(De)composing sociality: disentangling individual-specific from dyad-specific propensities to interact

Christof Neumann[1,2] and Julia Fischer[1,2,3]

[1]Cognitive Ethology Laboratory, German Primate Center, 37077 Göttingen, Germany
[2]Leibniz Science Campus Primate Cognition, 37077 Göttingen, Germany
[3]Department of Primate Cognition, Georg-August-Universität Göttingen, 37077 Göttingen, Germany

2023-08-14

Abstract

1. In socially living animals, relationships between group members are typically highly differentiated. Some dyads maintain strong and long-lasting relationships, while others are only connected by weak ties. There is growing evidence that the number and strength of social bonds are related to reproductive success and survival. Yet, few of these analyses have considered that frequent or prolonged affiliative interactions between two individuals are driven by two different processes: namely, the overall gregariousness of the individuals involved and their dyadic affinity, i.e., the preference the members of the dyad have to interact specifically with one another. Crucially, these two axes of sociality cannot be observed directly, although distinguishing them is essential for many research questions, for example, when estimating kin bias or when studying the link between sociality and fitness.

2. We present a principled statistical framework to estimate the two underlying sociality axes using dyadic interaction data. We provide the R package `bamoso`, which builds on Stan code to implement models based on the proposed framework and allows visual and numerical evaluation of the estimated sociality axes.

3. We demonstrate the application and some of the critical advantages of our proposed modeling framework with simulated and empirical data: (1) the possibility of checking model fit against observed data, (2) the assessment of uncertainty in the estimated sociality parameters, and (3) the possibility to extend it to more complex models that use interaction data to estimate the relationship between individual-level sociality and individual-level outcomes in a unified model.

4. Our model will help to understand how and why individuals interact with each other and will help address questions about the relationship between variation in sociality and other features of interest, both within and across species.

Keywords:

social relationships, dyadic interactions, Bayesian modeling, social network, animal behavior

1 Introduction

The majority of socially living species maintain highly differentiated dyadic relationships. Analyses of these relationships allow us to infer key components of a society's social structure (Hinde, 1976; Kappeler & van Schaik, 2002), including its social complexity (Bergman & Beehner, 2015; Fischer et al., 2017; Kappeler, 2019). Typically, individuals maintain meaningful affiliative relationships with only a subset of other group members. Numerous studies have shown positive associations between fitness and various aspects of affiliative relationships, for example cumulative relationship strength (empirical examples: Cameron et al., 2009; Frère et al., 2010; Silk et al., 2010; reviews: Ostner & Schülke, 2018; Snyder-Mackler et al., 2020). However, other studies failed to find evidence for or even reported negative associations between sociality- and fitness-related measures (Dal Pesco et al., 2022; Duboscq et al., 2023; Menz et al., 2020; Thompson & Cords, 2018; Wey & Blumstein, 2012). Dyadic relationships are also crucial in models of pathogen and knowledge transmission (e.g., Drewe & Perkins, 2015; Nightingale et al., 2015).

One crucial step in any study about sociality is how to quantify it. Unfortunately, this is not a trivial endeavor. Since sociality cannot be observed or measured directly, we rely on recording behavior that is assumed to reflect sociality. More specifically, we rely on recording data that captures dyadic interactions and/or spatiotemporal associations between individuals. These data (typically frequencies or durations) are then typically processed to represent adjusted rates, proportions or behavioral indices (figure 2). For example, a number of highly influential studies used the dyadic sociality index (DSI, Silk et al., 2006, 2013) to quantify dyadic relationship strength. This index is based on a variable number of dyadic behaviors (typically frequencies of grooming interactions and proximity). Studies have shown that DSI varies with other dyadic features, for example relatedness and age similarity (Silk et al., 2006). More importantly, the DSI has also been used to express individual aspects of sociality. For example, it has been used to categorize dyadic relationships into strong versus weak, which in turn allowed to express the number of strong bonds an individual maintained (Silk et al., 2006). Using the sum of the top X (typically, three) relationships for each individual also has been done to arrive at an individual-level measure of sociality, which then has been used as a predictor for various components of individual fitness across a wide range of species (e.g, Silk et al., 2010; Young et al., 2014) (see also Ostner & Schülke, 2018). All these analyses rest on the implicit assumption that the derived values, regardless of whether they are dyadic or individual, represent the latent construct of sociality.

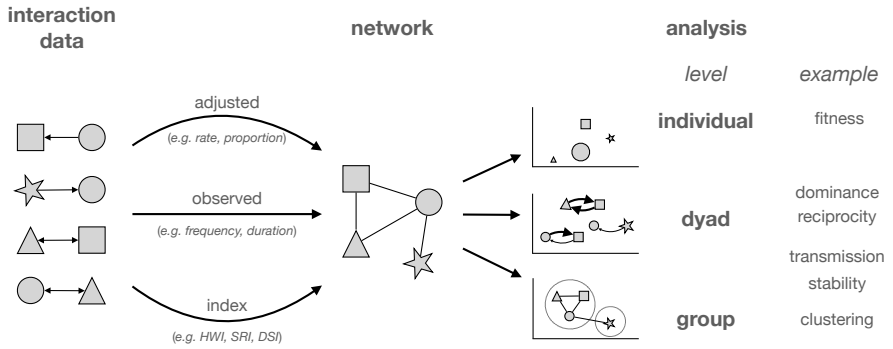

Figure 1: When data on dyadic interactions or associations are collected, they are often processed to represent social networks. These networks, or rather the dyadic connections, can either be set up directly from observed interactions (e.g., mating or dominance networks), via adjusted observations (e.g., rates of grooming, proportion of time co-feeding), or derived indices based on observations (association indices (HWI, SRI), composite dyadic sociality index (DSI)). Setting up the network is usually a means to an end in this workflow; the goal is to learn about the behavior of individuals, dyads, or groups.

As noted above, the behavior that we observe to quantify social relationship strength is dyadic in nature and it is essential to realize that dyadic interaction patterns are not driven by a single unobservable trait 'sociality' that can be assigned to each dyad. How often (or how long) two

individuals interact not only depends on the dyad-specific relationship between the two, but also on the individual propensities of the two individuals to interact with anyone (Godde et al., 2013; Pepper et al., 1999; Whitehead & James, 2015) (figure 2).

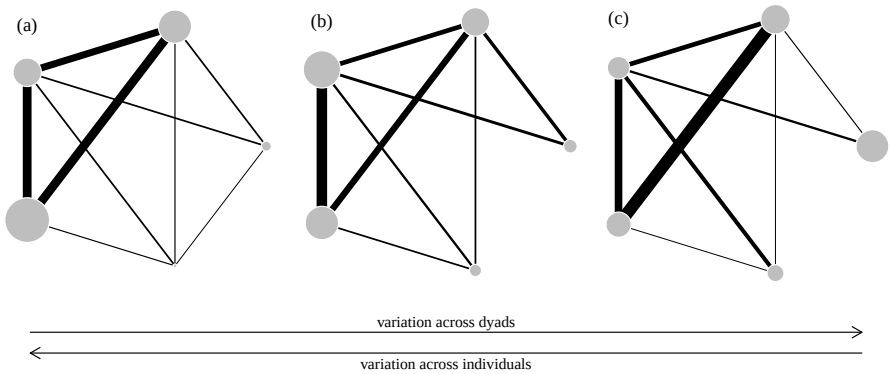

Figure 2: Interaction networks that look similar can arise from different processes. Across the three networks shown here, the edges (representing the *observed* interactions) appear similar despite being generated from combining two different processes that led to the depicted interactions: any particular dyad has roughly the same number of interactions in each network. At the same time, nodes, which represent the individuals (illustrated by the size of the circles), differ substantially between the three networks. In the network to the left, individuals vary considerably in their propensities to interact. In the network on the right, individuals differ relatively little. In contrast, dyadic propensities vary very little in the left network but differ substantially in the network on the right. The resulting interaction numbers (i.e., the edges in the network) look very similar. Interactions were generated from a Poisson process, i.e., they represent counts.

Two individuals may frequently interact because they are both intrinsically very social (have high propensities to interact with anyone) and not because they have a strong affinity with each other. Likewise, two individuals may interact because they have a strong affinity with each other, even though they might both have low propensities to interact in general. Hence, simply using observed interactions tells us nothing about the relative contributions of the individual- and dyad-level propensities because the observed interactions can arise from a mixture of both processes (figure 2). Expressing this thought slightly differently: we consider that the frequency or duration of dyadic interactions is due to at least two different components: the predispositions of individuals to *interact in general*, and the predispositions of dyads to *interact with one another*.

This distinction is important. Suppose we want to quantify 'pure' pairwise relationships. Such an analysis would be crucial to assess whether dyadic preferences exist in the first place. In that case, we are interested in excess interactions (or lack thereof), i.e., the duration/frequency of interactions that individual-level propensities alone cannot explain. As we show below in more detail, interaction patterns that look as if they are indicative of differentiated dyadic relationships might very well be the result of variation in individual-level propensities alone. Calculating dyadic indices and interpreting them as if they represented dyadic relationships would be questionable unless we can establish that a dyadic component exists in the first place.

For the sake of this study, we label the two sociality components individual gregariousness and pairwise affinity (Pepper et al., 1999). Because dyadic interactions are at the core of understanding social relationships, gregariousness and pairwise affinity reflect critical, and distinct, processes for how often (or long, or intensely) two individuals interact affiliatively (or associate). Unfortunately, the vast majority of analyses to date do not speak to the difference between these two processes generating affiliative interactions and likely conflate the two.

Previous attempts to disentangle gregariousness and pairwise affinity can be mainly found in the context of analyzing spatiotemporal associations (e.g., Godde et al., 2013; Pepper et al., 1999; Surbeck et al., 2017; Whitehead & James, 2015). The overall idea of these approaches is to arrive at dyadic affinity values while 'controlling' for gregariousness. There are at least two drawbacks

that these approaches have in common. First, individual gregariousness is not actually estimated in these procedures, but only approximated by summing pairwise affinity values for each individual or averaging the number of individuals the focal animal associates with. Second, the resulting affinity estimates are hard to interpret and there is no straightforward way to assess how well the estimated model parameters map onto observed association data, i.e., on the actual behavior of the studied animals. Comparing predictions from estimated model parameters to the observed data would be highly desirable. This would for example allow comparing competing models with respect to their predictive performance. If for example one model accounts for gregariousness and another one does not, it would be informative to see which of the two predicts interaction patterns closer to the actually observed patterns.

