Five misunderstandings in animal social network analysis

Daniel Redhead[1,2,3 †, ‡], Ben Kawam[3,4, †, ‡], Jean-Gabriel Young[5,6], Daniel Franks[7,8], Conner S. Philson[9,10], Marijtje van Duijn[1,2], Jordan Hart[9], Mary Brooke McElreath[3,11], Richard McElreath[3], Eleanor A. Power[12, 3], Cody T. Ross[3], Sebastian Sosa[3], Christian Steglich[1,2], Michael Weiss[9,14], and Lauren J. N. Brent[9]

Abstract

Animal social network analysis has become central to behavioural ecology, offering powerful tools to explore the links between social behaviour and ecological or evolutionary processes. While rooted in the broader field of social network analysis, the methods used in animal studies have diverged from contemporary practices in the broader field. This divergence has led to conflicting guidance on best practices and in confusion among behavioural ecologists on how to analyse animal network data. Here, we identify and resolve five key misunderstandings in animal social network analysis. We start by tracing a brief history of the field. We then define each misunderstanding, discuss the flaws in the methodology that they are premised on, and outline their consequences for scientific inference. Finally, we examine how these issues might be overcome by using models that reflect the generative mechanisms that underlay the structural features of social network data—building upon tools and ideas from the wider social networks literature. Our goal is to help bridge the gap between behavioural ecologists and the broader social network analysis community, encouraging methodological realignment and facilitating fundamental advances in how we understand the ecological and evolutionary foundations of animal social behaviour.

Keywords: Animal Behaviour, Behavioural Ecology, Sociality, Social Network Analysis, Social Relationships.

---

[1] Department of Sociology, University of Groningen, Grote Rozenstraat 31, 9712 TG Groningen, The Netherlands.
[2] Inter-University Center for Social Science Theory and Methodology, University of Groningen, Groningen, The Netherlands.
[3] Department of Human Behaviour, Ecology and Culture, Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, 04103 Leipzig, Germany.
[4] Department of Behavioural Ecology, University of Göttingen, Kellnerweg 6, 37077 Göttingen, Germany.
[5] Department of Mathematics and Statistics, The University of Vermont, Burlington, Vermont, USA.
[6] Vermont Complex Systems Institute, The University of Vermont, Burlington, Vermont, USA.
[7] Department of Biology, University of York, York, UK.
[8] The University of York Heslington, York, UK.
[9] Center for Research in Animal Behaviour, Department of Psychology, University of Exeter, Exeter, UK.
[10] Natural Reserve System, University of California - Santa Barbara, Santa Barbara, California, USA.
[11] Department for the Ecology of Animal Societies, Max Plank Institute of Animal Behavior, 78467 Konstanz, Germany
[12] Department of Methodology, London School of Economics and Political Science, London, UK.
[13] Santa Fe Institute, Santa Fe, New Mexico, USA.
[14] Center for Whale Research, Friday Harbor, UK.
[†] Corresponding authors: *Daniel Redhead*, d.j.redhead@rug.nl, and *Ben Kawam*, bkawam@dpz.eu.
[‡] These authors contributed equally to this work.

Introduction

Social network analysis (SNA) is an approach within the behavioural sciences that encompasses theoretical models and analytical tools centred on the study of groups, and the relations between units (*e.g.,* individuals) within these groups (Wasserman & Faust, 1994). Emerging from advances around the 1930s in sociology, physics, and graph theory (*e.g.,* Moreno, 1934), network analysis is used today in multiple areas of academic inquiry, from studies of food webs (Dunne et al., 2002) to brain connectivity (Bassett & Sporns, 2017), and from studies of workplace dynamics (Kilduff & Brass, 2010) to online behaviour (Falkenberg et al., 2024)—the latter being examples of "social network analysis" (Hidalgo, 2016). The use of social network analysis has now become widespread in ecology and evolution. This has, for example, allowed behavioural ecologists to examine the fitness benefits related to different types of sociality (*e.g.,* Formica et al., 2012), to identify the routes through which cooperation is maintained within and between populations (*e.g.,* Samuni & Surbeck, 2023) and how the structure of such cooperation networks produces social and material inequalities (Redhead, 2024), and to determine the social strategies of individuals in the face of ecological upheaval (*e.g.,* Testard et al., 2024), across a broad range of taxa.

