interactions. Evolution 69:916–925.

Croft, D., R. James, and J. Krause. 2008. Exploring animal social networks. Princeton University Press, Princeton, NJ.

Croft, D. P., J. R. Madden, D. W. Franks, and R. James. 2011. Hypothesis testing in animal social networks. Trends Ecol. Evol. 26:502–507.

Donohue, K. 2004. Density-dependent multilevel selection in the Great lakes sea rocket. Ecology 85:180–191.

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.

Farine, D. R. 2017. A guide to null models for animal social network analysis. Methods Ecol. Evol., doi: 10.1111/2041-210X.12772.

Farine, D. R. 2014. Measuring phenotypic assortment in animal social networks: weighted associations are more robust than binary edges. Anim. Behav. 89:141–153.

18

Farine, D. R., P.-O. Montiglio, and O. Spiegel. 2015. From Individuals to Groups and Back: The Evolutionary Implications of Group Phenotypic Composition. Trends Ecol. Evol. 30:609–621.

Farine, D. R., and B. C. Sheldon. 2015a. Selection for territory acquisition is modulated by social network structure in a wild songbird. J. Evol. Biol., doi: 10.1111/jeb.12587.

Farine, D. R., and B. C. Sheldon. 2015b. Data from: Selection for territory acquisition is modulated by social network structure in a wild songbird. Dryad Digit. Repos., doi: http://dx.doi.org/10.5061/dryad.bn27c

Farine, D. R., and H. Whitehead. 2015. Constructing, conducting and interpreting animal social network analysis. J. Anim. Ecol. 84:1144–1163.

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

Formica, V. A., J. W. McGlothlin, C. W. Wood, M. E. Augat, R. E. Butterfield, M. E. Barnard, and E. D. Brodie III. 2011b. Data from: Phenotypic assortment mediates the effect of social selection in a wild beetle population. Dryad Digit. Repos., doi: http://dx.doi.org/10.5061/dryad.2r4gj.

Foster, K. R., T. Wenseleers, and F. L. W. Ratnieks. 2006a. Kin selection is the key to altruism. Trends Ecol. Evol. 21:57–60.

Foster, K. R., T. Wenseleers, F. L. W. Ratnieks, and D. C. Queller. 2006b. There is nothing wrong with inclusive fitness. Trends Ecol. Evol. 21:599–600.

Frank, S. A. 1997. Multivariate Analysis of Correlated Selection and Kin Selection, with an ESS Maximization Method. J. Theor. Biol. 189:307–316.

19

Frank, S. A. 2013. Natural selection. VII. History and interpretation of kin selection theory. J. Evol. Biol. 26:1151–1184.

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

Gardner, A., S. A. West, and G. Wild. 2011. The genetical theory of kin selection. J. Evol. Biol. 24:1020–1043.

Goodnight, C. 2013. On Multilevel Selection and Kin Selection: Contextual Analysis Meets Direct Fitness. Evolution 67:1539–1548.

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. Pp. 28–89 *in* Oxford Surveys in Evolutionary Biology. Oxford University Press, Oxford, UK.

Hamilton, W. D. 1972. Altruism and Related Phenomena, Mainly in Social Insects. Annu. Rev. Ecol. Syst. 3:193–232.

Hamilton, W. D. 1971. Selection of selfish and altruistic behavior in some extreme models. Pp. 57–91 *in* Man and beast: comparative social behavior. Smithsonian Institution Press, Washington, D.C.

Hamilton, W. D. 1970. Selfish and Spiteful Behaviour in an Evolutionary Model. Nature 228:1218–1220.

Hamilton, W. D. 1964. The genetical evolution of social behaviour, I & II. J. Theor. Biol. 7:1–52.

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

20

Kasumovic, M. M., M. J. Bruce, M. C. B. Andrade, and M. E. Herberstein. 2008. Spatial and temporal demographic variation drives within-season fluctuations in sexual selection. Evolution 62:2316–2325.

Krause, J., A. D. M. Wilson, and D. P. Croft. 2011. New technology facilitates the study of social networks. Trends Ecol. Evol. 26:5–6.

Laiolo, P., and J. R. Obeso. 2012. Multilevel selection and neighbourhood effects from individual to metapopulation in a wild passerine. PloS One 7:e38526.

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

Leung, C. C., and H. F. Chau. 2007. Weighted assortative and disassortative networks model. Phys. Stat. Mech. Its Appl. 378:591–602.

McDonald, G. C., R. James, J. Krause, and T. Pizzari. 2013. Sexual networks: measuring sexual selection in structured, polyandrous populations. Philos. Trans. R. Soc. B Biol. Sci. 368:1–10.