In contrast to spatiotemporal association data, we do not know of any analog approach to investigate variation in actual dyadic interactions.[1] Instead, dyadic interaction patterns, typically adjusted for observation effort or derived dyadic indices, are often taken to directly reflect dyadic relationships, thereby confounding individual and dyadic propensities in how these interaction patterns arose in the first place.

In order to tackle these issues, we outline a general, flexible, and extendable modeling framework that allows differentiating the contributions of individual and dyadic propensities to observed interaction and association patterns. We use *observed* dyadic interaction patterns to estimate the underlying but unobservable axes of individual-level sociality (gregariousness) and dyad-level sociality (affinity). We do this by fitting a model that reflects an assumed generative process based on contributions from individuals and dyads to explain variation in the observed interactions between dyads in a Bayesian framework.

This approach has at least four advantages. First, we can obtain estimates for dyadic *and* individual propensities from a single model fitted to a given data set (including the potential to directly assess uncertainty in these estimates, see appendix A2, see also Lusseau et al. (2008) and Farine & Strandburg-Peshkin (2015)). Second, we can make predictions from this model, which allows for objectively assessing model fit, i.e., we can directly compare predicted interaction patterns to observed interaction patterns (posterior predictive checks, Gabry et al. (2019)). Third, we can model multiple interaction types simultaneously, even with different error distributions (for example, the frequency of grooming bouts and durations of proximity). Fourth, the model can be extended such that the sociality features can be estimated and simultaneously be used as predictors in downstream models of other quantities of interest, carrying over uncertainty from the sociality estimates if so desired. It is also possible to extend the model such that we can use the estimated sociality features as a response variable (Rathke & Fischer, 2021).

2 Outline of the modelling approach

In line with the notion that dyadic interactions are the outcome of individual and dyadic propensities to engage in interactions, we can summarize our approach as follows:

$$\text{interactions}_{ij} = \text{baseline} + \frac{1}{2}(\text{individual sociality}_i + \text{individual sociality}_j) + \text{dyadic sociality}_{ij} \quad (1)$$

where interactions_{ij} reflects the observed interactions between individuals i and j. The terms $\text{individual sociality}_i$ and $\text{individual sociality}_j$ correspond to the individual gregariousness values of i and j, and $\text{dyadic sociality}_{ij}$ is the pairwise affinity for dyad ij. The *baseline* term reflects the propensities for two individuals with average gregariousness that have an average affinity, i.e., it is an intercept term. We multiply the sum of the two individual gregariousness values by 0.5 in order to bring the two underlying distributions of gregariousness and affinity to the same scale.

[1] In contrast to association patterns (as used in the mentioned studies, group-by-individual data), data on interactions often come in the form of counts or durations for a given observation effort.

More formally, we can express the model as follows, assuming for the sake of illustration that the observed behavior was a count variable (y_{ij}, for example, the number of how often two individuals groomed or the number of times two individuals were associated):

$$
\begin{aligned}
y_{ij} &\sim \text{Poisson}(\lambda_{ij}) \\
log(\lambda_{ij}) &= b + \frac{1}{2}(g_i + g_j) + r_{ij} \\
g &= N(0, \sigma_g^2) \\
r &= N(0, \sigma_r^2)
\end{aligned}
\tag{2}
$$

In this implementation, we estimate three parameters from the observed interactions. First, we estimate the variance σ_g^2 for individual propensities, i.e., how much do individuals vary in their propensities to interact (gregariousness). Second, we estimate the variance σ_r^2 due to each dyad's propensity to interact (affinity). The third component, b, is a baseline for the modeled behavior, which reflects the expected interactions for two individuals of average individual gregariousness with an average pairwise affinity.

After fitting the model to data, two main pieces of information can be obtained. First, we can extract the actual values for each individual's gregariousness and dyad's affinity. These individual and dyadic values reflect how much a given individual or dyad differs from an average individual or dyad. In statistical terms, we estimate varying intercepts for individuals and dyads. The obtained individual and dyadic values might then be used downstream to construct networks based on dyadic connections that are not confounded by individual propensities.

Second, we can assess the relative magnitude of individual and dyadic contributions to interactions. For example, if the individual variation is large relative to the variation in dyads, then one might conclude that individual propensities are more important than dyadic relationships in determining who interacts with whom. In the extreme case of the dyadic component being near 0, one might even conclude that interaction patterns are determined by individual propensities alone. If, on the other hand, the variation in the individual component is near 0, while the dyadic component is large, we could conclude that individuals are practically identical with respect to their general propensity to interact and that any variation in observed interactions is largely due to dyadic, i.e., relationship, effects.

Note that in this context, 'interactions' is used in a generic sense, i.e., it could mean *duration*, *proportion*, *count*, *frequency*, etc. The model can handle multiple kinds of interactions that were observed simultaneously and were recorded in different scales, for example, grooming (duration) and proximity (proportion of point samples).

A further advantage is that the model we present is fitted in a Bayesian framework, i.e., it produces posterior distributions for all the quantities of interest, which allow for assessment of certainty in individual quantities or differences between quantities (e.g., has individual A a stronger connection to B than to C?, see Kajokaite et al. (2022) and Neumann & Fischer (2023) for examples).

The model makes two key assumptions. First, both individual gregariousness and dyadic relationships are normally distributed. We use this as a convenient statistical shorthand to indicate that the majority of individual and dyadic propensities are 'average', while fewer individuals/dyads have smaller and stronger propensities. We also assume that there are no confounding factors. For example, an important driver for dyadic interaction patterns, in mammals at least, appears to be relatedness, which in this implementation is ignored but could be incorporated. Similarly, spatiotemporal associations may be driven by external factors, for example, aggregating around food resources (Whitehead & James, 2015).

At the same time, the model does not require thresholding in the sense that there is a minimum number of observations or observation effort required (Farine & Whitehead, 2015; Schülke et al., 2022), which also removes researcher degrees of freedom since there are no universally accepted rules for thresholding.

2.1 Model details and provided code

In essence, the model we outlined is a multi-membership model with an additional varying intercept for dyad ID and is also closely related to the social relations model (Koster et al., 2019, 2020). We implemented the model in Stan using the `cmdstanr` interface (Gabry & Češnovar, 2022) and provide helper functions in the accompanying R package `bamoso` to fit the model from R. The simple implementation with only one response variable/behavior can also be fitted directly with the `brms` package (Bürkner, 2017).

We used weakly informative priors for the individual and dyadic variability (50% of probability mass for the two standard deviations falls between approximately 0.3 and 1.5, which still allows for smaller and larger estimates). For the baseline values for each behavior, we also used priors that were weakly informative. Here, however, we used the data to inform how these priors were set exactly. This step is justified because we are estimating an intercept, and there is good reason to assume that the empirical median is a good approximation of the quantity we want to estimate (Bürkner, 2017).

Currently, the model we implemented accepts interaction data in four different data categories, which cover a wide range of how behavioral data are typically recorded:

1) counts (optionally with offset for observation effort)
2) discrete proportions (counts for a given number of discrete observation events, also like association indices (number of times seen together and number of times seen either dyad member))
3) continuous proportions (proportion of time)
4) continuous durations (optionally with offset for observation effort)

In its multivariate implementation, the model accepts any combination of behaviors (e.g., two count behaviors; three behaviors of which one is a count, one is a duration and one a discrete proportion). In addition, association data can also be supplied in group-by-individual format (Whitehead, 2008).

The accompanying R package `bamoso` provides helper functions to convert interaction data to a format that can be passed to Stan and some basic model evaluation functions. The supplement contains a manual that illustrates how to fit the model from R.

3 Illustrative examples with simulated data

In the following, we provide three examples using simulated data to illustrate the intuition behind our model. In the first two examples, our goal is simply to estimate the latent individual and dyadic sociality components, i.e., gregariousness and pairwise affinity, using one and two observed behaviors, respectively. The third example illustrates an application where the ultimate goal is to estimate the relationship between individual sociality and an individual-level outcome, which represents a template for doing an empirical study.

3.1 Example 1: one behavior observed

We created four scenarios that differed with respect to how individual propensities vary relative to dyadic propensities while keeping the group-level baseline constant.

- **scenario 1)** Both individual propensities and dyadic propensities varied little, which reflects a system in which interaction patterns vary stochastically around the group baseline. In other words, all individuals had practically the same propensities to interact (i.e., gregariousness), and all pairwise affinities were practically the same.

- **scenario 2)** Individual propensities varied substantially relative to dyadic propensities. This scenario reflects a system in which differentiated interaction patterns emerge, which are predominantly driven by individual differences and where dyad-specific propensities are practically absent and do not contribute to interaction patterns (although looking at interaction patterns only, this might not become apparent, see figure 2a).
- **scenario 3)** Individual propensities varied little relative to dyadic propensities. This situation reflects a system in which individuals do not differ from each other, while across dyads, pairwise affinity varies from having very weak to very strong affinities (see figure 2c).
- **scenario 4)** Both components varied substantially, which reflects a system in which both individual and dyadic propensities drive interaction patterns.

In each scenario, we simulated a group of 20 individuals (190 dyads) and set up a vector of individual propensities and a vector of dyadic propensities. Each of these vectors was drawn from a normal distribution with a mean of 0 and a standard deviation of 0.15 or 1.7, depending on whether the component varied a little or substantially in a given scenario. We then generated a count outcome from a Poisson distribution for each dyad using the equation above with a baseline of 2.5, which was the same across all scenarios. This means that if two individuals each had an individual propensity of 0 and their dyadic affinity was 0 as well, we expected them to interact about 12 times[2] ($exp(\frac{1}{2}(0+0)+0+2.5) = 12.2$). As in all our examples, these data are undirected.

Before applying our model, we first show the emerging interaction patterns under the different relative magnitudes of individual to dyadic contributions to interactions (figure 3). In the first scenario (with small individual and dyad variation), the emerging dyadic interaction numbers were all fairly similar and clustered around the intercept. The second and third scenarios led to fairly similar interaction patterns even though the underlying mechanisms were different. In scenario 2, interactions were driven predominantly by individual differences, while in scenario 3, interactions were driven by differences between dyads (see also figure 2). Scenario 4 also showed a differentiated interaction pattern although here both individual and dyadic components contributed to its emergence.

