Nepotism masks evidence for reciprocity in cooperation networks

Gerald G. Carter[1,2], Gabriele Schino[3], Damien Farine[1,2]

[1] Department of Collective Behaviour, Max Planck Institut für Ornithologie, Konstanz, Germany;

[2] Chair of Biodiversity and Collective Behaviour, Department of Biology, University of Konstanz, Germany;

[3] Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, Rome, Italy

Keywords: cooperation, kin discrimination, mandrills, macaques, vampire bats

The authors wish to be identified to the reviewers.

Abstract

Nepotism and reciprocity are not mutually exclusive explanations for cooperation; helping decisions can depend on both kinship cues and past reciprocal help. The importance of these two factors can therefore be difficult to disentangle with observational data. We developed a resampling procedure for inferring the statistical power to detect observational evidence of nepotism and reciprocity, and applied this procedure to simulated and real datasets. We simulated datasets resulting from perfect reciprocity, where the probability and duration of helping events from individual A to B equaled B to A. We also simulated varying degrees of simultaneous nepotism. We then assessed how nepotism and sampling effort influenced the probability of detecting evidence of reciprocity. We applied the same analysis to empirical data on food sharing in vampire bats and allogrooming in mandrills and Japanese macaques. Nepotism consistently masked evidence for reciprocity. With perfect reciprocity and imperfect nepotism, nepotism was more likely to be detected and overestimated. We explain the causes and consequences. To compare the relative importance of genetic and social ties, researchers should measure the relative reliability of both estimates. We provide R scripts to allow others to assess the reliability of kinship and reciprocal help estimates in their own datasets.

Introduction

An adaptive cooperative trait, such as a heritable propensity for helping others, is an investment that yields a net return for the actor's inclusive fitness by increasing direct fitness, indirect fitness, or both (Hamilton 1964; West et al. 2007a). In cooperative groups with limited dispersal, such as microbial biofilms and eusocial insect colonies, groupmates will typically be close kin and selection can favor indiscriminate cooperative investments (Cornwallis et al. 2009; Hamilton 1964; Kümmerli et al. 2009; Lehmann et al. 2007). In well-mixed populations however, indiscriminate cooperation will lead to less cooperative phenotypes gaining the fitness benefits of receiving help while paying lower costs by helping less, and this exploitation can result in an evolutionary tragedy of the commons (West et al. 2007a). Individuals can ensure indirect or direct fitness returns on their investments by preferentially helping closer kin, *nepotism* (e.g. Cornwallis et al. 2009; Griffin and West 2003) or more cooperative partners, *reciprocity* (Carter 2014; Dolivo and Taborsky 2015; e.g. Rutte and Taborsky 2008; Schino and Aureli 2017; Taborsky et al. 2016; Trivers 1971). Crucially, these strategies can coexist if helping decisions are influenced by both kinship and past experience of reciprocal help, and they can also interact if, for instance, the degree of reciprocity is greater or smaller among kin (Axelrod and Hamilton 1981; Trivers 1971; Van Cleve and Akçay 2014). Evidence for nepotism and reciprocity can therefore be difficult to disentangle. In this paper we show that one should expect asymmetries in the accuracy of the estimates of kinship and helping rates, and these biases will lead nepotism to mask evidence of even perfect reciprocity. Given this phenomenon, we provide methods that help with interpreting the effects of kinship and reciprocal help in cooperation networks.

The claim of reciprocity is far more contentious than nepotism for any nonhuman species. Despite growing correlational and experimental evidence, there is still disagreement about whether reciprocity regulates cooperation in any nonhuman animal (Carter 2014; Clutton-Brock 2009; Schino and Aureli 2017; Taborsky et al. 2016), but a primary reason for this debate is that authors do not agree on what the term 'reciprocity' means or should mean (Bshary and

Bergmüller 2008; Carter 2014; Lehmann and Keller 2006; Noë 2006; West et al. 2007b). Definitions of reciprocity (also called reciprocation, reciprocal altruism, reciprocal cooperation, contingent cooperation, and direct reciprocity) have varied between authors and sub-disciplines (Carter 2014). Although the original concept was quite broad (Trivers 1971), more narrow subsequent definitions of reciprocity restricted its general importance to humans (Carter 2014). For example, the term has been used to describe: a broad category of enforced mutual benefit (analogous to kin selection), a correlation between cooperation given and received across dyads or over time (analogous to kin-biased association), a conditional helping behavior that causes this correlation (analogous to nepotism), and a specific psychological mechanism that might cause this conditional behavior (analogous to phenotype matching)(Carter 2014). For our purposes here, we define *reciprocity* broadly as help given that is influenced by rates of help received (i.e. reciprocal help), where help can involve different behaviors integrated over short or long timespans. Reciprocity has been most evident in controlled experiments where helping rates are immediate responses to past help between partners without social bonds (Dolivo and Taborsky 2015; Rutte and Taborsky 2008; Schweinfurth et al. 2017a; Schweinfurth and Taborsky 2018a; Taborsky et al. 2016). However, if individuals form enduring social relationships, they should integrate social experiences over long timespans and reciprocal help is expected to take multiple forms, such as allogrooming, food sharing, coalitionary support (Carter 2014; Jaeggi et al. 2013; Schino and Aureli 2017; Seyfarth and Cheney 2012). In this case, observed helping rates are a measure of an underlying causal relationship rather than the immediate cause of observed reciprocal help. Because mere correlations do not demonstrate reciprocity, we use the term *symmetry* for the observed correlation between rates of help given and received. Symmetry in helping is the most common observational evidence for reciprocity. Similarly, we use the term *nepotism* for help that is influenced by kinship cues, and we use the term *kinship bias* for the observed correlation between help given and kinship. Kinship bias is the most common observational evidence for nepotism.

Semantics aside, the importance of reciprocity is also contentious because it is difficult to test especially in the presence of nepotism. Moreover, kinship bias is considered sufficient evidence for nepotism, while symmetry, even in the absence of nepotism, is not sufficient evidence of reciprocity (Carter 2014). A demonstration of reciprocity requires experimentally manipulating helping rates and then measuring a change in reciprocal helping (Dolivo and Taborsky 2015; Fruteau et al. 2009; Krama et al. 2012; Krams et al. 2013; Krams et al. 2008; Rutte and Taborsky 2008; Schweinfurth and Taborsky 2018a). In such tests, kinship can be excluded as a factor by testing only nonkin. However, reciprocity in nature could also play a role in cooperation among kin (Jaeggi and Gurven 2013; Schino and Aureli 2010c; Schweinfurth and Taborsky 2018b; Taborsky et al. 2016; Wilkinson 1984; Wilkinson 1988). An experimental test of both nepotism and reciprocity therefore requires simultaneously manipulating cues to both kinship and past experience of cooperation (Schweinfurth and Taborsky 2018b; Zöttl et al. 2013). Given the logistical difficulty of this test, the vast majority of evidence for both reciprocity and nepotism is correlational (e.g. Carter and Wilkinson 2013; Schino and Aureli 2010c). In addition to nepotism and reciprocity not being mutually exclusive or independent, an additional challenge arises in observational studies because they are also not equally detectable.

In observational studies, nepotism will lead to collinearity between reciprocal help and kinship as predictors of helping rates. Due to inherent differences in sampling effort, kinship estimates will generally be more precise than estimates of helping rates (see discussion for details). As a consequence, the more precise estimate of the correlation between kinship and helping will often be over-estimated relative to the less precise estimate of the correlation between help given and received, and nepotism should tend to mask evidence of reciprocity in correlational datasets. To assess this idea, we developed a resampling procedure for inferring power to detect both kinship bias and symmetry in mixed-kinship groups. To simulate perfect reciprocity, we simulated data of helping events resulting from pairs of individuals that based their decisions to help on an unobserved history of past reciprocal help that is perfectly symmetrical. We then systematically changed two variables: (1) the degree of nepotism and (2) the number of observed helping events. Finally, we used permutation and bootstrapping to assess how these two factors interactively influenced the probability of detecting evidence for reciprocity.