McGlothlin, J. W., A. J. Moore, J. B. Wolf, and E. D. Brodie III. 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. Lond. B Biol. Sci. 369:20130358.

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

Newman, M. 2010. Networks: An Introduction. OUP Oxford.

Newman, M. E. J. 2002. Assortative mixing in networks. Phys Rev Lett 89:208701.

21

Newman, M. E. J. 2003. Mixing patterns in networks. Phys. Rev. E 67:26126.

Nunney, L. 1985a. Female-Biased Sex Ratios: Individual or Group Selection? Evolution 39:349–361.

Nunney, L. 1985b. Group selection altruism and structured-deme models. Am. Nat. 126:212–230.

Okasha, S. 2006. Evolution and the Levels of Selection. Oxford University Press.

Okasha, S. 2004. Multilevel selection and the partitioning of covariance: a comparison of three
approaches. Evolution 58:486–494.

Pepper, J. W. 2000. Relatedness in trait group models of social evolution. J. Theor. Biol. 206:355–
368.

Price, G. R. 1970. Selection and covariance. Nature 227:520–21.

Queller, D. C. 1992. A General Model for Kin Selection. Evolution 46:376–380.

Shizuka, D., and D. B. McDonald. 2015a. Data from: The network motif architecture of dominance
hierarchies. Dryad Digit. Repos., doi: http://dx.doi.org/10.5061/dryad.f76f2.

Shizuka, D., and D. B. McDonald. 2015b. The network motif architecture of dominance hierarchies. J.
R. Soc. Interface 12:20150080.

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

Tsuji, K. 1995. Reproductive conflicts and levels of selection in the Ant *Pristomyrmex pungens*:
contextual analysis and partitioning of covariance. Am. Nat. 146:586–607.

Weinig, C., J. A. Johnston, C. G. Willis, and J. N. Maloof. 2007. Antagonistic multilevel selection on
size and architecture in variable density settings. Evolution 61:58–67.

22

West, S. A., A. S. Griffin, and A. Gardner. 2007. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. J. Evol. Biol. 20:415–432.

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

Wyatt, G. a. K., S. A. West, and A. Gardner. 2013. Can natural selection favour altruism between species? J. Evol. Biol. 26:1854–1865.

Table 1. Network assortment leads to incorrect estimates of social selection. Using data on the winter arrival time of great tits, from Farine & Sheldon (2015a,b), we performed social selection analyses using alternative measures of phenotypic assortment—the assortment coefficient from social selection theory (β_A for unstandardized data, in [a]; β_{A*} for variance-standardized data, in [b]) and network assortment (Newmann's assortativity, r_N), calculated using weighted networks (Farine & Sheldon 2015a). The "% change in social component" column shows how much the social component of selection would differ by using network assortment instead of the correct measure from social selection theory ($[\beta_S r_N - \beta_S \beta_A]/ \beta_S \beta_A$ in [a] or $[\beta_S r_N - \beta_S \beta_{A*}]/ \beta_S \beta_{A*}$ in [b]).

(a) Non-standardized data

Year	β_N	β_S	β_A	r_N	% change in social component
2012	-0.036	0.055	0.251	0.288	+ 14.7
2013	-0.020	0.043	0.170	0.188	+ 10.6

23

Year	β_N	β_S	β_{A*}	r_N	% change in social component
2014	-0.021	0.030	0.298	0.352	+ 18.1

(b) Variance-standardized data

Year	β_N	β_S	β_{A*}	r_N	% change in social component
2012	−0.271	0.168	0.628	0.288	− 54.1
2013	−0.157	0.104	0.555	0.188	− 66.1
2014	−0.171	0.112	0.652	0.352	− 46.0

Notes: In our analyses, fitness (w) is an individual's territory establishment success measured relative to the population average success. We calculated non-social and social selection gradients (β_N and β_S, respectively) as a function of individuals' winter arrival time (z) and the weighted mean arrival time of their social partners (z'). All analyses were conducted using data from the whole population (including both first year great tits and older birds) and using least-squares regression. Our methods and selection estimates therefore differ slightly from the original study (Farine & Sheldon [2015a] analysed selection on first year birds only and used a binomial regression rather than least-squares regression).

Figure legends

Figure 1. Comparing the calculation of the assortment coefficient, from social selection theory, and network assortment (Newman's assortativity). (a) Network and sociomatrix representations of a social network of four individuals (A, B, C, and D). Social interactions are represented by links in the network representation and denoted as "1"s in the sociomatrix.