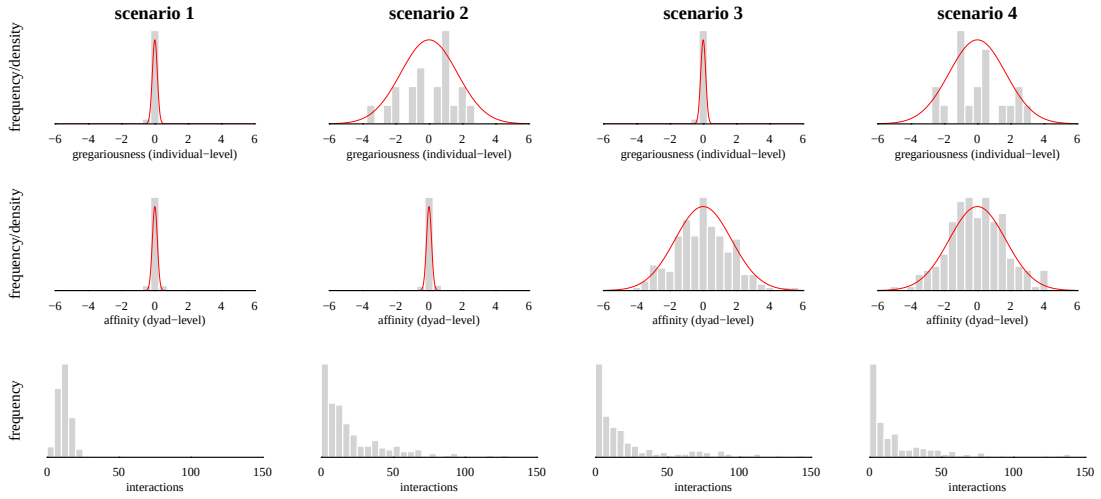

Figure 3: Emerging dyadic interaction patterns in groups of 20 individuals forming 190 dyads. Scenarios differed by combining low and high variability for individual and dyadic interaction propensities. See text for details.

Next, we turned the question around, i.e., if we only know the observed interaction patterns, can our model recover and separate the underlying individual and dyadic contributions? To this end, we fitted our model to each of the four data sets created by our four scenarios (figure 4).

[2]We need to exponentiate (i.e., use a link function) because we use a Poisson process with a log-link to generate the data.

First, we looked at our model's ability to predict interaction patterns that are similar to what we generated above, i.e., we performed posterior predictive checks (Gabry et al., 2019). In all four scenarios, the model was able to make predictions that matched the observed data very well (top row in figure 4). These plots show the densities of the observed interactions of 190 dyads and overlays 20 densities of interactions predicted by the model. As for recovering individual and dyadic values, we observed that in case the underlying true variation was large, the model recovered the true individual and dyadic values faithfully. When the underlying variation was small, the model recovered the fact that the variation was small (small estimated SDs).

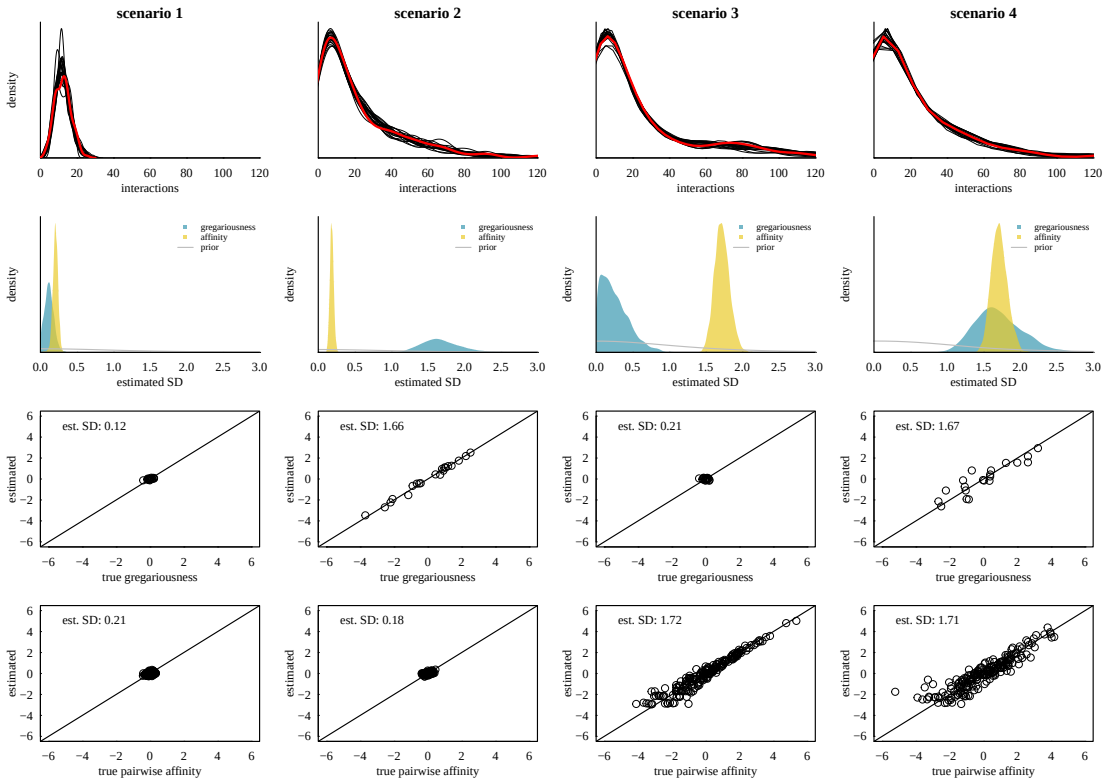

Figure 4: Model results of the four scenarios. In the top row are predicted patterns from model (red is observed as in the histograms above). Below are relationships between true (as per simulation) and estimated (posterior mean) individual and dyadic sociality values. The inset numbers are the posterior medians of the estimated sociality values (originally set to 0.15 and 1.7). The estimated intercepts for the four scenarios were 2.49, 2.52, 2.48 and 2.50 (originally set to 2.5).

3.2 Example 2: two behaviors observed

The second example illustrates a case when data from more than one behavior, for example, grooming and co-feeding, should be used to estimate the underlying individual and dyadic sociality components. A widely used index to quantify dyadic sociality, the composite dyadic sociality index (Silk et al., 2006, 2013), attempts precisely that but does not account for individual sociality. The core idea here is that we simultaneously fit our model to two behavioral response distributions and estimate a single (shared) individual and a single (shared) dyadic component while estimating two separate baselines for the two different behaviors. We created two scenarios where we first generated two behaviors resembling a count and a duration (figure 5). In the first scenario, dyads varied a lot while individuals varied little (similar to scenario 2 above). In the second scenario, individuals varied little, but dyads varied greatly (similar to scenario 3 above).

One additional central assumption that the model makes when applied in this way (i.e., with more than one modeled behavior) is that there is indeed a single underlying individual component and

a single underlying dyadic component. This assumption is made implicitly, for example, by the widely used composite sociality index (Silk et al., 2006, 2013). Whether this assumption holds can be checked by visually examining the model's predictions for the two behaviors. If the underlying processes differ substantially, the predictions will deviate from the observed data for either or both behaviors. In the case of our example, the predictions match the observed data well for both behaviors in both scenarios (figure 5, top two rows). Likewise, the estimated individual and dyadic values were recovered well for the cases where the variation in the given component was considerable.

3.3 Example 3: a complete workflow

Imagine we want to study the relationship between individual sociality and an outcome related to fitness, like the number of offspring produced or survival (e.g., Dal Pesco et al., 2022; McFarland et al., 2017; Schülke et al., 2010; Wey & Blumstein, 2012). Typically, such a study requires at least two distinct steps. First, the individual sociality variable is quantified from raw dyadic interaction data (in the examples above, these were degree from a sociality index network, cumulative strength of the three strongest edges from a sociality index network, and principal component scores from network metrics). The second step then takes these values and uses them as a predictor variable in the statistical model, in which the response variable is the (fitness-related) outcome of interest. Although a common practice, this workflow has at least two related downsides. First, it assumes that step 1 produces values that genuinely reflect the construct of interest. Second, it completely ignores any uncertainty in the values reflecting sociality.

Here we provide a simulated example that combines the estimation of gregariousness and feeds it directly into a single model of interest, carrying over uncertainty from estimating gregariousness into modeling the outcome of interest. In other words, we explicitly treat the quantification of sociality as a statistical model, as illustrated in the two examples above.

To achieve this, we set up a group of 25 individuals, in which each individual was assigned a value of sociality. These values represented each individual's true gregariousness and were drawn from a normal distribution with mean $= 0$ and $\sigma_g = 0.9$. Then we assigned each dyad an affinity value (normally distributed with mean $= 0$ and $\sigma_r = 2.3$, $n = 300$ dyads). With these values, we generated interactions for each dyad following equation 2, with a baseline rate of -1.5, i.e., a dyad of two average individuals with an average affinity would be expected to interact approximately 0.2 times ($exp(-1.5 + 0.5 * (0 + 0) + 0) = 0.22$). Next, we generated a count variable representing our outcome for each individual (e.g., the number of offspring produced). These numbers were drawn from a Poisson distribution with a mean of $exp(c_0 + c_1 \times g)$, where g is the individuals' gregariousness values generated above and c_0 and c_1 are intercept and slope set to $c_0 = -0.5$ and $c_1 = 1.2$, respectively. This resulted in a positive relationship between gregariousness and outcome, as illustrated in figure 6a. We are ultimately interested in this relationship, but crucially, we do not know the gregariousness predictor and need to estimate it from the observed interaction data.

To this aim, we set up the following model.[3] In brief, we estimated an individual sociality component (gregariousness) and a dyadic sociality component (affinity) from the interaction data and in parallel used the estimated gregariousness as a predictor for modeling the outcome. We assumed that we sampled the behavior with equal observation effort per individual and dyad.[4]

[3] Code is in appendix A1.

[4] Accounting for varying observation effort via offset terms can easily be integrated.