To demonstrate the application of our approach to empirical data, we applied the same permutation and bootstrapping procedures to three datasets where both kinship and reciprocal help are suspected to influence cooperative behavior: allogrooming in female mandrills (*Mandrillus sphinx*), allogrooming in female Japanese macaques (*Macaca fuscata*), and food sharing in female common vampire bats (*Desmodus rotundus*). Mandrills appear to form large groups structured by matriline (Abernethy et al. 2002; Bret et al. 2013), and show evidence for reciprocal allogrooming (Schino and Pellegrini 2009) and kin discrimination (Charpentier et al. 2007; Levréro et al. 2015). Japanese macaques are nepotistic, have a despotic social network with a steep dominance hierarchy based largely on maternal kinship, and direct allogrooming to dominant individuals and consistently preferred partners (Balasubramaniam et al. 2018). Regurgitated food sharing in vampire bats has been a classic example of the possible co-occurrence of reciprocity and nepotism (Wilkinson 1984; Wilkinson 1988). We focus on cooperative interactions among females because females in all three species form matrilines and interact with both kin and nonkin.

To test hypotheses and estimate Type I and Type II error rates with both simulated and real datasets, we combined subsampling with bootstrapping to generate resampled datasets and plot the changes in the slopes and detection rates for kinship bias and symmetry across varying levels of sampling effort. These plots show whether the relative roles of kinship and reciprocal help are either remaining ambiguous or becoming clearer with more data. To generate coefficients and p-values, we used a permutation test designed to deal with both non-independence in the data and collinearity between the predictors (Dekker et al. 2007). R scripts are available online (Carter et al. 2018) so that readers can adapt them to their own observations of helping and kinship among humans or nonhuman animals.

Methods

Inferring power and precision of symmetry and kinship bias

To test if and how nepotism prevents the detection of reciprocity, we inferred the power to detect both kinship bias and symmetry in simulated and real datasets of various sizes. To infer power, we estimated how estimates of kinship bias and symmetry vary with an increasing number of observations (N). Note that N is the number of observed helping events not individuals. We began with about 20 equally-spaced steps of N, starting at N = 20 observations and with each next step adding about 5% of the total sample of observed interactions. For example, for a dataset of 500 observations we would have 20 sample size steps of N = 20, 45, 70, 95, 120 …. 500. At each step, we randomly sampled N observations from the total dataset. We sampled with replacement (bootstrapping) to avoid confounding smaller variances at larger samples sizes with smaller variances in our samples. We bootstrapped the datasets 1000 times at each sample size. For instance, at the first step, we randomly sampled 20 observations with replacement 1000 times. To analyze the simulated data (described below), we created a different dataset of size N observations by sampling from the input probability distributions 1000 times (i.e. Monte Carlo simulation) rather than bootstrapping a single dataset 1000 times.

For each observed dataset, we extracted the observed coefficients of kinship and reciprocal help from a multiple regression quadratic assignment procedure permutation test with double semi-partialling (MRQAP-DSP, (Dekker et al. 2007)). We defined the response variable ‘help’ for individual A to B as the total of duration of help from A to B, divided by the total duration of help received by B for all times where A could have helped B. This measure controls for differences in sampling time, and current situational factors such as need (Farine 2015). ‘Reciprocal help’ for A to B is then defined as help from B to A. We applied a log transformation to the empirical allogrooming and food sharing durations because they were lognormal. We z-transformed all variables to obtain standardized beta coefficients, so that an observed coefficient of X for kinship indicates that a one standard deviation increase in kinship predicts an increase of X standard deviations in help.

To calculate p-values for the observed coefficients, we used a network-level permutations (Farine 2017) randomizing each input variable independently using the standard approach from the MRQAP-DSP function in the R package ‘asnipe’ (Farine 2013). We used this procedure to generate one null coefficient from a randomized network for each observed coefficient, resulting in 1000 observed and 1000 paired null coefficient values for the two predictors, kinship and for reciprocal help, at each sample size step. At each sample size, we then calculated (1) the mean and 95% confidence interval (CI) for the observed coefficients, which are the observed symmetry and kinship bias estimates, (2) the mean and 95% CI for the null coefficients, which are the symmetry and kinship bias estimates expected under the null hypothesis, (3) the proportion of samples where the observed coefficient was greater than the paired null coefficient, which indicates if the effect is real using all the samples, and (4) the proportion of observed coefficients that were greater than 95% of the expected null coefficients, which indicates the power to detect an effect with one sample of a given size.

Simulating data resulting from 100% reciprocity and 0-100% simultaneous nepotism

We simulated 500 observations of help among 20 individuals. To simulate perfect reciprocity, we generated a network of reciprocal helping history, and set this to be perfectly symmetric across dyads, such that the helping history from A to B was always equal to helping history of B to A. We then created a dataset of 500 observed helping events by randomly sampling one individual as an actor and selecting a remaining individual as a recipient with a probability that was proportional to the helping history. The duration of help was equal to this symmetrical helping history. All observed helping was therefore determined by the symmetrical helping history, which is not directly observable but that probabilistically informs the observed data.

To simulate nepotism as an additional behavior, we made the history of past reciprocal help correlate with kinship to varying degrees. To do this, we constructed helping history network as the combination of two independent matrices of random values from 0 to 0.5 representing the kinship and non-kinship components of the helping history for each dyad. Nepotism determines the degree to which kinship predicts the helping history, so we calculated ‘helping history (h), as a combination of the kinship (r) and non-kinship components (c), weighted by a ‘nepotism coefficient’ (n), which ranges from 0 to 1:

$$h = nr + (1-n)c$$

The nepotism coefficient represents the degree to which past unobserved helping history correlated with kinship. We created populations where nepotism equaled either 0, 0.25, 0.5, 0.75, or 1. Increasing nepotism means a higher correlation between unobserved helping history and kinship (Table S1), so the observed correlations between observed help and kinship (i.e. kinship bias) will also increase. Finally, we added a step to ensure that all individuals were observed helping at least one other individual.

In sum, these simulations generate an observed set of helping events where individuals base their helping decisions entirely on an unobserved foundation of past reciprocal help. However, this reciprocity co-existed across a spectrum of nepotism from 0% nepotism, where helping rates are symmetrical and kinship played no role, to 100% nepotism, where helping was symmetrical but only occurred among kin, and the relative causal roles of reciprocal help and kinship are therefore unknown. Of course, there are many possible causes of symmetrical helping besides helping others based on reciprocal help. However, the point of this simulation is simply to ask: If perfect reciprocity did exist among individuals that were also somewhat nepotistic, how likely are we to overestimate the evidence for kinship bias relative to reciprocal helping?

Real datasets

We also applied this resampling procedure as a power analysis for three real datasets. The first two studies were conducted on mandrills and Japanese macaques housed at the Rome Zoo (Bioparco) in Italy. An observer recorded the duration of all female-female grooming episodes involving a focal subject, either as actor or recipient. All female subjects were available as potential grooming partners throughout the study. Kinship was based on maternal pedigrees. The first dataset contained 1703 observations of mandrill allogrooming collected between July 2014 and June 2015 from 10 sexually mature female mandrills in a group that also included two mature males. A past study of six female mandrills from the same captive population found that allogrooming A to B predicted allogrooming B to A, when controlling for kinship (Schino and Aureli 2010c), or when controlling for kinship and rank and excluding recent reciprocal grooming (Schino and Pellegrini 2009).

The second dataset contained 737 observations of macaque allogrooming collected between April and November 1996 from 22 sexually mature female Japanese macaques in a group of 71 that also included mature males and juveniles. Similar to the mandrills, analyses of allogrooming in the same captive population of Japanese macaques found symmetry in female allogrooming, and also found that allogrooming predicted support in social conflicts when controlling for kinship, rank, or time spent in proximity (Schino et al. 2007), and allogrooming was better predicted by kinship than by grooming received (Schino and Aureli 2010c).

The third dataset included 408 regurgitated food-sharing donations among 15 female common vampire bats from previous studies where food sharing was induced by fasting a subject (for details, see Carter and Wilkinson 2013; Carter and Wilkinson 2015). Each donation size was estimated by the total seconds that the unfed subject licked the mouth of a fed groupmate.

Kinship was estimated using a maternal pedigree and maximum likelihood estimates applied to genotypes of 19 polymorphic microsatellite markers (for details see Carter and Wilkinson 2015). Past analyses of these same data found that food sharing was better predicted by reciprocal sharing than by kinship, when controlling for grooming and donor sex (Carter and Wilkinson 2013a; Carter and Wilkinson 2013b), and this conclusion was supported by later experiments showing that the bats were attracted to the calls of nonkin donors more than non-donor kin (Carter and Wilkinson 2016) and that females that previously fed more nonkin were less affected by the removal of a donor from their food-sharing network (Carter et al. 2017).