Animal social network analysis (ASNA) often uses "centrality" metrics to quantify an individual's position within their social system, ratio indexes (*e.g.,* Simple Ratio Index, Dyadic Sociality Index) to quantify dyadic structure, and permutation tests to assess statistical significance. These methodological traditions and practices set ASNA apart from the broader field of SNA, which now largely focuses on the development and application of statistical models that directly, *i.e.*, parametrically, estimate the structural features of social networks. We will refer to such statistical models as *generative network models* throughout the manuscript. Here, the use of the word *generative* reflects the ability of these statistical models to generate predictions (or pseudo-data) about a group or population from parameters that instantiate theoretical assumptions or constraints about the individuals that are situated within a network (see Epstein, 2012; Kandler & Powell, 2018; Steglich & Snijders, 2022, for detailed outlines and approaches from relevant fields). Such generative models can make explicit how measurement error, reporting biases, sampling bias, and censoring affect the data that are ultimately recorded by practising scientists (*e.g.*, Redhead et al., 2023b; Sosa et al., 2025), which permits researchers to make scientific inferences that account for such factors.

In recent years, numerous studies have commented on the limitations of common statistical practices in the field of ASNA, highlighting the need for methodological innovation in how animal network data are analysed. Here, we identify five key misunderstandings in ASNA that underlie these methodological limitations. We first provide a brief history of ASNA, and structure our following sections by defining each misunderstanding—where each section title directly refers to the misunderstanding. We provide clarity on the inferential pitfalls of common approaches, and review extant solutions to these analytical issues that exist in the wider social networks literature—our subsection titles refer to these contributions. In doing this, we aim to help re-integrate ASNA into the broader field. Additionally, we aim to motivate important developments in statistical models for animal social networks, which will fundamentally advance our ability to understand the structure of animal societies.

A brief history of animal social network analysis

Researchers studying behaviour in non-human animals first turned their attention to social network analysis (then typically referred to as "sociometry") in the 1940s. In 1945, the primatologist Charles Ray Carpenter stated that "*Sociometric techniques are suggested as one means of measuring changes that result from social conditioning and learning*" (Carpenter, 1945, p. 61). Carpenter was an early proponent of the study of the social behaviour of animals, founding one of behavioural ecology's longest-running field stations in the world, Cayo Santiago (Clutton-Brock, 2016). It is perhaps no surprise that some of the first researchers to adopt ASNA had ties to this Caribbean field station. In the 1970s and 80s, Donald Sade and Diane Chepko-Sade published a series of papers on the social networks of the Cayo Santiago rhesus macaques (Sade, 1972; Sade, 1989), including the quantification of grooming clusters in a group before it fissioned (Chepko-Sade et al., 1989). E.O. Wilson, who was sat on Cayo Santiago when he had the eureka moment leading to his foundational 2000 book, *Sociobiology*, also promoted a network approach to compare and contrast the structure of animal societies.

Around the same time of Sociobiology's publication, the zoologist Robert Hinde published a conceptual framework to consolidate the growing number of findings on animal societies into a holistic theory (Hinde, 1976). Hinde's framework posited that relationships between individuals emerge from the patterning of their interactions, and social structure emerges from the patterning of relationships, with feedback loops existing between the three levels (Brent et al., 2011; Hinde, 1976). Referencing the Sades, and work on human social networks, Hinde's framework sat squarely within the sociometric approaches that were applied by his contemporaries to the study of human and non-human animals.

Although prominent early animal behaviour researchers, such as Tinbergen and Allee, wrote about relations between actors and receivers (Allee, 1931; Tinbergen, 1939), early examples of animal research that used social network analysis come from non-human primates. The social networks of non-human primates were quantified extensively in the 1970s-1990s, including work on sex differences in the social behaviour of bonnet macaque monkeys (Simonds, 1974), the social integration of high-ranking compared to low-ranking female baboons (Seyfarth, 1976), and on the strength of weak ties in chimpanzees and gorillas (Maryanski, 1987).

It was after the turn of the 21st century when studies of other types of animals began to adopt social network analysis *en masse* (*e.g.*, Krause et al., 2009; Wey et al., 2008; Whitehead, 2008). Proffered reasons why social network research on other animals was not widely adopted sooner include limited computational power, and the perception of a lack of network descriptors deemed "meaningful" to the study of animal behaviour (Krause et al., 2007; Whitehead, 2008). Along with social network analysis's increasing popularity in animal research, came regular use of terms like "new" and "novel" to describe the approach. These descriptions are misleading, as social networks have been fundamental in the study of behavioural ecology for over 40 years (Brent et al., 2011).