24

(b & c) The set of paired variables used for the calculation the assortment coefficient (β_A or β_{A*}; panel b) and for the calculation of network assortment (r_N; panel c).
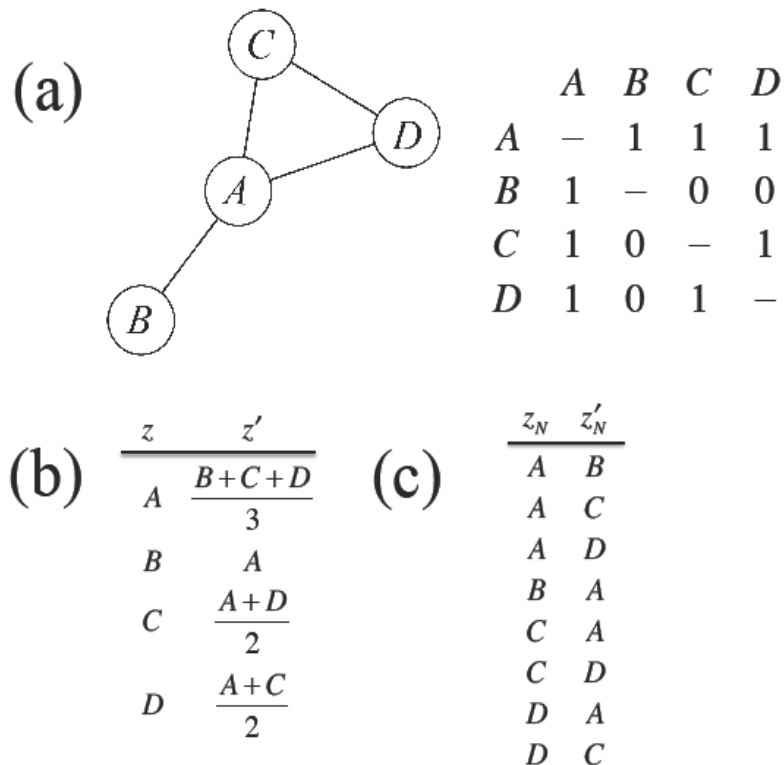

(a)

	A	B	C	D
A	–	1	1	1
B	1	–	0	0
C	1	0	–	1
D	1	0	1	–

(b)

	z	z'
A		$\dfrac{B+C+D}{3}$
B		A
C		$\dfrac{A+D}{2}$
D		$\dfrac{A+C}{2}$

(c)

z_N	z'_N
A	B
A	C
A	D
B	A
C	A
C	D
D	A
D	C

Figure 2. Network assortment (r_N) is not generally equivalent to the assortment coefficient from social selection theory (β_A or β_{A*}). (a & b) Regression measures of assortment were calculated using non-standardized trait values (yielding β_A; panel a) or variance-standardized trait values (yielding β_{A*}; panel b). In both panels, we generated 17200 random networks using empirical interactions networks from 172 animal groups (Shizuka and McDonald 2015a,b). Note that network assortment does not change with data standardization.
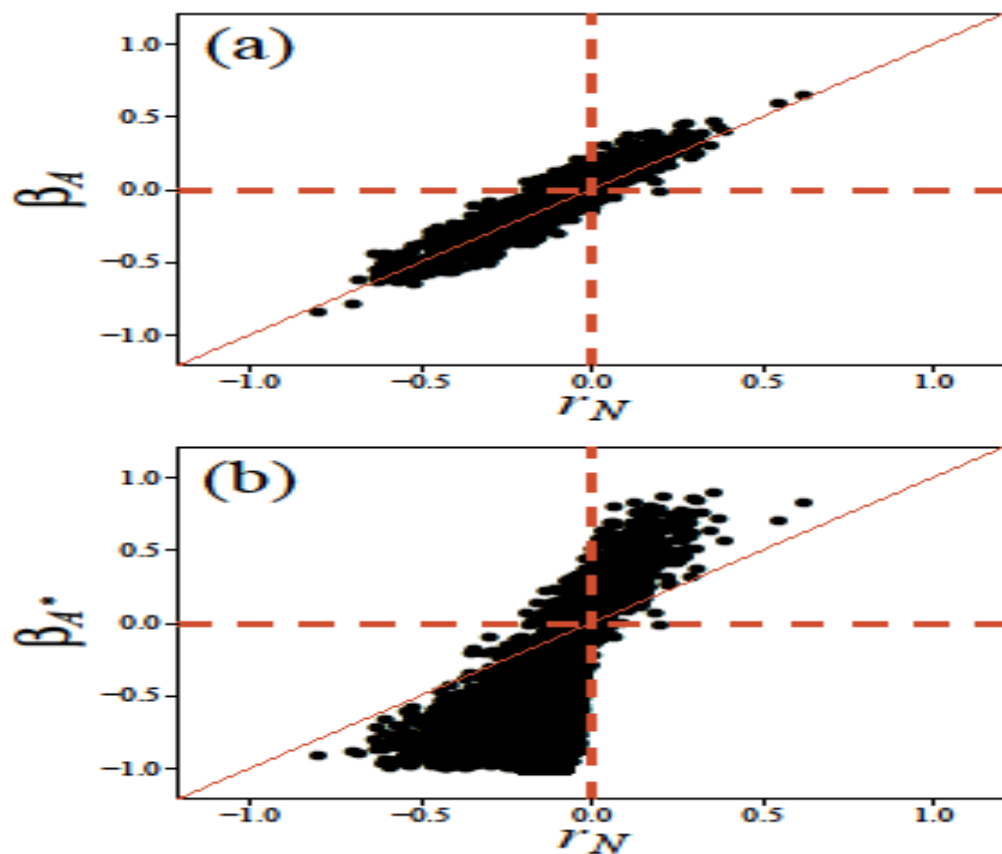
Figure 3. Comparing null models of phenotypic assortment with empirical data. (a & b) The