Code availability

Data and R code, including functions to apply the same procedure to other datasets, is available online at the *figshare* data repository (Carter et al. 2018a; Carter et al. 2018b).

Results

Simulated data

Nepotism reduced the ability to detect the perfect symmetry in helping that exists independent of kinship. The simulations were constructed such that reciprocal help played a more important role than kinship in determining helping rates at every level of nepotism up to 100% nepotism, where reciprocal help and kinship were completely confounded. Yet as nepotism increased above zero, permutation tests were more likely to incorrectly infer that nepotism was more important than reciprocity (Figure 1). The reason for this can be seen in the plots of the size and precision of the observed and null coefficient estimates over increased sampling effort (Figure S1-S5). As nepotism increased, so too did the collinearity between kinship and the observed reciprocal help (Figure S6). Compared to helping symmetry, the kinship bias estimates were less variable. Starting at 50% nepotism, kinship biases were estimated to be larger than symmetry even though the generative probabilities and the actual durations of helping were always perfectly symmetrical (Figure 1). In these scenarios, where we know the real contribution of both kinship and reciprocal help as drivers of helping, we see that kinship bias was consistently overestimated relative to symmetry. At zero nepotism, estimates of symmetry in helping were still underpowered at 500 observations (Figure 1, see also Figure S7).

Real data

Results with empirical data are consistent with expectations from the analysis of the simulated data. For female mandrill allogrooming, symmetry was eventually detected to be significantly greater than kinship bias but this required more than 1250 observations, and the power for either effect was similar across sampling efforts (Figure 2). A sample of 1703 observations provided adequate power to detect both effects, but the relative effect size estimates of symmetry and kinship appear to still be diverging with more observations (Figure 2), suggesting that the relative contribution of nepotism may still be over-estimated despite this large sample size of helping events. For female macaque allogrooming, where nepotism is quite strong, a sample size of about 60 observations provided enough power to reliably detect a positive kinship bias, but the full set of 737 observations did not provide enough power to reliably detect positive symmetry (i.e. power < 80%, Figure 3). This highlights the combined impacts of greater nepotism and fewer helping events per dyad. For food sharing among female vampire bats, kinship bias is detected more reliably than symmetry at all sample sizes (Figure 3). The estimate of kinship bias is relatively stable at about 200 observations, but at 400 observations the estimate of symmetry is still increasing with additional sampling. There is enough power to detect positive kinship bias and symmetry, but power is lacking to reliably estimate the amount of symmetry (Figure 3). Together the plots from our simulations combined with those from the empirical datasets illustrate how and why detecting kinship bias and symmetry requires much less sampling effort than identifying their relative importance.

Discussion

Why nepotism masks evidence for reciprocity

If kinship and reciprocal help are simultaneous predictors of helping, then nepotism will often be easier to detect and overestimated relative to symmetry in helping rates. Looking at the simulation results of Figure 1, consider a scenario of 100 helping observations, where individuals are performing perfect reciprocity. At zero nepotism, symmetry will be detected (i.e. 'statistically significant') with a 50% probability, and nepotism will have a low rate of false positives. At 25% nepotism, however, a researcher is now equally likely to detect symmetry versus a kinship bias. At 50% nepotism, kinship bias further masks symmetry and the probability of correctly detecting symmetry drops to less than 25% and kinship bias is detected about 90% of the time. At 75% and 100% nepotism, kinship bias is always detected while symmetry is usually not detected. Even if both effects are detected, the relative effect sizes give the clear but incorrect appearance that helping decisions were influenced more by kinship than by reciprocal help. In the 100% nepotism scenario where the symmetrical helping history and kinship are exactly equal, kinship bias is always detected, but with the same sampling effort in the 0% nepotism scenario, where helping is symmetrical and independent of kinship, there is only 50% power to detect evidence of reciprocity. Why does this happen?

Estimating a correlation or slope coefficient requires having accurate data on the predictor as well as the response variables. Cases with two highly correlated predictors (collinearity) will lead to imprecise estimates of both coefficients, making each effect harder to detect. More critically, the most precisely estimated predictor, in this case kinship, will often appear to have a relatively larger coefficient and to be more important, regardless of its true causal role. The joint contributions of nepotism and reciprocity create collinearity between the predictors kinship and reciprocal help, which weakens the power to detect either effect, and it also leads to masking of the reciprocal help effect at lower sampling efforts, because the precision of kinship estimates is independent of the number of helping events observed.

Genetic relatedness is becoming increasingly easy to measure compared to the history of helping interactions. Pedigrees can be supplemented or replaced by genetic or genomic data, which is becoming cheaper and easier to collect (Städele and Vigilant 2016), and the precision of marker-based relatedness estimates are often estimated by resampling methods. In contrast, the history of cooperation between any two individuals is often unknown. The existence and strength of a cooperative relationship must be estimated by sampling many helping events given and received (e.g. allogrooming, food sharing, or coalitionary support) to calculate dyadic helping rates. When such cooperative interactions are rare or difficult to observe, these dyadic helping rates will be imprecise measures of the true cooperative relationship between the individuals. For example, we might have observed A helping B, but never B helping A, despite the fact that B has helped A in the past, or may do so in the future. Such imprecision in the estimate of helping symmetry will have twice the effect on the ability to find support for reciprocity because it will affect both the predictor (how much B helps A) and the response (how much A helps B), whereas it affects kinship estimates only once through the response variable (how much A helps B). The sample size of helping observations does not affect the estimates of kinship, but it does determine the accuracy of the dyadic rates of help given and received. Fewer helping observations will directly impact estimates of symmetry but not kinship.

To give another example, if we observe only one sharing event for each dyad those data could be sufficient to detect and estimate an existing kinship bias (albeit the estimate might be poor), whereas we cannot use the same data to detect or estimate an existing symmetry in helping rates. If a cooperative relationship exists, then kinship estimates will generally be more precise than estimates of helping rates. Because the more precise estimate will be over-estimated relative to the less precise estimate, nepotism will tend to mask evidence of reciprocity.

This asymmetry in precision is compounded in larger study populations. The number of possible dyadic helping rates is almost the square of the number of individuals, n*(n-1). Observing enough dyadic helping interactions to accurately estimate a cooperation network

(Farine and Strandburg-Peshkin 2015) is therefore a far greater challenge than estimating pairwise kinship among the same number of individuals. The former requires many unique behavioral samples per individual, whereas the latter requires applying the same panel of genetic markers to one genetic sample per individual. As population size increases, so too does the inequality between the effort required to estimate helping rates versus pairwise kinship across all dyads. As a result, the power to detect helping symmetry relative to kinship bias will almost always decrease with more individuals.

Even if kinship plays a less important role than reciprocal help, nepotism will mask some or all of the evidence of reciprocity, especially as the sample of helping events becomes smaller. This is important because observations of costly helping events, such as food sharing, are often quite rare. For example, more than 400 hours of focal sampling of wild vampire bats over 26 months of fieldwork yielded 110 observations of regurgitated food sharing and only 21 observations of non-maternal sharing between adults with known association and relatedness (Wilkinson 1984). Researchers studying meat sharing in chimpanzees observed an average only 1.3 sharing episodes between adult male chimpanzees per successful hunt (Mitani and Watts 2001). When captive capuchin monkeys were separated by mesh barrier and given opportunities to share food, observations of 9896 interactions over food yielded only 18 instances of apparent food sharing and 4 unambiguous cases of one monkey directly giving food to the other (de Waal 1997a). More frequent interactions such as grooming are therefore useful for providing a window into cooperative relationships that may drive more costly helping decisions, but effect sizes based on these most frequent interactions are still underpowered with surprisingly large numbers of helping events (Figure 1).

How to infer power to detect symmetry and kinship bias

Given the presence of collinearity if both nepotism and reciprocity exist, one cannot interpret the standardized coefficients in the standard way (as the expected increase in helping with an increase in kinship holding reciprocal help constant), nor can one be sure from the typical permutation test that the relative magnitude of the coefficients at a given sample size reflect their actual predictive importance, because the relative precision of the estimates will be unclear. One can, however, estimate the precision of helping symmetry through resampling. Researchers often use resampling to assess the reliability of genetic marker-based estimates of relatedness (Kalinowski et al. 2006; Queller and Goodnight 1989; Wang 2011). One can also use resampling to create similar confidence estimates for social relationships (Farine and Strandburg-Peshkin 2015; Sánchez-Tójar et al. 2018).