At the same time as the broadening taxonomic interest in ASNA came an increasing shift in the methodological approaches being applied in ASNA and SNA. During this time, researchers working in SNA began to increasingly develop and apply generative models to social network data (*e.g.,* Anderson et al., 1999; Snijders, 1996; Van Duijn et al., 2004), while ASNA remained focused on the computation of network metrics, permutation tests and the development of solutions to sampling problems that were somewhat unique to the field (Whitehead, 2008). This divergence might have been caused by a number of factors. First, behavioural ecologists who first adopted a network approach were housed in anthropology, psychology, or sociology departments because of their focus on non-human primates. The growing number of ecologists studying non-primate social networks were housed in biology, ecology, and evolution departments, and were less connected to human behavioural sciences, and therefore with the methodological advances in the broader field of SNA (Brent et al., 2011; Haraway, 2023). The divergence may further be explained by differences in the data that are collected when studying the sociality of humans (*e.g.*, through self-reports during surveys; De Bacco et al., 2023) compared to other animals (*e.g.*, focal sampling protocols, biologgers). Such data types are typically directly reflected in the statistical software used by analysts, meaning that the behavioural data collected by animal behavioural ecologists was often incompatible with the software developed in SNA. Regardless of its causes, the divide between SNA and ASNA has persisted, and even grown, to this day.

Current practices in animal social network analysis

Today, ASNA is often associated with a set of specific methodological practices, which consist of the following steps. First, researchers construct a social network from behavioural data—typically either by following certain individuals and observing their behavioural interactions (*i.e.*, 'focal follows') or through observations of spatial or behavioural association between individuals in a certain physical space (*e.g.*, scan sampling, 'gambit of the group'). Second, the network is constructed by computing an adjacency- or socio-matrix, whose entries are obtained by *aggregating* behavioural data collected over a given period of time into a number, quantifying the social aspect of interest. If individual animals were not sampled equally (*e.g.*, if some focal individuals were observed more frequently than others), researchers may then, subsequently, divide the chosen social quantity by a measure of sampling effort, by computing a Simple Ratio Index (SRI) or related measure, in an attempt to *control* for sampling imbalance (Whitehead, 1999; Whitehead, 2008). The typical methods used by behavioural ecologists subsequently derive *social network metrics* from the adjacency matrix, which may describe individual-level features (*e.g.*, strength, degree, eigenvector centrality) or network-level properties (*e.g.*, network density). Variation in these metrics is finally explained—usually through Null Hypothesis Significance Testing—using permutation procedures and/or standard multiple regression (Farine, 2017).

Although these methods are widely used in the field of behavioural ecology (*e.g.*, Earl et al., 2025; Samuni & Surbeck, 2023), they have received growing critical attention in recent years from applied statisticians and social network analysts. One major issue is that these methods assume that the social network of interest is *directly observed* (we will return to this point later). Moreover, the biological processes that generate the social network data are often implicit in the formal analysis, if not outright ignored. The only consistent focus on the actual process generating the data is through the tangential concept of dependent, or non-independent, social network data. However, such dependencies seem to create considerable confusion and are rarely treated formally. Researchers using multiple regressions to study the causes or consequences of social network structure rarely justify their statistical models using transparent causal assumptions about the process that generated the data. In addition, behavioural ecologists continue to use permutation procedures (Farine, 2017). Although these tools were once popular in SNA in the 1980s (Krackhardt, 1988), they were critiqued on logical grounds shortly after their introduction (Krackhardt, 1992), and have now largely been phased out outside of animal SNA due to their limitations (*e.g.,* Anderson & Robinson, 2001).