relationship between group/network size (number of individuals) and the assortment

coefficient, measured with non-standardized trait values (yielding β_A; panel a) or variance-

standardized trait values (yielding β_{A*}; panel b). Points show means ± 1 standard error, and

the numbers in (b) denote the interaction density (number of interacting pairs / total number

of possible interacting pairs). (c & d) Points show empirical estimates of body size

assortment measured within groups of male beetles, using non-standardized data (yielding

β_A) or variance-standardized data (yielding β_{A*}) from Formica et al. (2011a,b). Error bars

bound the 95% range of simulated assortment coefficients from 'node-based' randomization

of the empirical data (see Supplemental Information for full details). The dotted lines in

panels (a) and (c) represent the null prediction $\beta_A = -1/(n-1)$, where n is the number of individuals in the group/network.
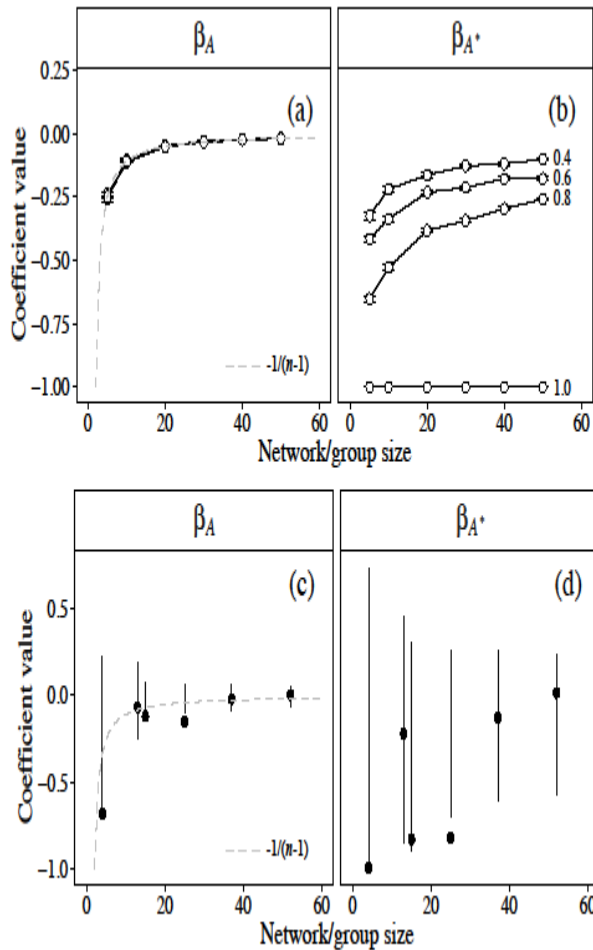
Appendix

Further details of the phenotypic assortment coefficient (β_A)

Here, we give further description of the phenotypic assortment coefficient and derive a null model that assumes completely random social interactions (i.e., no assortment based on phenotype). As described in the main text, the assortment coefficient is the slope from a least-