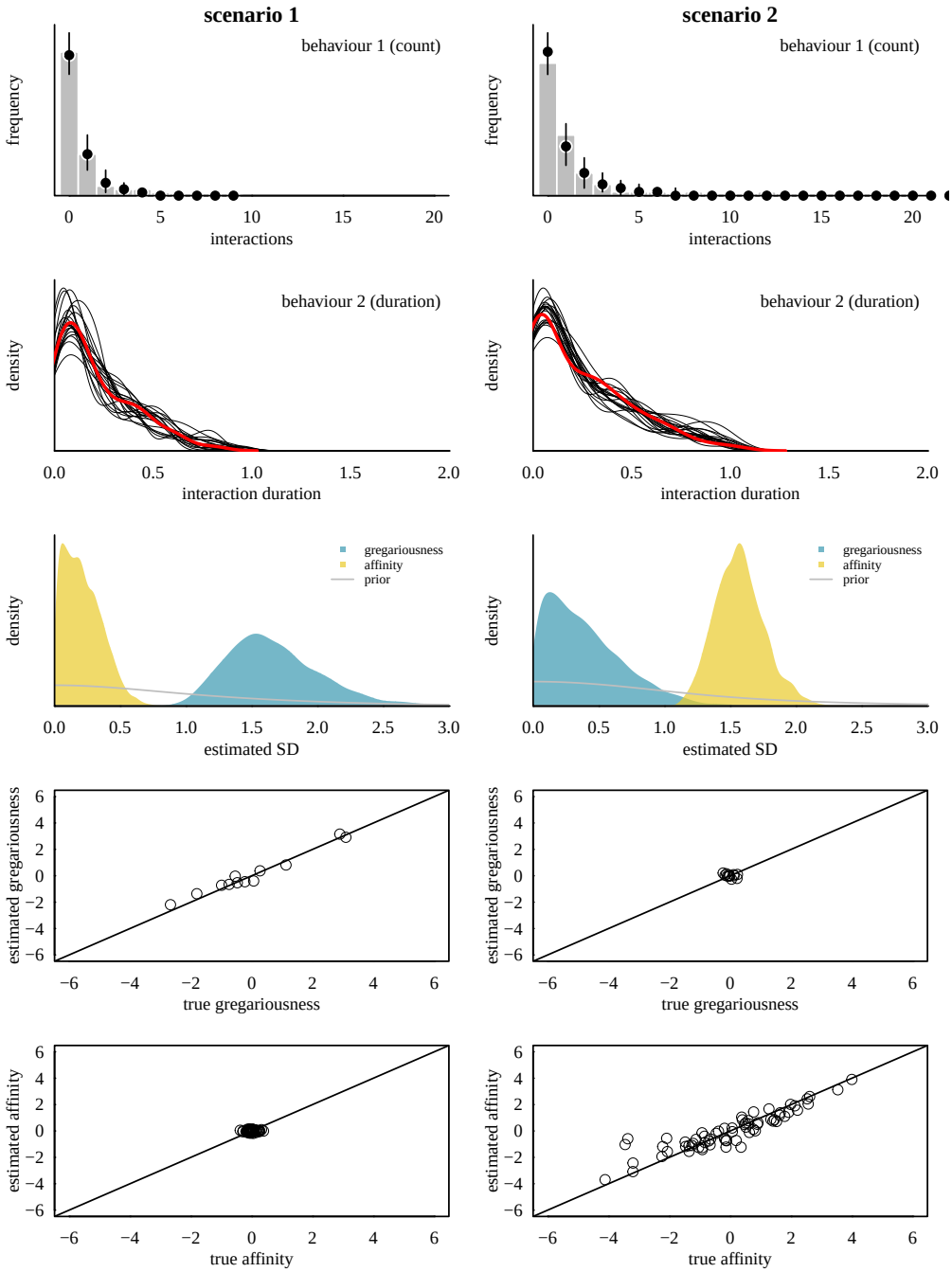

Figure 5: Model results of the two scenarios, each using two different interaction types (a frequency and a duration) to infer the underlying individual and dyadic sociality axes. The two top rows show posterior predictive checks for both behaviors. The third line summarises variation in gregariousness and affinity. The two bottom row show the relationship between true and estimated values for gregariousness and affinity.

$$
\begin{aligned}
\text{outcome}_i &\sim \text{Poisson}(\exp(c_0 + c_1 \times g_i)) \\
\text{interactions}_{ij} &\sim \text{Poisson}(\exp(\frac{1}{2}(g_i + g_j) + r_{ij} + b)) \\
g &\sim \text{Normal}(0, \sigma_g^2) \\
r &\sim \text{Normal}(0, \sigma_r^2) \\
\sigma_g^2 &\sim \text{Exponential}(1) \\
\sigma_r^2 &\sim \text{Exponential}(1) \\
c_0 &\sim \text{Normal}(0,\ 1) \\
c_1 &\sim \text{Normal}(0,\ 1) \\
b &\sim \text{Normal}(0,\ 1)
\end{aligned}
\tag{3}
$$

Fitting this model to our simulated interaction and outcome data produced estimates for all our quantities of interest. Foremost, the model recovered the true relationship between gregariousness and outcome fairly well (posterior medians; $c_0 = -0.49, c_1 = 1.43$, figure 6b) and was able to predict the observed outcomes faithfully overall (figure 6c). Similarly, also the parameters corresponding to modeling the interactions closely resembled their true values (posterior medians; $b = -1.60, \sigma_g = 0.89, \sigma_r = 2.40$). Thus, we were able to predict interaction patterns from the model that resemble the input interactions (figure 6d).

For illustration, we also ran two analyses representing how such a study question would have been addressed. First, we calculated individual degrees for the simulated network and then used degree as a predictor in a GLM with the simulated outcome as the response variable. Second, we constructed a network using DSI (Silk et al., 2013) from our simulated interaction network and then extracted the summed strength of its three strongest partners for each individual. We then used each metric as a predictor in a simple GLM for our simulated outcome. These models yielded very different results than our integrative model (figure 6e and f). Notably, the relationships in both cases were much closer to zero than our estimate, suggesting that these measures of individual sociality are not good predictors of the outcome in this simulated data set.[5]

This simulated example illustrates how a model could be designed to answer questions about the association between individual sociality and an outcome variable of interest. It also demonstrates some of the key advantages of our approach: allowing for model checking on the level of the observed interactions (and also at the level of the outcome of interest), and it illustrates how uncertainty in the estimation of the sociality component is carried over into the actual model of interest.

Nevertheless, the model we fitted here could be improved. Importantly, we did not run prior predictive checks (Gabry et al., 2019), which likely would have led to setting more reasonable priors and hence more solid inference given that our model, while overall recovering the true parameters reasonably well, also resulted in draws that made seemingly implausible predictions.[6]

4 Evaluation

We created artificial data sets to evaluate the model's performance across a range of biologically plausible scenarios. This analysis extends the results we presented above (figures 4 and 5). In particular, we varied individual gregariousness variation (SD ranging continuously between 0 and 2), dyadic affinity variation (SD ranging continuously between 0 and 2), and behavioral baseline (ranging continuously between -3 and 0). We also varied group size (number of individuals: 5 through 30), the number of behaviors (one or two), and the distribution type of the behavior (frequency, discrete proportion, continuous proportion, and duration). We first evaluated the

[5] And it raises the question as to whether these metrics are suitable as a proxy for gregariousness/social integration.

[6] In figure 6b, some of the predictions for the outcome for the individual with the highest gregariousness were larger than 50 (not visible in the figure), which seems not very likely given the range of the *observed* outcomes.

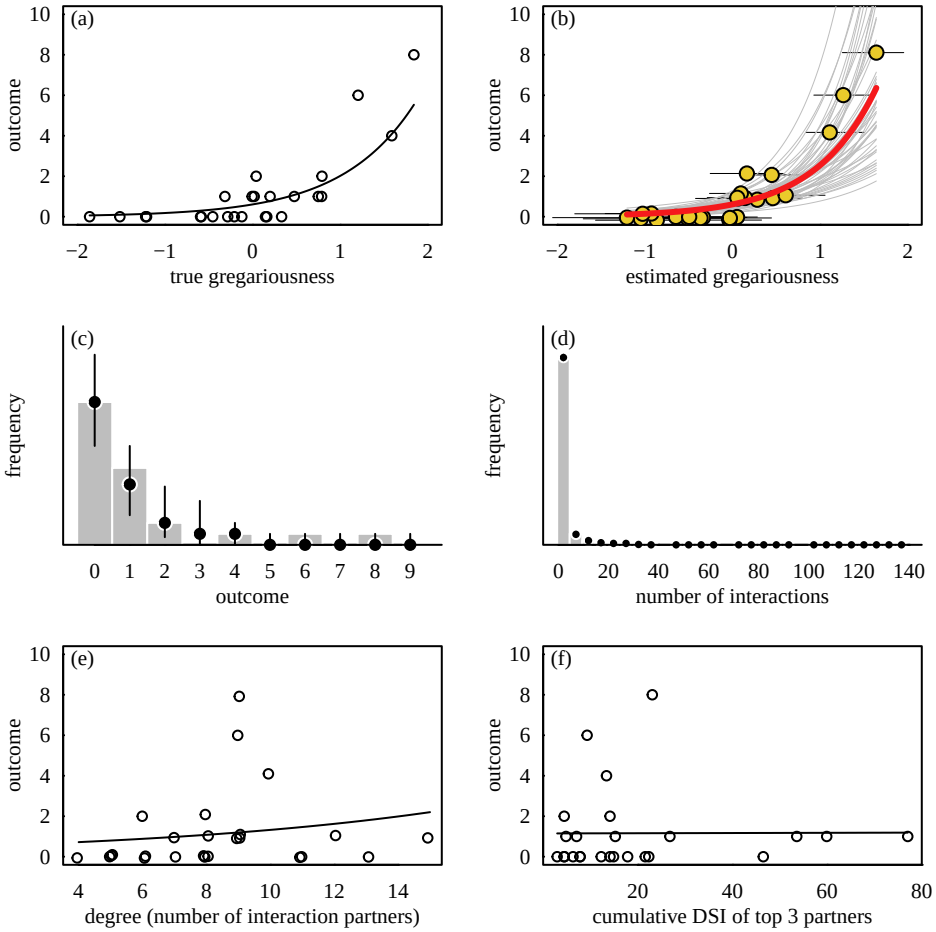

Figure 6: Estimating the relationship between gregariousness and an outcome using simulated data. Panel (a) shows the true underlying association between gregariousness and outcome, which in our case, is a relatively strong positive relationship. Panel (b) shows the results of our model, which *estimated* gregariousness from interaction data. The panel shows as circles the posterior medians of estimated gregariousness and the horizontal lines represent the 50% credible intervals for each estimate. In red is the estimated model of the outcome (using the median posterior intercept and slope). The thin grey lines correspond to model predictions using 50 random draws. Panels (c) and (d) depict posterior predictive checks for the outcome variable and the interaction data (using 50 random draws). The filled bars represent histograms of the observed counts, while the circles and vertical bars represent the median of the 50 draws and their 89% quantiles. Panels (e) and (f) illustrate the common practice of calculating individual-level scores as point estimates from the raw data and using them as predictors in a model.

relationship across data sets in how well the model recovered the variation in the underlying process (i.e., the relationship between estimated variation and ground truth). The four metrics we evaluated for each simulated data set were the correlation between true and estimated individual gregariousness values, the correlation between true and estimated dyadic affinity values, the difference between true and estimated variance in gregariousness, and the difference between true and estimated variance in dyadic affinity. If the model worked as promised, we expect the correlations to be near 1 and the differences near 0, without biases.