Our resampling procedure allows one to determine how symmetry and kinship bias respond to increased sampling. Subsampling and bootstrapping generate confidence intervals and trajectories for each effect size with increasing sampling. By inspecting the resulting plots, one can infer if more power is needed to test a hypothesis about symmetry or kinship bias. For instance, if the coefficients are on a crossing trajectory, then more data are required before drawing any conclusions. If the trajectories are diverging, then more reliable conclusions can be drawn about which predictor is more important, but not by how much. If the trajectories are stable, this suggests that the precision of the estimate would not be improved with more sampling.

Five other reasons why reciprocity is hard to detect

Symmetrical helping rates in a kin-structured populations can be hard to detect for purely statistical reasons, but several biological factors can also make correlational evidence for reciprocity difficult to detect. First, unlike kinship, the helping rates between any two individuals can change during the course of a study. This is generally true for any measure of a social relationship such as dominance or association rate. The concept of symmetry as a correlation in helping rates across individuals rather than as a correlation in rates over time assumes the existence of a stable social relationship (Seyfarth and Cheney 2012), where stable helping rates are based on an extended history of past interactions that are likely to have occurred before the study began (Carter 2014; Schino and Aureli 2009; Schino and Aureli 2010a; Schino and Aureli 2010b). However, even when animals form such stable long-term bonds, social decisions are clearly influenced by immediate costs and benefits (Newton-Fisher and Kaburu 2017). For instance, one study reported that an alpha male chimpanzee was attacked and killed by his most strongly affiliated and most frequent grooming partner, that many observers might label as a ‘friend’ (Kaburu et al. 2013). Social relationships can not only change suddenly, but often asynchronously across dyads, meaning that defining a sampling period for helping rates can prove difficult.

Second, stable relationships are also expected to have a longer timescale for reciprocation, which makes symmetry more difficult to detect in a short timespan (Carter and Wilkinson 2013; de Waal 1997; Gomes and Boesch 2009; Gomes and Boesch 2011; Gomes et al. 2008; Jaeggi et al. 2013; Sabbatini et al. 2012; Schino and Aureli 2017; Seyfarth and Cheney 2012). Strongly bonded individuals show dyadic helping rates that are more predicted by a foundation of past events and less predicted by recent events, compared to weakly bonded individuals (de Waal 1997; Jaeggi et al. 2013; Schino and Aureli 2009; Seyfarth and Cheney 1984; Seyfarth and Cheney 2012). If individuals invest in relationships for long-term benefits, then a strong social bond might only show symmetry over a year but not a month or week. Hence, greater social stability makes it harder to detect helping symmetry over short time periods, whereas lower social stability makes it harder to detect symmetry over long time periods. The relative importance of recent events versus longer-term past experience in driving decisions to cooperate is an open question for most behaviors and species.

Third, for any given relationship, the relative importance of reciprocity and nepotism can change over time. For example, the help invested by a mother in her two daughters when they are young might be 100% nepotistic and 0% reciprocal, with equal helping allocated to each daughter. However, when her daughters become adults, the mother’s investment might also be influenced by each daughter's reciprocal investment in her, and she may have a stronger relationship with one daughter over another.

Fourth, reciprocity can be more difficult to detect than nepotism because cooperative returns can take different forms (Borgeaud and Bshary 2015; de Waal and Berger 2000; de Waal 1997; Fruteau et al. 2009; Gomes and Boesch 2009; Gomes and Boesch 2011; Seyfarth and Cheney 1984). If each decision to help is based on a weighted sum of different forms of past help, then any one form of help might not show much symmetry. If reciprocity in a stable long-term relationship involves some form of 'emotional book-keeping' (Schino and Aureli 2009), then 'grooming on Tuesday can create an emotional bond that causes meat sharing on Saturday afternoon' (p. 167, Seyfarth and Cheney 2012). Asymmetries in any one service are expected if subordinates groom more dominant individuals in exchange for tolerance (Borgeaud and Bshary 2015; Port et al. 2009; Tiddi et al. 2011; Ventura et al. 2006), or if individuals adjust their grooming rates based on the ability of partners to provide food relative to others (Fruteau et al. 2009). Such asymmetries play a key role in biological market theory (Noë and Hammerstein 1994; Noë and Hammerstein 1995).

Finally, the degree of contingency in a reciprocal relationship may be more or less strict among kin. In other words, reciprocity and nepotism might interact. This occurs if the contingency between helping given and received differs between kin and nonkin. A negative interaction between contingency and kinship was observed in cooperatively breeding cichlids where dominants share their nests with subordinate helpers that must 'pay-to-stay', but subordinates nonkin help more than kin because dominants tolerate subordinate kin regardless of their degree of alloparental care (Zöttl et al. 2013). Such differences in reciprocity between kin and nonkin may also exist in primate long-term cooperative relationships, but they can only be detected through repeated manipulations of helping among both kin and nonkin.

Experimental evidence for simultaneous nepotism and reciprocity

To test the causal roles of reciprocity and nepotism, experiments must manipulate both the helping history and kinship cues that influence decisions to help. To our knowledge, this has only been accomplished once using an experimental paradigm where rats are trained to understand how to pull a bar to deliver a food reward to a partner rat. In a series of experiments, reciprocity was evident because decisions to pull for a partner were influenced by factors such as past food received or allogrooming received from the partner (Rutte and Taborsky 2008; Schweinfurth and Taborsky 2018a). To test for a simultaneous kinship effect, outbred wild-type male rats were separated from littermates, housed with non-kin, tested for an ability to recognize kin, and then tested in the same food-pulling task with partners that varied in both their past reciprocal help and kinship (Schweinfurth and Taborsky 2018b). The rats demonstrated kin discrimination by preferring to associate with unfamiliar kin over unfamiliar nonkin, but they did not show nepotism in the food pulling task; kinship did not increase food pulling nor did it change the symmetry of reciprocal pulling rates (Schweinfurth and Taborsky 2018b).

When is nepotism harder to detect than symmetry?

There are several conditions under which nepotism is unlikely to be detected. First, Type 1 error is greater when kinship estimates are imprecise or when there is insufficient variation in kinship among dyads (Csilléry et al. 2006). Although genetic and genomic data is becoming cheaper, easier, and more available (Städele and Vigilant 2016), kinship estimates based on genetic data can still be quite imprecise (Csilléry et al. 2006; Pemberton 2008; van Horn et al. 2008). Second, relatedness estimates become severely biased using allele frequencies calculated from only a few animals (Wang 2017). If genetic samples are used to score relatedness in a small subset of individuals, it is crucial to calculate the baseline population allele frequencies from a much larger sample. Third, in studies using pedigrees based on births, ‘kinship’ is actually maternal kinship. These estimates may be ecologically valid if the animals themselves cannot recognize paternal kin, but increasing evidence suggests that some primates for instance can recognize unfamiliar paternal kin (Charpentier et al. 2007; Levréro et al. 2015).

Association rates might mask evidence for nepotism when association is kin-biased and both association and kinship are tested as simultaneous predictors of helping rates. Although helping rates are often imprecise, association rates might often be more precise than kinship estimates when dyadic association rates are based on automated methods that can involve many thousands of observations (Alarcón-Nieto et al. 2018; Aplin et al. 2015). In such cases, the social networks could be described more accurately than the genetic relationships, and association will be easier to detect than kinship as a predictor of cooperation.

Quantifying if and how nepotism contributes to symmetrical helping in network data

If both reciprocity and nepotism exist, decomposing the inclusive fitness benefits of a cooperative trait into the relative fitness and indirect fitness components is probably implausible using empirical data. On the other hand, controlled experiments can identify proximate mechanisms by identifying the relative importance of different conflicting cues used to make helping decisions (Schweinfurth and Taborsky 2018b). Observational studies can also play an important role in looking at what proximate factors best predict cooperation in nature. Quite different proximate mechanisms can lead to the same pattern of symmetry and kinship bias in a given system, and investigating how cooperation network structures arise could help with distinguishing between possible mechanisms. For example, in one scenario kinship could determine proximity and proximity could then drive symmetrical grooming. Alternatively, individuals might associate in space independent of kinship and preferentially groom their kin. Recent developments in social network analysis (Farine and Whitehead 2015) and null models (Farine 2017) provide potentially useful tools for distinguishing mechanisms by constructing different mechanistic networks (Ilany and Akcay 2016; VanderWaal et al. 2014) or by constructing different null models that allow different aspects of associations to vary (Farine et al.