27

squares regression of focal individuals' mean social partner phenotype (z') on the phenotype of focal individuals (z). Hence, according to the standard description of a regression slope, the assortment coefficient is

$$\beta_A = \frac{\text{cov}(z, z')}{\text{var}(z)} \, , \qquad \qquad \text{(A1)}$$

where the covariance and variance are taken over all focal individuals under consideration (e.g., in a population or local group). Another way to write the assortment coefficient is to start with the relevant regression equation, predicting mean social partner phenotype as a function of focal individual phenotype: $z' = \overline{z}' + \beta_A(z - \overline{z}) + \varepsilon$. We rearrange this to yield

$$\beta_A = \frac{z' - \overline{z}'}{z - \overline{z}} + \frac{\varepsilon}{z - \overline{z}} \, , \qquad \qquad \text{(A2)}$$

where the second term will be zero because ε is, on average, zero. Hence, Equation A2 implies that the assortment coefficient can be written as $\beta_A = (z' - \overline{z}')/(z - \overline{z})$.

We derive a null model from Equation A2, assuming that a focal individual interacts randomly and equally with all other individuals in its population or group. In this case, the mean of all focal individuals' social environments will equal the mean phenotype of all focal individuals ($\overline{z}' = z$). Substituting this into Equation A2, and dropping the second term, gives

28

$$\beta_A = \frac{z' - \bar{z}}{z - \bar{z}} \; . \tag{A3}$$

Moreover, in this null model, the mean phenotype of social partners (z') will be the mean of a focal individuals' population/group mates, and so the mean of the entire population/group can be written as $\bar{z} = (1/n)z + ([n-1]/n)z'$, where n is the total number of focal individuals. Substituting this into Equation A3 yields the null prediction $\beta_A = -1/(n-1)$, as given in the main text. When measuring phenotypic assortment across all individuals in a large population (large n), the null prediction will be close to zero ($\beta_A \approx 0$). In contrast, when measuring assortment at a small scale (e.g., within a local group with small n), the null prediction will take a non-trivial negative value.

Further details of network assortment (Newman's assortativity; r_N)

Here, we give further description of network assortment (Newmann's assortativity; r_N) and compare it with the assortment coefficient from social selection theory. Network assortment is the Pearson correlation coefficient between the traits at either end of an edge (i.e. the link between one individual and a social partner), across all edges in a social network. A standard formal definition is

$$r_N = \frac{\sum_i^M j_i k_i - M^{-1}\sum_i^M j_i \sum_i^M k_i}{\sqrt{\left[\sum_i^M j_i^2 - M^{-1}(\sum_i^M j_i)^2\right]\left[\sum_i^M k_i^2 - M^{-1}(\sum_i^M k_i)^2\right]}} \; , \tag{A4}$$

29

where M is the total number of edges in the network and j and k can be interpreted as the phenotypic values of individuals at either end of the i_{th} edge (Newman, 2002, 2003). In the context of social selection, the total number of edges on a network is twice the number of interacting pairs, because inclusion in a social group is reciprocal (i.e. if individual A shares a link a social link with individual B, then individual B also shares a social link with A; Newman, 2003). For example, in the network of four individuals (A to D) in Figure 1, there are four interacting pairs (i.e. A-B, A-C, A-D, C-D) and therefore eight edges (i.e., eight rows of data) used in the calculation of network assortment. A similar approach can be used to calculate assortativity on networks with edges weighted by interaction frequency/strength (Farine & Sheldon 2015a).

To compare network assortment with the assortment coefficient from social selection theory, it is helpful to write network assortment in a simpler form. To do this, we first note that the j and k from Equation A3 correspond to the vectors z_N and z'_N, respectively, illustrated in Figure 1c in the main text. Furthermore, Equation A3 is a Pearson coefficient for the correlation between z_N and z'_N (i.e., the correlation between a focal individual's phenotype and the phenotype of each its social partners individually, across all focal individuals). This correlation coefficient can be written as

$$r_N = \frac{\mathrm{cov}(z_N, z'_N)}{\sqrt{\mathrm{var}(z_N)\mathrm{var}(z'_N)}} \ . \tag{A5}$$

30

Comparing Equations A5 and A1, it is clear that network assortment will not generally equal the assortment coefficient from social selection theory, except under special circumstances. For example, if every focal individual interacts with only one other focal individual, then $z_N = z$, $z'_N = z'$, and network assortment will match all versions of the assortment coefficient ($r_N = \beta_A = \beta_{A*}$). At the other extreme where every individual interacts equally with every other individual, $\text{cov}(z_N, z'_N) = \text{cov}(z, z')$ and $\sqrt{\text{var}(z_N)\text{var}(z'_N)} = \text{var}(z)$, so in this case network assortment will match the unstandardized assortment coefficient ($r_N = \beta_A$).