Our model recovered ground truth well for both components: when the simulated variation (i.e., ground truth) was small, the estimated variation was also small, and vice versa (figure 7). Notably, for the gregariousness component, our model sometimes underestimated the true variation (figure 8), which was primarily prevalent in smaller groups (blue/violet points in figure 7).

Finally, we also found that the correlations between the true and estimated individual and dyadic propensities were generally positive and approached 1 when the true values became larger (figure 9).

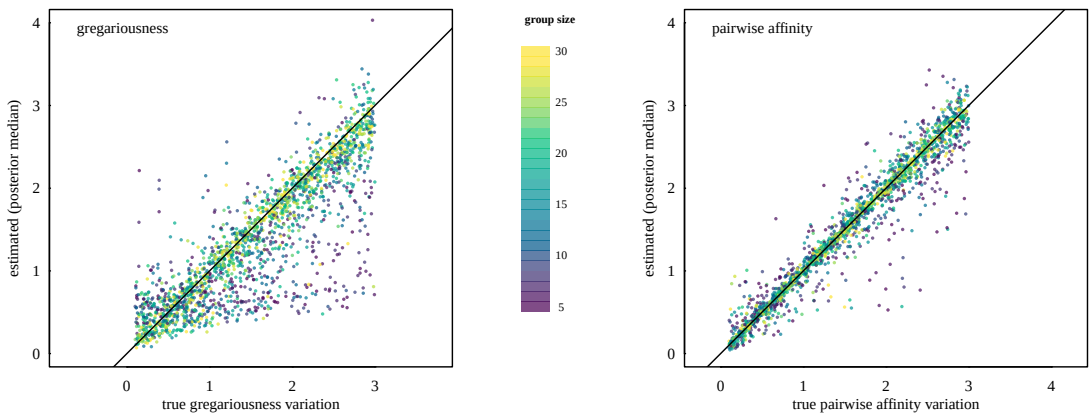

Figure 7: Recovering individual and dyadic sociality variation. Each circle represents one simulated interaction network.

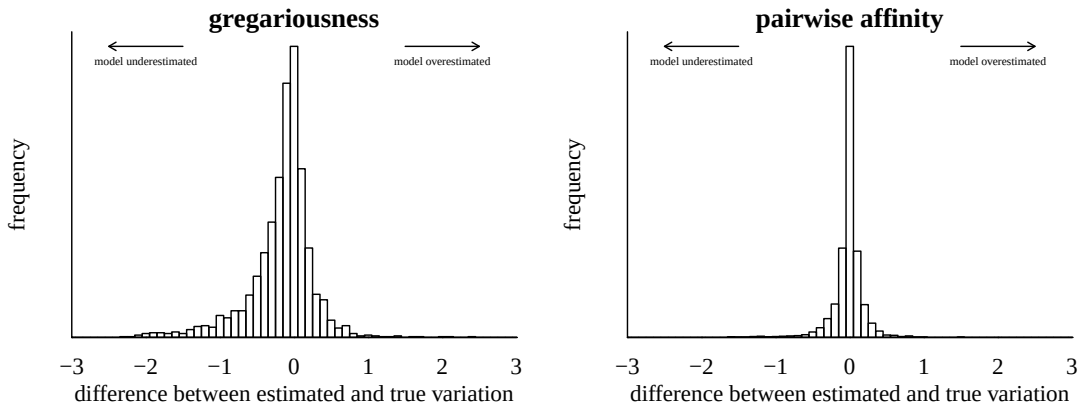

Figure 8: Differences between true and estimated gregariousness and affinity variation in simulated data.

5 Applied examples

In the following section, we illustrate the application of the framework with two examples. The first example descriptively compares the variation in the two sociality axes in groups of four

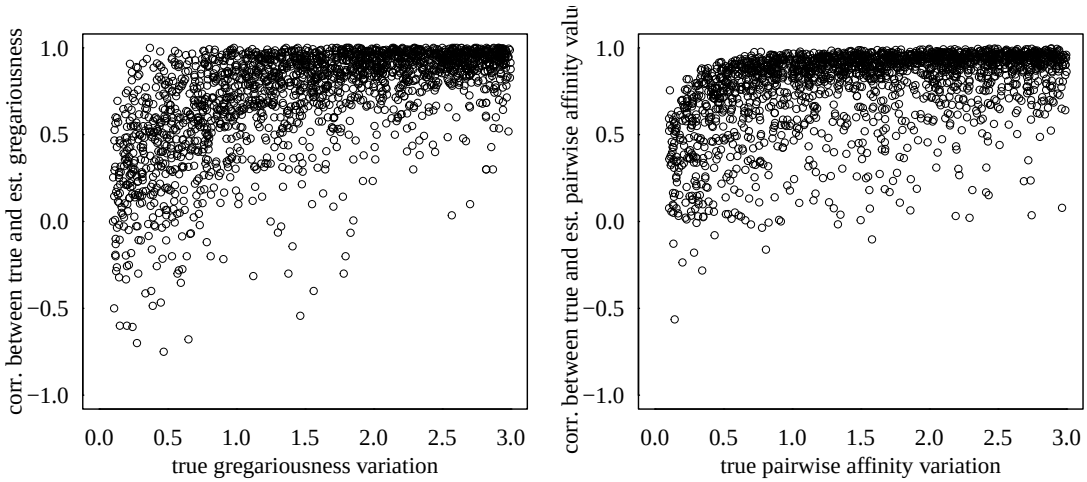

Figure 9: Correlations between true and estimated individual gregariousness values and between true and estimated pairwise affinity values within simulated networks. Each circle represents one simulated interaction network

primate species. The second example illustrates a 'complete' workflow where sociality is used as a predictor in a model of reproductive success in kangaroos.

5.1 Grooming in macaques

To illustrate a potential application, we analyzed grooming networks of four social groups belonging to four macaque species (*Macaca* sp.). The genus is widely used in comparative research in behavioral ecology (e.g., Adams et al., 2015; Amici et al., 2020; Balasubramaniam et al., 2012; Neumann & Fischer, 2023; Thierry et al., 2008). Here, we chose four species representing the diversity of macaque social styles (Thierry, 2007) to describe how much gregariousness and pairwise affinity vary across the different species and possibly in relation to their social style. For each species, we obtained data for one group. All but one of the groups (*sylvanus*) were observed in the wild, using focal animal sampling. The interaction data were undirected, and grooming was recorded as frequencies (duration for *nigra*). Group sizes ranged between 15 and 20 individuals, all of which were adult females.

We found substantial variation between the groups in the absolute and relative magnitudes of estimated gregariousness and affinity variation (figures 10 and 11). In two groups (*assamensis* and *nigra*), affinity variation was smaller than gregariousness variation (figure 10). In the *nigra* group, the estimate for the affinity component was so small that it could be argued that it appears to play only a minor role, if any, in explaining interaction patterns. In other words, the results suggest that for this group, grooming patterns were primarily driven by individual differences in grooming propensities. Across species, we found no obvious pattern. However, our results are tentatively more compatible with the notion that variation in both gregariousness and affinity is stronger in more despotic species (figure 11).

In the future, it would be desirable to incorporate a more diverse species data set, with possibly more than one group per species, to understand how much sociality components vary both *within* and *between* species.

5.2 Reproductive success in kangaroos

In this example, we reanalyzed a data set from Carter and colleagues (Carter et al., 2020; Menz et al., 2020) from a study that modeled the relationship between sociality and several measures of reproductive success in grey kangaroos (*Macropus giganteus*). The primary goal here is to illustrate

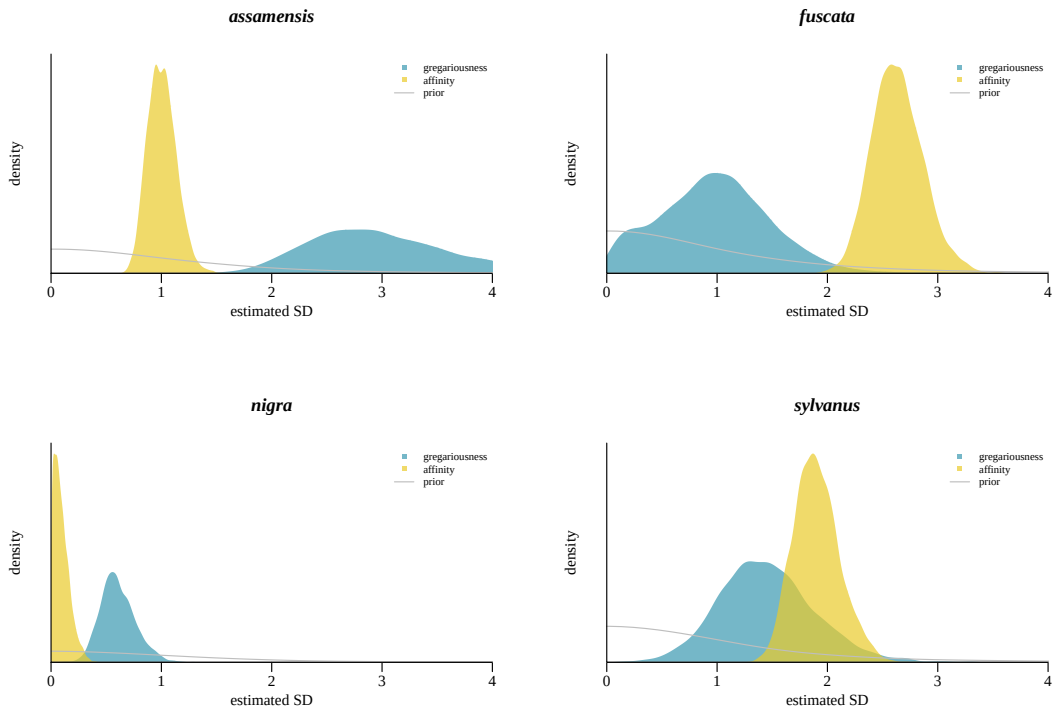

Figure 10: Variation in individual and dyadic grooming propensities in four groups of macaques. The four interaction networks represented undirected grooming frequencies in groups of 15 to 20 adult females.