2015). Linking these models to tests of symmetry and kinship bias could yield greater insights into whether kinship or past experience shape patterns of helping directly or via proximity (Puga-Gonzalez et al. 2015).

Ultimately, patterns generated by reciprocal helping should have a temporal signature in that helping given should reflect some degree of helping received in the past. However, as we noted above, a major challenge is determining over what timeframe reciprocity takes place in a stable cooperative relationship. Few studies to date are likely to have a sufficiently complete dataset of helping behaviors to use temporal analyses. However, once such data are available, temporal social network analysis (Blonder et al. 2012; Farine 2018; Pinter-Wollman et al. 2013) could provide useful tools for investigating these topics.

Practical recommendations

When testing simultaneously for evidence of nepotism and reciprocity in cooperation networks, researchers should be aware of the necessity to estimate the reliability of estimates of both kinship and helping rates (the network edge weights) before comparing their relative importance. The most obvious way to improve inferences is to collect more interactions and to estimate kinship using more pedigree or genetic data. Adding more individuals, however, cannot compensate for a lack of repeated measures of the same individuals, which is what determines the precision of network edge weights (Farine & Strandburg-Peshkin 2015). One way to increase such repeated measures is to induce acts of cooperation. For example allogrooming can be induced by applying substances to the fur (Hemelrijk 1994; Schweinfurth et al. 2017b), cooperative mobbing can be induced with fake predators (Krama et al. 2012; Krams et al. 2010; Krams et al. 2013; Krams et al. 2008), food sharing can be induced by fasting individuals (Carter and Wilkinson 2013; Wilkinson 1984) or by creating opportunities to provide (Silk et al. 2013).

We used network permutations which account for the network structure and hold the total help given and received by each individual constant. Alternate versions of our analyses suggested that the multiple regression quadratic assignment procedure we used here (Dekker et al. 2007) is better at reducing collinearity, and hence the masking of symmetry, compared to permutation tests applied to standard multiple linear regression coefficients. However, network permutations do not account for biased sampling, so the helping rates (network edge weights) must take into account the relative opportunity for individuals to help each other. We accomplished this by defining edge weights as the proportion of help received from individual X divided by the total help received from all other individuals that could have otherwise come from individual X because X was present at the time. Another possibility is to define edges as the help from X over the opportunity for X to help. If helping events are scored as yes/no events, then an even more rigorous approach is to use pre-network permutations (Farine 2017), where the helping acts in the dataset are permuted across individuals present at the time, rather than permuting the helping rates in the network. Pre-network permutations allow for precise control over the null hypothesis by swapping within time periods or locations, and also control for biased sampling; however, they are most appropriate when the helping events are binary (0/1) and hence interchangeable.

In conclusion, because nepotism masks evidence of reciprocity in cooperation networks, it is useful to assess the reliability of symmetry and kinship bias as a function of sampling effort. We provide R code (Carter et al. 2018) to produce plots that allow one to assess the relative power for detecting evidence of nepotism and reciprocity in simulated datasets or in a given dataset of helping observations in humans or nonhuman animals.

Acknowledgements

We thank Erol Akçay, Jorg Massen, and an anonymous reviewer for constructive feedback which improved the manuscript. We thank the Rome Zoo (Bioparco) for allowing us to study their Japanese macaque and mandrill colonies, and Francesca Lasio and Raffaella Ventura for help with the data collection. GGC was supported by fellowships from the Smithsonian Institute and the Humboldt Foundation. DRF was funded by the Max Planck Society.

References

Abernethy, K. A., L. J. White, and E. J. Wickings. 2002. Hordes of mandrills (*Mandrillus sphinx*): extreme group size and seasonal male presence. Journal of Zoology 258:131-137.

Alarcón-Nieto, G., J. M. Graving, J. A. Klarevas-Irby, A. A. Maldonado-Chaparro, I. Mueller, and D. R. Farine. 2018. An automated barcode tracking system for behavioural studies in birds. Methods in Ecology and Evolution 9:1536-1547.

Aplin, L. M., D. R. Farine, J. Morand-Ferron, A. Cockburn, A. Thornton, and B. C. Sheldon. 2015. Experimentally induced innovations lead to persistent culture via conformity in wild birds. Nature 518:538-541.

Axelrod, R., and W. D. Hamilton. 1981. The evolution of cooperation. Science 211:1390-1396.

Balasubramaniam, K. N., B. A. Beisner, C. M. Berman, A. De Marco, J. Duboscq, S. Koirala, B. Majolo et al. 2018. The influence of phylogeny, social style, and sociodemographic factors on macaque social network structure. American Journal of Primatology 80.

Blonder, B., T. W. Wey, A. Dornhaus, R. James, and A. Sih. 2012. Temporal dynamics and network analysis. Methods in Ecology and Evolution 3:958-972.

Borgeaud, C., and R. Bshary. 2015. Wild vervet monkeys trade tolerance and specific coalitionary support for grooming in experimentally induced conflicts. Current Biology 25:3011-3016.

Bret, C., C. Sueur, B. Ngoubangoye, D. Verrier, J.-L. Deneubourg, and O. Petit. 2013. Social structure of a semi-free ranging group of mandrills (*Mandrillus sphinx*): a social network analysis. PLOS One 8:e83015.

Bshary, R., and R. Bergmüller. 2008. Distinguishing four fundamental approaches to the evolution of helping. Journal of Evolutionary Biology 21:405-420.

Carter, G., G. Schino, and D. R. Farine. 2018a. Data for "Nepotism masks evidence of reciprocity in cooperation networks", Figshare.

Carter, G. G. 2014. The reciprocity controversy. Animal Behavior and Cognition 1:368-386.

Carter, G. G., G. Schino, and D. R. Farine. 2018b. R code for “Nepotism masks evidence for reciprocity in cooperation networks”, Figshare.

Carter, G. G., and G. S. Wilkinson. 2013. Food sharing in vampire bats: reciprocal help predicts donations more than relatedness or harassment. Proceedings of the Royal Society B 280:20122573.

—. 2015. Social benefits of non-kin food sharing by female vampire bats. Proceedings of the Royal Society B 282:20152524-20152524.

Charpentier, M. J., P. Peignot, M. Hossaert-McKey, and E. J. Wickings. 2007. Kin discrimination in juvenile mandrills, *Mandrillus sphinx*. Animal Behaviour 73:37-45.

Clutton-Brock, T. 2009. Cooperation between non-kin in animal societies. Nature 462:51-57.

Cornwallis, C., S. West, and A. Griffin. 2009. Routes to indirect fitness in cooperatively breeding vertebrates: kin discrimination and limited dispersal. Journal of Evolutionary Biology 22:2445-2457.

Csilléry, K., T. Johnson, D. Beraldi, T. Clutton-Brock, D. Coltman, B. Hansson, G. Spong et al. 2006. Performance of marker-based relatedness estimators in natural populations of outbred vertebrates. Genetics 173:2091-2101.

de Waal, F. B., and M. L. Berger. 2000. Payment for labour in monkeys. Nature 404:563.

de Waal, F. B. M. 1997. The chimpanzee's service economy: food for grooming. Evolution and Human Behavior 18:375-386.

Dekker, D., D. Krackhardt, and T. A. Snijders. 2007. Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. Psychometrika 72:563-581.

Dolivo, V., and M. Taborsky. 2015. Norway rats reciprocate help according to the quality of help they received. Biology Letters 11:20140959.

Farine, D. R. 2013. Animal social network inference and permutations for ecologists in Rusing asnipe. Methods in Ecology and Evolution 4:1187-1194.

—. 2015. Proximity as a proxy for interactions: issues of scale in social network analysis. Animal Behaviour 104:e1-e5.

Farine, D. R. 2017. A guide to null models for animal social network analysis. Methods in Ecology and Evolution 8:1309-1320.

Farine, D. R. 2018. When to choose dynamic vs. static social network analysis. Journal of Animal Ecology 87:128-138.