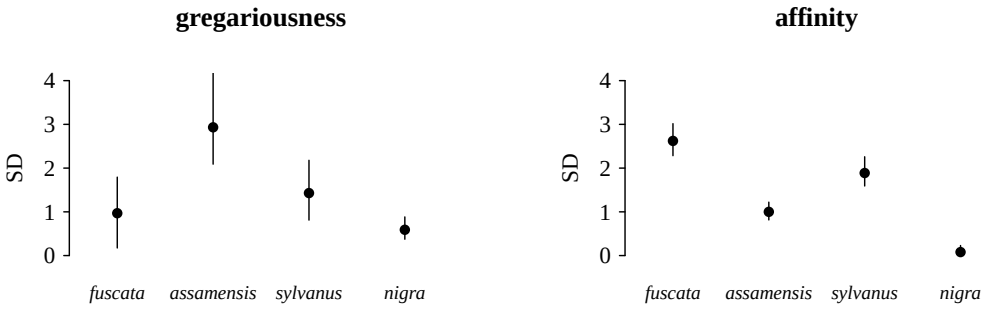

Figure 11: Variation in gregariousness and relationships in four groups from four species (tentatively ordered by their social style, from despotic (*fuscata*) to tolerant (*nigra*)). Shown are posterior medians with 89% credible intervals. This figure presents essentially the same information as the figure above, displayed in a different way.

how the framework can be extended using observed interaction data to quantify gregariousness and pairwise affinity. Specifically, we aimed for an analysis that uses raw behavioral data (data on spatio-temporal associations in this example) to estimate the actual model of interest, i.e., the relationship between individual sociality and reproductive success. In this way, we avoid doing this analysis in two separate steps, allowing us to carry over the uncertainty from estimating the two sociality components into our model of interest. This example can be considered an extension of the simulated example above (figure 6) with empirical, as opposed to simulated, data.

The authors of the original study constructed networks from association data using the Half-Weight Index (HWI: Cairns & Schwager, 1987). The authors extracted two metrics from these networks that represented individual sociality: strength and the local cluster coefficient. These two metrics then served as predictor variables in models of three measures of reproductive success.[7]

Implementing this approach in our framework came with one key challenge. The authors calculated derived metrics from HWI networks: strength and clustering. In order to follow this aspect of the original analysis, we had to code the algorithms to obtain strength and clustering in Stan (see appendix A3).

In addition, we needed to account for the fact that data came from individuals across multiple sampling periods (years) that only partly overlapped in composition, i.e., some individuals were represented in more than one period, and some only occurred in one period. Since this is not a direct replication attempt and only serves as an illustration for a real-world application, we made additional changes to the model design. First, we used a subset of 47 individuals (representing around one-third of the original sample size) to speed up model fitting. We also allowed the slopes for the two network measures (strength and clustering) to vary between the four years. Finally, we analyzed only one outcome variable (weaning success), instead of all three. Code for this analysis is in the supplementary files (repl_06_kangaroos.R).

The results show that our model successfully predicted the association patterns and the eventual outcome variable (did the female successfully wean an infant in a given year) (figure 12a and b). Furthermore, looking at the magnitude of variation in the two sociality components, we found that variation in both was substantially larger than 0 and that the dyadic affinity component was slightly larger than the individual sociality component (affinity: median = 1.38, 89%CI: 1.29−1.47; individual: median = 1.24, 89%CI: 1.05 - 1.46, figure 12c). Finally, the effect of individual network strength on weaning probability our model recovered was strongly negative and therefore agreed with the original findings of Menz et al. (2020): Females with higher strengths were less likely to wean successfully (figure 12d).

6 Discussion

In this study, we provide a framework to use dyadic interaction data to obtain estimates for individual-specific and dyad-specific propensities to interact. We argue that observed interactions are outcomes driven by the combination of these two processes. Estimating them both simultaneously has at least two distinct advantages. First, the parameters the model estimates have clear interpretations and allow checking whether our estimated sociality axes are able to predict our observed interaction patterns. This can serve as a sanity check on model validity, a step that should be incorporated in current approaches. Second, since we actually *estimate* the two sociality components, it is possible to grasp the uncertainty of the estimated sociality values for individuals and dyads (and the group baseline).

However, estimating the sociality components is rarely the ultimate goal in empirical questions (but see our comparative example above). Instead, we typically use sociality estimates as predictors in further downstream analyses to study variation in biologically relevant traits, for example, reproductive performance, survival or disease, or knowledge transmission. We, therefore, also illustrated how this could be achieved while carrying over the uncertainty in the sociality estimation to the actual model of interest (i.e., the model that links some particular sociality axis to an outcome of interest) instead of treating sociality values as point estimates.

[7] There were also a number of 'control' predictors.

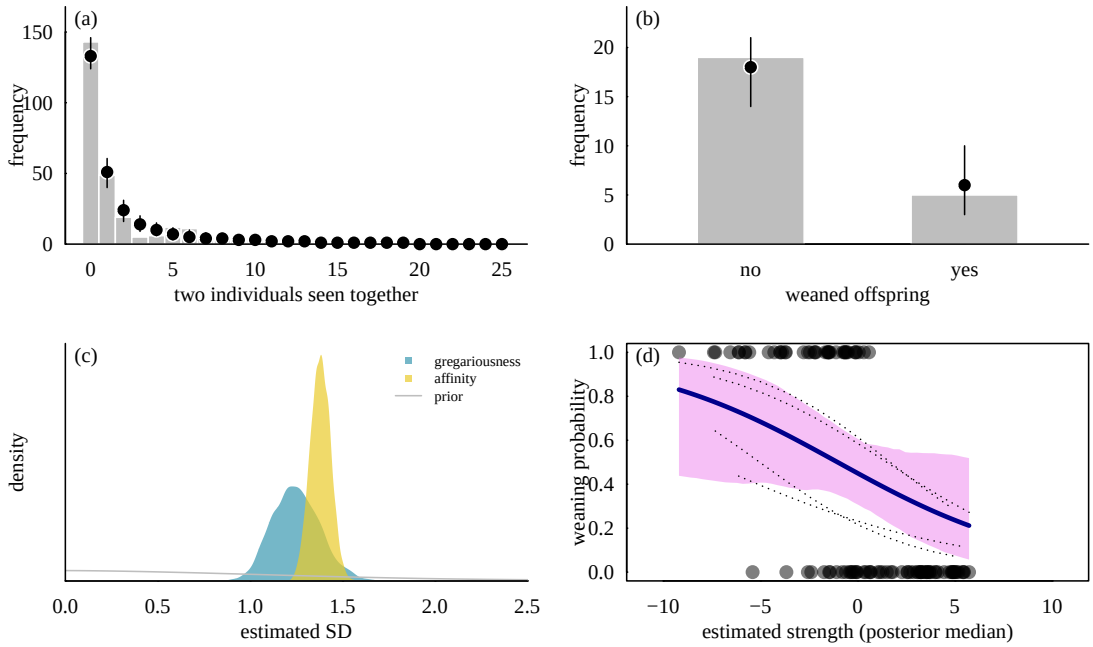

Figure 12: Reanalyzing Menz et al. 2020. Panels (a) and (b) show posterior predictive checks for association patterns (how often dyads were seen together) and the outcome for the reproductive success measure (number of females that weaned successfully) from 50 draws. The bars in each panel correspond to the observed frequencies. The circles and the vertical lines represent the median value predicted by the model alongside the range of predicted values from the 50 draws. These checks only reflect year 1. Panel (c) shows the posterior distributions of the estimated variation in the two sociality components. Panel (d) shows the model predictions for weaning using estimated strength from association data (plus 89% credible interval). Dotted lines are predictions for the four different years.

The study of sociality, how and why individuals interact, has a long history (Hinde, 1976; Ostner & Schülke, 2018). In this paper, we were not directly concerned with the theoretical framework of social relationships per se, i.e., we ignored critical dimensions of relationships like predictability, stability, symmetry, diversity, and tenor (Cords & Aureli, 2000; Fischer et al., 2017; Hinde, 1976; Ostner & Schülke, 2018; Silk et al., 2013). Our primary focus was on understanding the strength of relationships in an affiliative context. We argued that we could use observed dyadic interactions to estimate contributions from the group (baseline), the dyad, and the individual to relationship strength. This decomposition is important because when studies describe fitness consequences of sociality, this is necessarily done on an individual level: *individuals* survive longer or produce more offspring (Snyder-Mackler et al., 2020). But it remains unclear what exactly is the underlying mechanism if quantities like 'stronger relationships' or 'higher centrality' are used to predict fitness outcomes if it is unknown whether a dyadic relationship exists in the first place. A more parsimonious explanation is that if interaction patterns are entirely driven by gregariousness such correlations *could be* misleading. In any case, getting a grasp on the mechanisms requires that we quantify the underlying individual and dyadic components. Only after that we can look meaningfully into individual-level summaries of sociality (e.g., 'stronger relationships' with the gregariousness component removed).

An issue that exacerbates the potential for generating misleading results is the nature of how dyad-level data are broken down to the individual level (often termed 'social integration', Snyder-Mackler et al. (2020)). The way social integration has been quantified from dyadic interaction or association data varies substantially between studies and this undermines meaningful comparisons across studies (Ellis et al., 2019; Ostner & Schülke, 2018; Schülke et al., 2022; Snyder-Mackler et al., 2020). One of the essential critical advantages of our approach is that we provide a principled workflow that eliminates at least some of the researcher degrees of freedom when estimating social integration. To be explicit, we are not saying that previous approaches necessarily led to wrong conclusions. However, we want to emphasize that the model we suggest is tractable and interpretable because of its generative nature and if applied should allow us to arrive at results that are more readily comparable across studies.

One crucial issue to keep in mind is what the overall question is when analyzing interaction patterns. There are scenarios in which the actual interaction frequency or duration may be more informative compared to estimating the underlying gregariousness and affinity. If the goal is to study, for example, disease transmission or transmission of cultural techniques or knowledge, then it might not be relevant to partial out the individual component of network connections (Farine & Whitehead, 2015; Godde et al., 2013). If a dyad interacts very often then this might facilitate *transmission* irrespective of whether the interaction frequency is due to individual or dyadic processes. But even then, it might be relevant to quantify the underlying processes because more gregarious individuals might be more susceptible to carry the disease or information in the first place. And more generally, it seems advisable to actually *model* interaction frequencies instead of using point estimates as truth.

One drawback of our approach is that it likely requires coding up the models in Stan when going beyond relatively simple descriptions of the sociality axes. It seems unlikely that it is possible to develop a software package that, while simplifying the implementation, does justice to all the possible combinations and peculiarities of all the different study systems and designs. For example, our package provides only functions to estimate the sociality components but does not allow fitting more complex models in a formula-syntax approach. However, the effort to code up models like the ones we presented is well worth it because of the advantages we can gain from the insights such models can generate.

In this light, we see several ways the modeling framework could be extended, which we already, at least partly, incorporated in the applied examples above. First, future extensions could handle directed behavior (given versus received). Second, if data were recorded over multiple sampling periods, this could be accounted for with varying estimates for period (see the kangaroo example). Third, actual network metrics could be included, if required, like clustering, centrality, or betweenness (kangaroo example). Fourth, it is possible that the individual and dyadic components correlate; for example, individuals with higher individual propensities also form stronger relationships.

Conclusion

Understanding interaction patterns is vital to understand relationships. This study presents a principled modeling framework to estimate the underlying latent sociality components from interaction data on the individual level, the dyad, and the group. We provided various simulated and empirical examples to illustrate and evaluate its application. This workflow provides a reasonable extension of our toolbox to understand how and why individuals interact and to investigate the consequences of variation in the different sociality components on other phenomena of interest.

Appendix

A1 Stan code

There are three ways to obtain the raw Stan code for the models presented in this manuscript. The easiest way is to browse to https://github.com/gobbios/bamoso/tree/main/inst/extdata, where all model code is accessible.

Raw Stan files can also be found in the source code of the accompanying R package `bamoso`. Unzip the file `bamoso_0.1.0.tar.gz` (which is part of the supplements). The Stan code files are in the directory `inst/extdata`.

The Stan code is also accessible via the `bamoso` package itself. This requires that the package has been successfully installed (see supplementary file `repl_00_readme.md` for instructions).

- general model (used in figures 3 - 5, 7 - 11)

```
f <- system.file("extdata/interaction_model.stan", package = "bamoso")
mod <- cmdstan_model(stan_file = f, compile = FALSE)
mod$print()
```

- workflow example (used in figure 6)

```
f <- system.file("extdata/workflow_example.stan", package = "bamoso")
mod <- cmdstan_model(stan_file = f, compile = FALSE)
mod$print()
```

- kangaroo example (used in figure 12)

```
f <- system.file("extdata/bayesian_kangaroos.stan", package = "bamoso")
mod <- cmdstan_model(stan_file = f, compile = FALSE)
mod$print()
```

A2 Uncertainty

In this section, we illustrate how varying sampling effort results in varying degrees of uncertainty in assessing the different components in our model. Imagine two researchers observing the same group of animals during different periods. For the sake of the example, we assume that the underlying behavior of the subjects is the same, as is group composition. The only difference between the two observation periods is the time spent observing the animals. Intuitively, the more time we spend observing, the more precise our estimates of behavioral rates should become. We illustrate this here for a simulated example in which we observed a countable behavior. The only difference in the data generation was the time we allowed individuals to interact (statistically speaking, we varied the offset term). In one case, we observed the group for, say, two hours (which represents low observation effort in our example). In the second case, 100 hours were spent observing the animals (high observation effort). We set the baseline rate of the behavior to -0.7. Since we simulated a count variable, we expected an average pair of two average individuals to interact $exp(-0.7) = 0.5$ times per hour, which amounts to approximately 1 expected interaction ($exp(-0.7 + log(2))$) with low observation effort and approximately 50 interactions ($exp(-0.7 + log(100))$) for high observation effort. The difference is that observing 50 interactions in 100 hours should leave us much more confident that the true rate is about 0.5 interactions per hour than observing 1 interaction in 2 hours.

Figure A1 visualizes this intuition. We fitted our model to data generated with the above-mentioned settings. Across all parameters shown in figure A1, the posteriors are narrower in the scenario with higher observation effort.

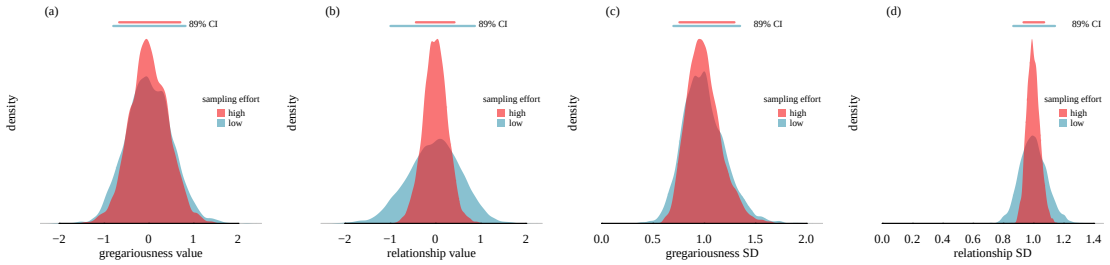

Figure A1: Uncertainty in gregariousness and relationship estimation. Observed behavior was a count variable. (a) and (b) show posteriors for one randomly selected individual and dyad, respectively. (c) and (d) show the posteriors for the estimated group-level variation for gregariousness and relationships. Within each panel, the posteriors were centered to aid visual comparison ((a) and (b) were centered at 0, (c) and (d) were centered at 1).

A3 Network strength and clustering in Stan

In order to implement the analysis of the kangaroo data (figure 12), we had to code functions that calculate strength and the local clustering coefficient from adjacency matrices in Stan. As reference we used the `igraph` package (Csardi & Nepusz, 2006), which the original authors also used. Figure A2 shows for a simulated data set that our Stan functions generated results identical to `igraph`'s `transitivity(graph, type = "weighted")` (for clustering) and `strength(graph)` (for strength).

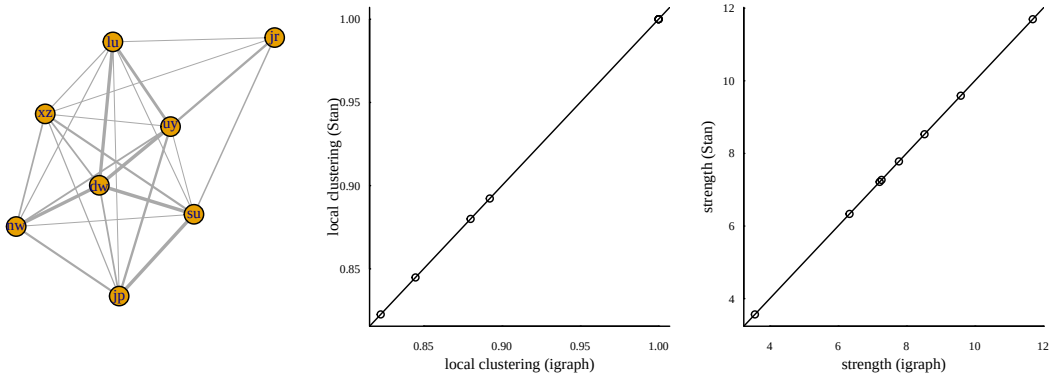

Figure A2: Functions for obtaining local clustering and strength written in Stan produce identical results as the `igraph` package.

References

Adams, M. J., Bonaventura, M., Ostner, J., Schülke, O., de Marco, A., Thierry, B., Engelhardt, A., Widdig, A., Gerald, M. S., & Weiss, A. (2015). Personality structure and social style in macaques. *Journal of Personality and Social Psychology*, *109*, 338–353. https://doi.org/10.1037/pspp0000041

Amici, F., Widdig, A., MacIntosh, A. J. J., Francés, V. B., Castellano-Navarro, A., Caicoya, A. L., Karimullah, K., Maulany, R. I., Ngakan, P. O., Hamzah, A. S., & Majolo, B. (2020). Dominance style only partially predicts differences in neophobia and social tolerance over food in four macaque species. *Scientific Reports*, *10*, 22069. https://doi.org/10.1038/s41598-020-79246-6

Balasubramaniam, K. N., Dittmar, K., Berman, C. M., Butovskaya, M. L., Cooper, M. A., Majolo, B., Ogawa, H., Schino, G., Thierry, B., & de Waal, F. B. M. (2012). Hierarchical steepness and phylogenetic models: Phylogenetic signals in *Macaca*. *Animal Behaviour*, *83*, 1207–1218. https://doi.org/10.1016/j.anbehav.2012.02.012

Bergman, T. J., & Beehner, J. C. (2015). Measuring social complexity. *Animal Behaviour*, *103*, 203–209. https://doi.org/10.1016/j.anbehav.2015.02.018

Bürkner, P.-C. (2017). `brms`: An `R` package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28. https://doi.org/10.18637/jss.v080.i01

Cairns, S. J., & Schwager, S. J. (1987). A comparison of association indices. *Animal Behaviour*, *35*(5), 1454–1469. https://doi.org/10.1016/S0003-3472(87)80018-0

Cameron, E. Z., Setsaas, T. H., & Linklater, W. L. (2009). Social bonds between unrelated females increase reproductive success in feral horses. *Proceedings of the National Academy of Sciences*, *106*, 13850–13853. https://doi.org/10.1073/pnas.0900639106

Carter, A. J., Menz, C. S., Best, E. C., Freeman, N. J., Dwyer, R. G., Blomberg, S. P., & Goldizen, A. W. (2020). *Higher sociability leads to lower reproductive success in female kangaroos*. Dryad. https://doi.org/10.5061/dryad.79cnp5hsb

Cords, M., & Aureli, F. (2000). Reconciliation and relationship qualities. In F. Aureli & F. B. M. de Waal (Eds.), *Natural conflict resolution* (pp. 177–198). University of California Press.