Farine, D. R., J. A. Firth, L. M. Aplin, R. A. Crates, A. Culina, C. J. Garroway, C. A. Hinde et al. 2015. The role of social and ecological processes in structuring animal populations: a case study from automated tracking of wild birds. Royal Society Open Science 2:150057.

Farine, D. R., and A. Strandburg-Peshkin. 2015. Estimating uncertainty and reliability of social network data using Bayesian inference. Royal Society Open Science 2:150367.

Farine, D. R., and H. Whitehead. 2015. Constructing, conducting and interpreting animal social network analysis. Journal of Animal Ecology 84:1144-1163.

Fruteau, C., B. Voelkl, E. van Damme, and R. Noë. 2009. Supply and demand determine the market value of food providers in wild vervet monkeys. Proceedings of the National Academy of Sciences of the United States of America 106:12007-12012.

Gomes, C. M., and C. Boesch. 2009. Wild chimpanzees exchange meat for sex on a long-term basis. PLOS One 4:e5116.

Gomes, C. M., and C. Boesch. 2011. Reciprocity and trades in wild West African chimpanzees. Behavioral Ecology and Sociobiology 65:2183-2196.

Gomes, C. M., R. Mundry, and C. Boesch. 2008. Long-term reciprocation of grooming in wild West African chimpanzees. Proceedings of the Royal Society B 276: 699-706.

Griffin, A. S., and S. A. West. 2003. Kin discrimination and the benefit of helping in cooperatively breeding vertebrates. Science 302:634-636.

Hamilton, W. D. 1964. The genetical evolution of social behavior. Journal of Theoretical Biology 7:1-51.

Hemelrijk, C. K. 1994. Support for being groomed in long-tailed macaques, *Macaca fascicularis*. Animal Behaviour 48:479-481.

Ilany, A., and E. Akcay. 2016. Social inheritance can explain the structure of animal social networks. Nature communications 7:12084.

Jaeggi, A. V., E. De Groot, J. M. Stevens, and C. P. van Schaik. 2013. Mechanisms of reciprocity in primates: testing for short-term contingency of grooming and food sharing in bonobos and chimpanzees. Evolution and Human Behavior 34:69-77.

Kaburu, S. S., S. Inoue, and N. E. Newton-Fisher. 2013. Death of the alpha: within-community lethal violence among chimpanzees of the Mahale Mountains National Park. American Journal of Primatology 75:789-797.

Krama, T., J. Vrublevska, T. M. Freeberg, C. Kullberg, M. J. Rantala, and I. Krams. 2012. You mob my owl, I'll mob yours: birds play tit-for-tat game. Scientific Reports 2:800.

Krams, I., A. Berzins, T. Krama, D. Wheatcroft, K. Igaune, and M. J. Rantala. 2010. The increased risk of predation enhances cooperation. Proceedings of the Royal Society B 277:513-518.

Krams, I., H. Kokko, J. Vrublevska, M. Abolins-Abols, T. Krama, and M. J. Rantala. 2013. The excuse principle can maintain cooperation through forgivable defection in the Prisoner's Dilemma game. Proceedings of the Royal Society B 280:20131475.

Krams, I., T. Krama, K. Igaune, and R. Mand. 2008. Experimental evidence of reciprocal altruism in the pied flycatcher. Behavioral Ecology and Sociobiology 62:599-605.

Kümmerli, R., A. Gardner, S. A. West, and A. S. Griffin. 2009. Limited dispersal, budding dispersal, and cooperation: an experimental study. Evolution 63:939-949.

Lehmann, L., and L. Keller. 2006. The evolution of cooperation and altruism - a general framework and a classification of models. Journal of Evolutionary Biology 19:1365-1376.

Lehmann, L., L. Keller, S. A. West, and D. Roze. 2007. Group selection and kin selection: two concepts but one process. Proceedings of the National Academy of Sciences of the United States of America 104:6736-6739.

Levréro, F., G. Carrete-Vega, A. Herbert, I. Lawabi, A. Courtiol, E. Willaume, P. Kappeler et al. 2015. Social shaping of voices does not impair phenotype matching of kinship in mandrills. Nature Communications 6:7609.

Newton-Fisher, N. E., and S. S. Kaburu. 2017. Grooming decisions under structural despotism: the impact of social rank and bystanders among wild male chimpanzees. Animal Behaviour 128:153-164.

Noë, R. 2006. Cooperation experiments: coordination through communication versus acting apart together. Animal Behaviour 71:1-18.

Noë, R., and P. Hammerstein. 1994. Biological markets: supply and demand determine the effect of partner choice in cooperation, mutualism and mating. Behavioral Ecology and Sociobiology 35:1-11.

—. 1995. Biological markets. Trends in Ecology and Evolution 10:336-339.

Pemberton, J. M. 2008. Wild pedigrees: the way forward. Proceedings of the Royal Society B 275:613-621.

Pinter-Wollman, N., E. A. Hobson, J. E. Smith, A. J. Edelman, D. Shizuka, S. d. Silva, J. S. Waters et al. 2013. The dynamics of animal social networks: analytical, conceptual, and theoretical advances. Behavioral Ecology 25:242-255.

Port, M., D. Clough, and P. M. Kappeler. 2009. Market effects offset the reciprocation of grooming in free-ranging redfronted lemurs, *Eulemur fulvus rufus*. Animal Behaviour 77:29-36.

Puga-Gonzalez, I., A. Hoscheid, and C. K. Hemelrijk. 2015. Friendship, reciprocation, and interchange in an individual-based model. Behavioral Ecology and Sociobiology:1-12.

Rutte, C., and M. Taborsky. 2008. The influence of social experience on cooperative behaviour of rats (*Rattus norvegicus*): direct vs generalised reciprocity. Behavioral Ecology and Sociobiology 62:499-505.

Sabbatini, G., A. De Bortoli Vizioli, E. Visalberghi, and G. Schino. 2012. Food transfers in capuchin monkeys: an experiment on partner choice. Biology Letters 8:757-759.

Schino, G., and F. Aureli. 2009. Reciprocal altruism in primates: partner choice, cognition, and emotions. Advances in the Study of Behavior 39:45-69.

—. 2010a. A few misunderstandings about reciprocal altruism. Communicative and Integrative Biology 3:561-563.

Schino, G., and F. Aureli. 2010b. Primate reciprocity and its cognitive requirements. Evolutionary Anthropology: Issues, News, and Reviews 19:130-135.

—. 2010c. The relative roles of kinship and reciprocity in explaining primate altruism. Ecology Letters 13:45-50.

—. 2017. Reciprocity in group-living animals: Partner control versus partner choice. Biological Reviews 92:665-672.

Schino, G., and B. Pellegrini. 2009. Grooming in mandrills and the time frame of reciprocal partner choice. American Journal of Primatology 71:884-888.

Schweinfurth, M., J. Neuenschwander, L. Engqvist, K. Schneeberger, A. Rentsch, M. Gygax, and M. Taborsky. 2017a. Do female Norway rats form social bonds? Behavioral Ecology and Sociobiology 71:98.

Schweinfurth, M. K., B. Stieger, and M. Taborsky. 2017b. Experimental evidence for reciprocity in allogrooming among wild-type Norway rats. Scientific Reports 7:4010.

Schweinfurth, M. K., and M. Taborsky. 2018a. Reciprocal trading of different commodities in Norway rats. Current Biology 28:1-6.

Schweinfurth, M. K., and M. Taborsky. 2018b. Relatedness decreases and reciprocity increases cooperation in Norway rats. Proceedings of the Royal Society B 285.

Seyfarth, R. M., and D. L. Cheney. 1984. Grooming, alliances and reciprocal altruism in vervet monkeys. Nature 308:541-543.

—. 2012. The evolutionary origins of friendship. Annual Review of Psychology 63:153-177.

Silk, J. B., S. F. Brosnan, J. Henrich, S. P. Lambeth, and S. Shapiro. 2013. Chimpanzees share food for many reasons: The role of kinship, reciprocity, social bonds and harassment on food transfers. Animal Behaviour 85:941-947.

Städele, V., and L. Vigilant. 2016. Strategies for determining kinship in wild populations using genetic data. Ecology and Evolution 6:6107-6120.