Csardi, G., & Nepusz, T. (2006). The `igraph` software package for complex network research. *InterJournal*, *Complex Systems*, 1695. https://igraph.org

Dal Pesco, F., Trede, F., Zinner, D., & Fischer, J. (2022). Male-male social bonding, coalitionary support and reproductive success in wild Guinea baboons. *Proceedings of the Royal Society B: Biological Sciences*, *289*, 20220347. https://doi.org/10.1098/rspb.2022.0347

Drewe, J. A., & Perkins, S. E. (2015). Disease transmission in animal social networks. In J. Krause, R. James, D. W. Franks, & D. P. Croft (Eds.), *Animal social networks* (pp. 95–110). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679041.003.0010

Duboscq, J., Micheletta, J., Perwitasari-Farajallah, D., Engelhardt, A., & Neumann, C. (2023). Investigating the relationship between sociality and reproductive success in wild female crested macaques, *Macaca nigra*. *International Journal of Primatology*. https://doi.org/10.1007/s10764-023-00376-5

Ellis, S., Snyder-Mackler, N., Ruiz-Lambides, A., Platt, M. L., & Brent, L. J. N. (2019). Deconstructing sociality: The types of social connections that predict longevity in a group-living primate. *Proceedings of the Royal Society B: Biological Sciences*, *286*, 20191991. https://doi.org/10.1098/rspb.2019.1991

Farine, D. R., & Strandburg-Peshkin, A. (2015). Estimating uncertainty and reliability of social network data using Bayesian inference. *Royal Society Open Science*, *2*, 150367. https://doi.org/10.1098/rsos.150367

Farine, D. R., & Whitehead, H. (2015). Constructing, conducting and interpreting animal social network analysis. *Journal of Animal Ecology*, *84*, 1144–1163. https://doi.org/10.1111/1365-2656.12418

Fischer, J., Farnworth, M. S., Sennhenn-Reulen, H., & Hammerschmidt, K. (2017). Quantifying social complexity. *Animal Behaviour*, *130*, 57–66. https://doi.org/10.1016/j.anbehav.2017.06.003

Frère, C. H., Krützen, M., Mann, J., Connor, R. C., Bejder, L., & Sherwin, W. B. (2010). Social and genetic interactions drive fitness variation in a free-living dolphin population. *Proceedings of the National Academy of Sciences*, *107*(46), 19949–19954. https://doi.org/10.1073/pnas.

1007997107

Gabry, J., & Češnovar, R. (2022). *`cmdstanr`: `R` interface to "CmdStan"*. https://mc-stan.org/cmdstanr/

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society: Series A*, *182*(2), 389–402. https://doi.org/10.1111/rssa.12378

Godde, S., Humbert, L., Côté, S. D., Réale, D., & Whitehead, H. (2013). Correcting for the impact of gregariousness in social network analyses. *Animal Behaviour*, *85*, 553–558. https://doi.org/10.1016/j.anbehav.2012.12.010

Hinde, R. A. (1976). Interactions, relationships and social structure. *Man*, *11*(1), 1–17. https://doi.org/10.2307/2800384

Kajokaite, K., Whalen, A., Koster, J., & Perry, S. (2022). Social integration predicts survival in female white-faced capuchin monkeys. *Behavioral Ecology*, *33*, 807–815. https://doi.org/10.1093/beheco/arac043

Kappeler, P. M. (2019). A framework for studying social complexity. *Behavioral Ecology and Sociobiology*, *73*(1), 13. https://doi.org/10.1007/s00265-018-2601-8

Kappeler, P. M., & van Schaik, C. P. (2002). Evolution of primate social systems. *International Journal of Primatology*, *23*(4), 707–740. https://doi.org/10.1023/A:1015520830318

Koster, J., Leckie, G., & Aven, B. (2020). Statistical methods and software for the multilevel social relations model. *Field Methods*, *32*, 339–345. https://doi.org/10.1177/1525822X19889011

Koster, J., Leckie, G., Aven, B., & Charlton, C. (2019). Methods and software for the multilevel social relations model: A tutorial. *arXiv:1907.13629 [Stat]*. https://doi.org/10.48550/arXiv.1907.13629

Lusseau, D., Whitehead, H., & Gero, S. (2008). Incorporating uncertainty into the study of animal social networks. *Animal Behaviour*, *75*, 1809–1815. https://doi.org/10.1016/j.anbehav.2007.10.029

McFarland, R., Murphy, D., Lusseau, D., Henzi, S. P., Parker, J. L., Pollet, T. V., & Barrett, L. (2017). The "strength of weak ties" among female baboons: Fitness-related benefits of social bonds. *Animal Behaviour*, *126*, 101–106. https://doi.org/10.1016/j.anbehav.2017.02.002

Menz, C. S., Carter, A. J., Best, E. C., Freeman, N. J., Dwyer, R. G., Blomberg, S. P., & Goldizen, A. W. (2020). Higher sociability leads to lower reproductive success in female kangaroos. *Royal Society Open Science*, *7*, 200950. https://doi.org/10.1098/rsos.200950

Neumann, C., & Fischer, J. (2023). Extending Bayesian Elo-rating to quantify the steepness of dominance hierarchies. *Methods in Ecology and Evolution*, *14*, 669–682. https://doi.org/10.1111/2041-210X.14021

Nightingale, G., Boogert, N. J., Laland, K. N., & Hoppitt, W. (2015). Quantifying diffusion in social networks: A bayesian approach. In J. Krause, R. James, D. Franks, & D. P. Croft (Eds.), *Animal social networks* (pp. 38–52). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679041.003.0005

Ostner, J., & Schülke, O. (2018). Linking sociality to fitness in primates: A call for mechanisms. *Advances in the Study of Behavior*, *50*, 127–175. https://doi.org/10.1016/bs.asb.2017.12.001

Pepper, J. W., Mitani, J. C., & Watts, D. P. (1999). General gregariousness and specific social preferences among wild chimpanzees. *International Journal of Primatology*, *20*, 613–632. https://doi.org/10.1023/A:1020760616641

Rathke, E.-M., & Fischer, J. (2021). Social aging in male and female barbary macaques. *American Journal of Primatology*, *83*, e23272. https://doi.org/10.1002/ajp.23272

Schülke, O., Anzà, S., Crockford, C., De Moor, D., Deschner, T., Fichtel, C., Gogarten, J. F., Kappeler, P. M., Manin, V., Müller-Klein, N., Prox, L., Sadoughi, B., Touitou, S., Wittig, R. M., & Ostner, J. (2022). Quantifying within-group variation in sociality—covariation among metrics and patterns across primate groups and species. *Behavioral Ecology and Sociobiology*, *76*(4), 50. https://doi.org/10.1007/s00265-022-03133-5

Schülke, O., Bhagavatula, J., Vigilant, L., & Ostner, J. (2010). Social bonds enhance reproductive success in male macaques. *Current Biology*, *20*(24), 2207–2210. https://doi.org/10.1016/j.cub.2010.10.058

Silk, J. B., Altmann, J., & Alberts, S. C. (2006). Social relationships among adult female baboons (*Papio cynocephalus*) i. Variation in the strength of social bonds. *Behavioral Ecology and*

Sociobiology, *61*, 183–195. https://doi.org/10.1007/s00265-006-0249-2

Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., Wittig, R. M., Seyfarth, R. M., & Cheney, D. L. (2010). Strong and consistent social bonds enhance the longevity of female baboons. *Current Biology*, *20*, 1359–1361. https://doi.org/10.1016/j.cub.2010.05.067

Silk, J. B., Cheney, D. L., & Seyfarth, R. M. (2013). A practical guide to the study of social relationships. *Evolutionary Anthropology*, *22*, 213–225. https://doi.org/10.1002/evan.21367

Snyder-Mackler, N., Burger, J. R., Gaydosh, L., Belsky, D. W., Noppert, G. A., Campos, F. A., Bartolomucci, A., Yang, Y. C., Aiello, A. E., O'Rand, A., Harris, K. M., Shively, C. A., Alberts, S. C., & Tung, J. (2020). Social determinants of health and survival in humans and other animals. *Science*, *368*, eaax9553. https://doi.org/10.1126/science.aax9553

Surbeck, M., Girard-Buttoz, C., Boesch, C., Crockford, C., Fruth, B., Hohmann, G., Langergraber, K. E., Zuberbühler, K., Wittig, R. M., & Mundry, R. (2017). Sex-specific association patterns in bonobos and chimpanzees reflect species differences in cooperation. *Royal Society Open Science*, *4*, 161081. https://doi.org/10.1098/rsos.161081

Thierry, B. (2007). Unity in diversity: Lessons from macaque societies. *Evolutionary Anthropology*, *16*, 224–238. https://doi.org/10.1002/evan.20147

Thierry, B., Aureli, F., Nunn, C. L., Petit, O., Abegg, C., & de Waal, F. B. M. (2008). A comparative study of conflict resolution in macaques: Insights into the nature of trait covariation. *Animal Behaviour*, *75*, 847–860. https://doi.org/10.1016/j.anbehav.2007.07.006

Thompson, N. A., & Cords, M. (2018). Stronger social bonds do not always predict greater longevity in a gregarious primate. *Ecology and Evolution*, *8*, 1604–1614. https://doi.org/10.1002/ece3.3781

Wey, T. W., & Blumstein, D. T. (2012). Social attributes and associated performance measures in marmots: Bigger male bullies and weakly affiliating females have higher annual reproductive success. *Behavioral Ecology and Sociobiology*, *66*, 1075–1085. https://doi.org/10.1007/s00265-012-1358-8

Whitehead, H. (2008). *Analyzing animal societies: Quantitative methods for vertebrate social analysis*. University of Chicago Press.

Whitehead, H., & James, R. (2015). Generalized affiliation indices extract affiliations from social network data. *Methods in Ecology and Evolution*, *6*, 836–844. https://doi.org/10.1111/2041-210X.12383

Young, C., Majolo, B., Heistermann, M., Schülke, O., & Ostner, J. (2014). Responses to social and environmental stress are attenuated by strong male bonds in wild macaques. *Proceedings of the National Academy of Sciences*, *111*(51), 18195–18200. https://doi.org/10.1073/pnas.1411450111