Taborsky, M., J. G. Frommen, and C. Riehl. 2016. Correlated pay-offs are key to cooperation. Philosophical Transactions of the Royal Society B 371:20150084.

Tiddi, B., F. Aureli, E. P. di Sorrentino, C. H. Janson, and G. Schino. 2011. Grooming for tolerance? Two mechanisms of exchange in wild tufted capuchin monkeys. Behavioral Ecology 22:663-669.

Trivers, R. L. 1971. The evolution of reciprocal altruism. Quarterly Review of Biology 46:35-57.

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

van Horn, R. C., J. Altmann, and S. C. Alberts. 2008. Can't get there from here: inferring kinship from pairwise genetic relatedness. Animal Behaviour 75:1173-1180.

VanderWaal, K. L., E. R. Atwill, L. A. Isbell, and B. McCowan. 2014. Linking social and pathogen transmission networks using microbial genetics in giraffe (G iraffa camelopardalis). Journal of animal ecology 83:406-414.

Ventura, R., B. Majolo, N. F. Koyama, S. Hardie, and G. Schino. 2006. Reciprocation and interchange in wild Japanese macaques: grooming, cofeeding, and agonistic support. American Journal of Primatology 68:1138-1149.

Wang, J. 2017. Estimating pairwise relatedness in a small sample of individuals. Heredity 119:302.

West, S., A. Griffin, and A. Gardner. 2007a. Evolutionary Explanations for Cooperation. Current Biology 17:R661-R672.

West, S. A., A. S. Griffin, and A. Gardner. 2007b. Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. Journal of Evolutionary Biology 20:415-432.

Wilkinson, G. S. 1984. Reciprocal food sharing in the vampire bat. Nature 308:181-184.

Wilkinson, G. S. 1988. Reciprocal altruism in bats and other mammals. Ethology and Sociobiology 9:85-100.

Zöttl, M., D. Heg, N. Chervet, and M. Taborsky. 2013. Kinship reduces alloparental care in cooperative cichlids where helpers pay-to-stay. Nature Communications 4:1341.

Figure Legends

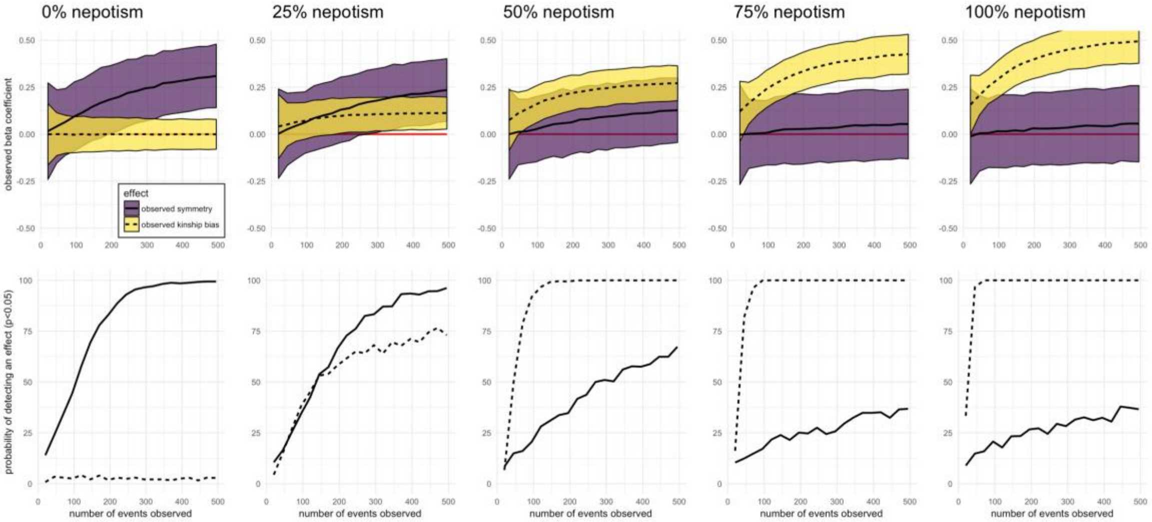


Figure 1. Simulated data: Kinship bias masks symmetry in cooperation networks resulting from perfect reciprocity. The probability and durations of help and reciprocal help are perfectly symmetrical within all dyads, and the degree of simultaneous nepotism increases from left to right. Top panels show the mean and 95% confidence interval of the standardized slope estimates for the effects of helping rate A to B (solid line, purple shading) and kinship between A and B (dotted line, yellow shading) as predictors of helping rate B to A. Bottom panels show the percentage of observed coefficients that were greater than 95% of the coefficients expected based on network permutations. Supplementary Figures S1-S5 in the appendix show plots for the null coefficients and for the probability of the observed coefficients being greater than expected coefficients.

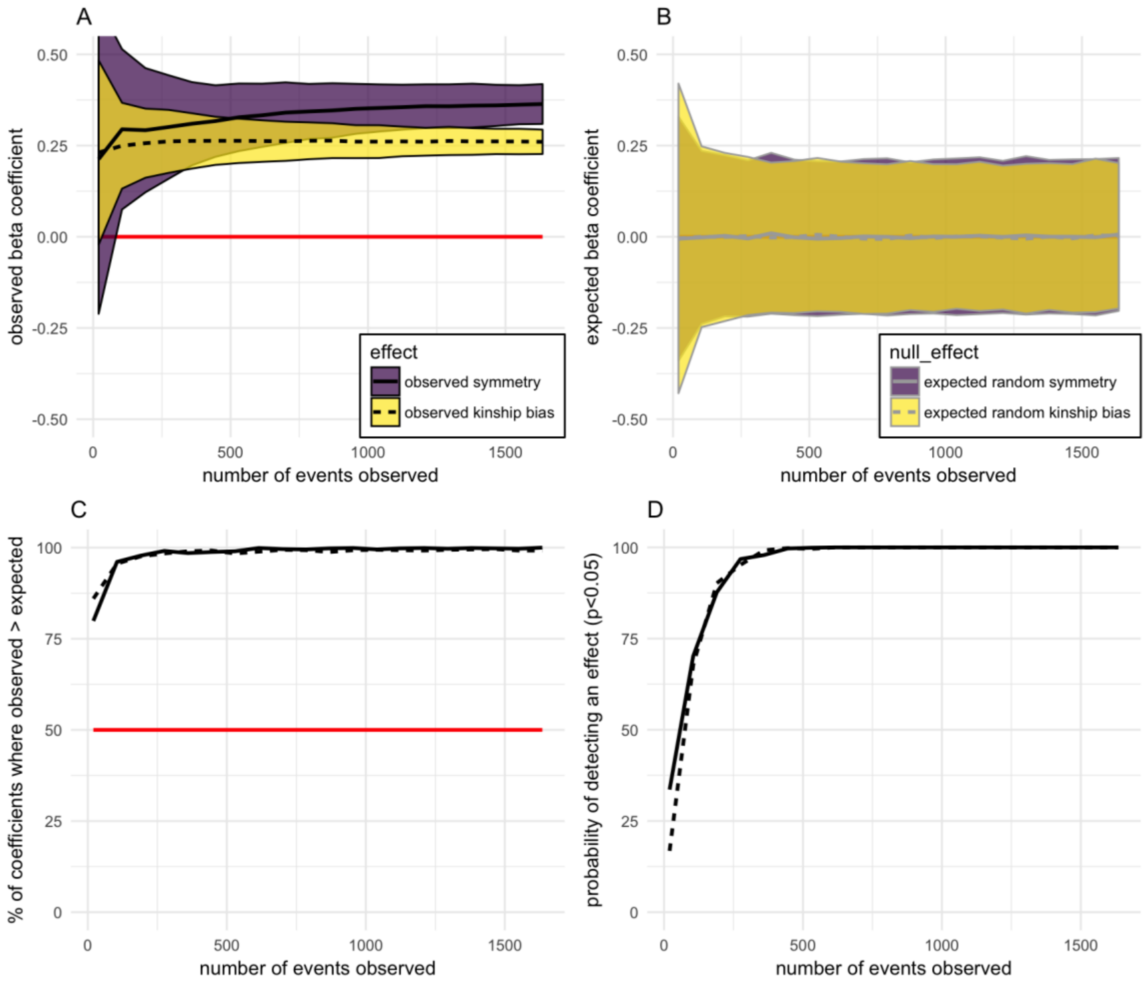
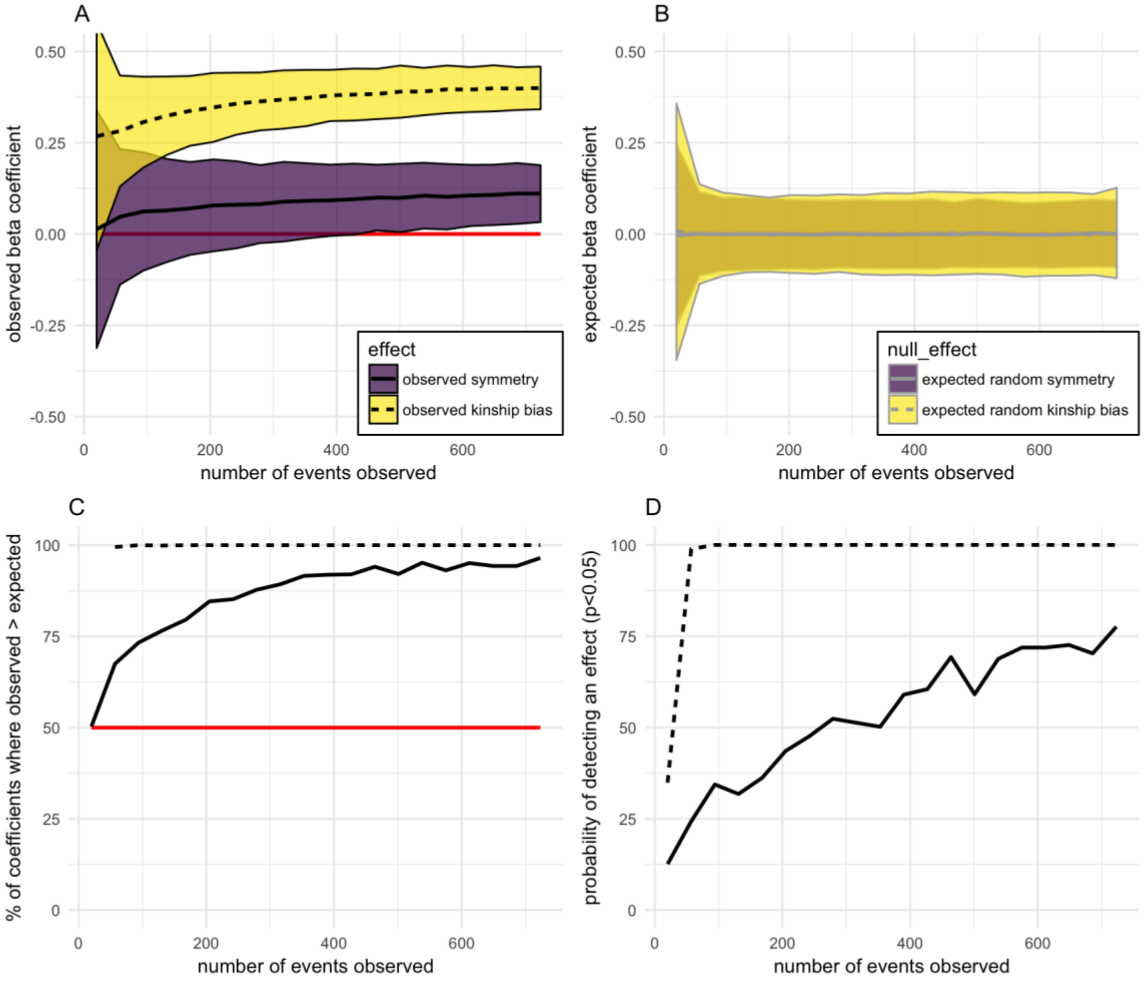


Figure 2. Female mandrills: observed and expected symmetry and kinship bias in allogrooming networks with increased sampling effort. Panel A shows the mean and 95% confidence interval of the standardized slope estimates for the effects of helping rate A to B (solid line, purple shading) and kinship between A and B (dotted line, yellow shading) as predictors of helping rate B to A. Panel B shows the same for the expected null coefficients generated by permutation. Panel C shows the percentage of observed coefficients that are greater than the paired null coefficient generated from the same subsample. If effects are real, then these values should be higher than 50% (red line). Panel D shows the percentage of observed coefficients that are greater than 95% of expected null coefficients for that sample size.



780

781 **Figure 3. Female Japanese macaques: observed and expected symmetry and kinship bias in**

782 **food-sharing networks with increased sampling effort.** See Figure 2 for explanation of plots.

783

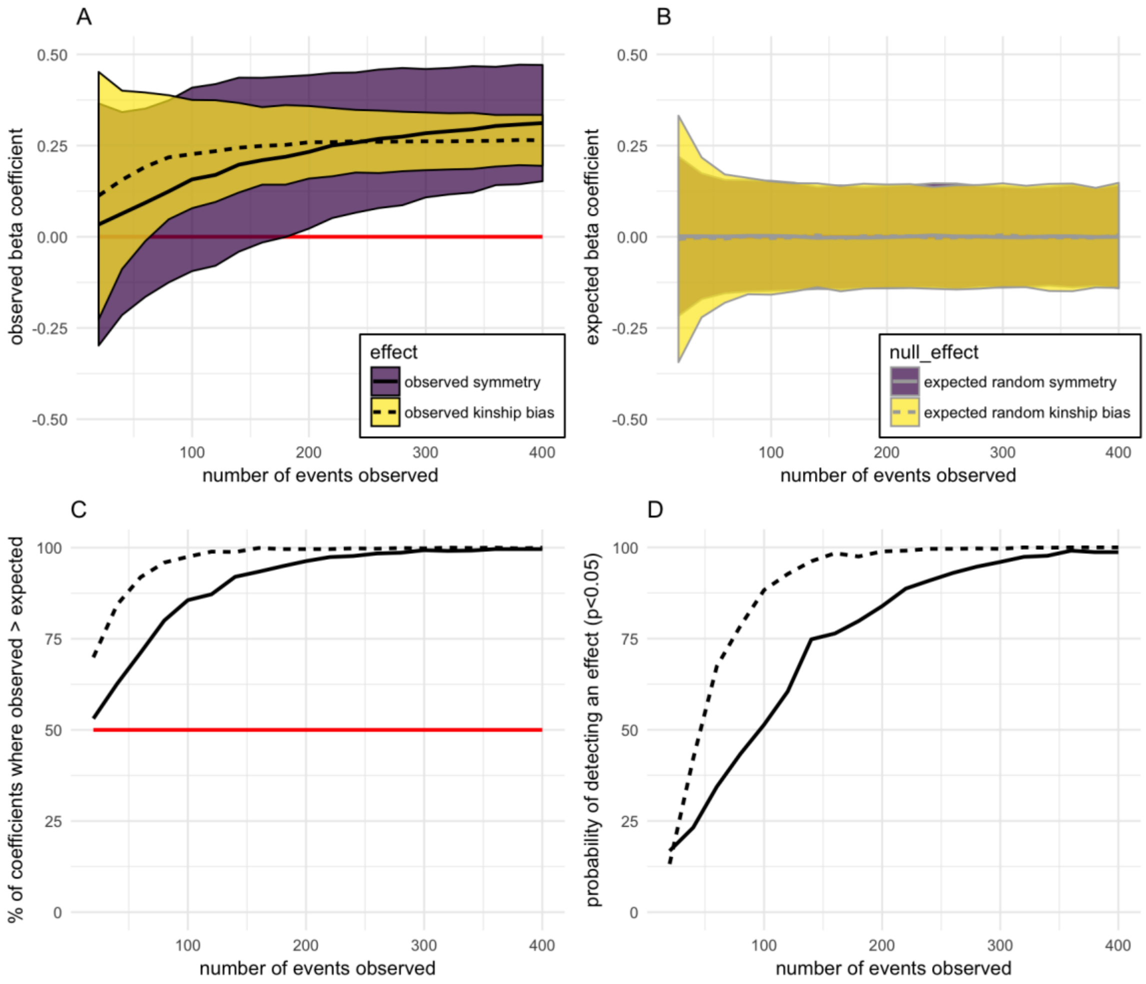


Figure 4. Female vampire bats: observed and expected symmetry and kinship bias in food-sharing networks with increased sampling effort. See Figure 2 for explanation of plots.