The statistical approaches often used in ASNA stand in contrast to contemporary methods in the rest of SNA. Generally speaking, models and methods in contemporary SNA have switched to estimating the *structure* of social networks explicitly. They directly incorporate theoretically-relevant social dependencies by including relevant parameters in the model (Robins et al., 2007a, 2007b; Snijders, 2017) or by specifying random-effects structures (Hart et al., 2023; Hoff et al., 2002; Pillinger et al., 2024; Ross et al., 2023; Snijders & Kenny, 1999; Van Duijn et al., 2004). These generative network models are rarely, if ever, applied in ANSA. Hesitance to apply such models remains no only due to a lack of awareness about these methodological developments in SNA, but is in large-part a consequence of several misunderstanding about how to analyse social interaction data. Here, we aim to clarify a number of these misunderstandings, and promote awareness of the methodological innovations made in SNA, bridging the gap between the two disciplines.

Misunderstanding 1: "Dependencies as a nuisance"

Social network data are known to be highly *inter-dependent*. In SNA, like in the rest of applied statistics, "dependent" simply means "statistically associated." A pair of random variables is dependent if knowing the value of one of them informs us about the value of the other. Statistical dependency is neither surprising nor unique to social network data: strong patterns of statistical associations are ubiquitous in nature.

A kind of dependency that has received a lot of attention in ASNA is the violation of a common statistical assumption: that data are independently and identically distributed (iid). This assumption essentially states that observations were sampled at random from the stochastic data-generating process implied by a statistical model (Lambert, 2018).[1] Social network data often violates the iid assumption baked into standard statistical models. This could be the case if some individuals have propensities (*e.g.*, like aggressiveness) that introduce correlations in behavioural outcomes (*e.g.*, increasing the aggressiveness of individuals a increases the probability of observing fighting events, y_{ab}, for all b), or if the network connections in one direction (the edges y_{ab}) correlate with those in the opposite direction (y_{ba}, *e.g.*, if there may be some form of reciprocity). When models that assume iid data are applied to data with dependencies, this can lead to biased parameter estimates (Kenny & Judd, 1986; Robinson, 1955) and to incorrect p-values (Anderson & Robinson, 2001). As a consequence, dependencies in animal social network data have generally been perceived as a nuisance that needs to be, in one way or another, eliminated (*e.g.*, Farine, 2017; Farine & Whitehead, 2015).

The problem with treating dependencies as a nuisance only is that dependencies in social network data partly result from the biological phenomena of interest. For instance, the objective of a study might be to investigate whether animals reciprocate behaviours in a population or to estimate how variation in an individual-level factor (*e.g.*, aggressiveness) affects social network structure. Such processes create specific patterns of correlation in social network data, which are precisely what one attempts to capture and *must not* be controlled away (Hart et al., 2022, 2023; Pearl et al., 2016). This situation—where dependencies bear the dual status of statistical threat and target of biological inference—has generated confusion about how to properly analyse animal social network data.

To solve this apparent conflict, it is useful to focus on two essential aspects of most social network analyses: (i) the target quantity of interest, the *estimand*, and (ii) the *data-generating process*.

The *estimand* is the quantitative goal of an analysis—a theoretical quantity, often causal, that is defined outside of any specific statistical model (Lundberg et al., 2021). Examples of estimands are: the average causal effect of a phenotypic trait on the propensity of individuals to direct interactions toward others, the direct effect of gregariousness (controlling for a set of mediators) on survival, or the causal effect of dyadic reciprocity (a dynamic process) on the pattern of interactions observed in a cross-sectional sample. In any case, it is crucial to clearly define the estimand when conducting an analysis of the causes or consequences of social structure, and to

---

[1] For example, consider the following statistical model:

$$y_{ab} \overset{\text{iid}}{\sim} \text{Poisson}(\lambda_{ab} = \exp(\alpha + \beta \cdot X_{ab})).$$

The variable y_{ab} refers to the number of observed social interactions, y, given by individuals, a, to other individuals, b. The $\overset{\text{iid}}{\sim}$ symbol means: "*is independently and identically distributed from*." λ_{ab} is the mean rate of interactions from a to b. It is described by a linear model that has been exponentiated to ensure that λ_{ab} remains positive. α is the intercept, and β is a slope describing the association between the log of λ_{ab} and the dyad-level covariate X_{ab}. The equation in this simple example implies that the variation in social network data y_{ab} is explained by two factors: the association with X_{ab}, and the stochastic variation described by a Poisson distribution (Blitzstein & Hwang, 2019). Any additional source of variation would result in a violation of the iid assumption.

state it transparently when communicating it to the rest of the community (Grosz et al., 2020; Kawam et al., 2024; McElreath, 2020).

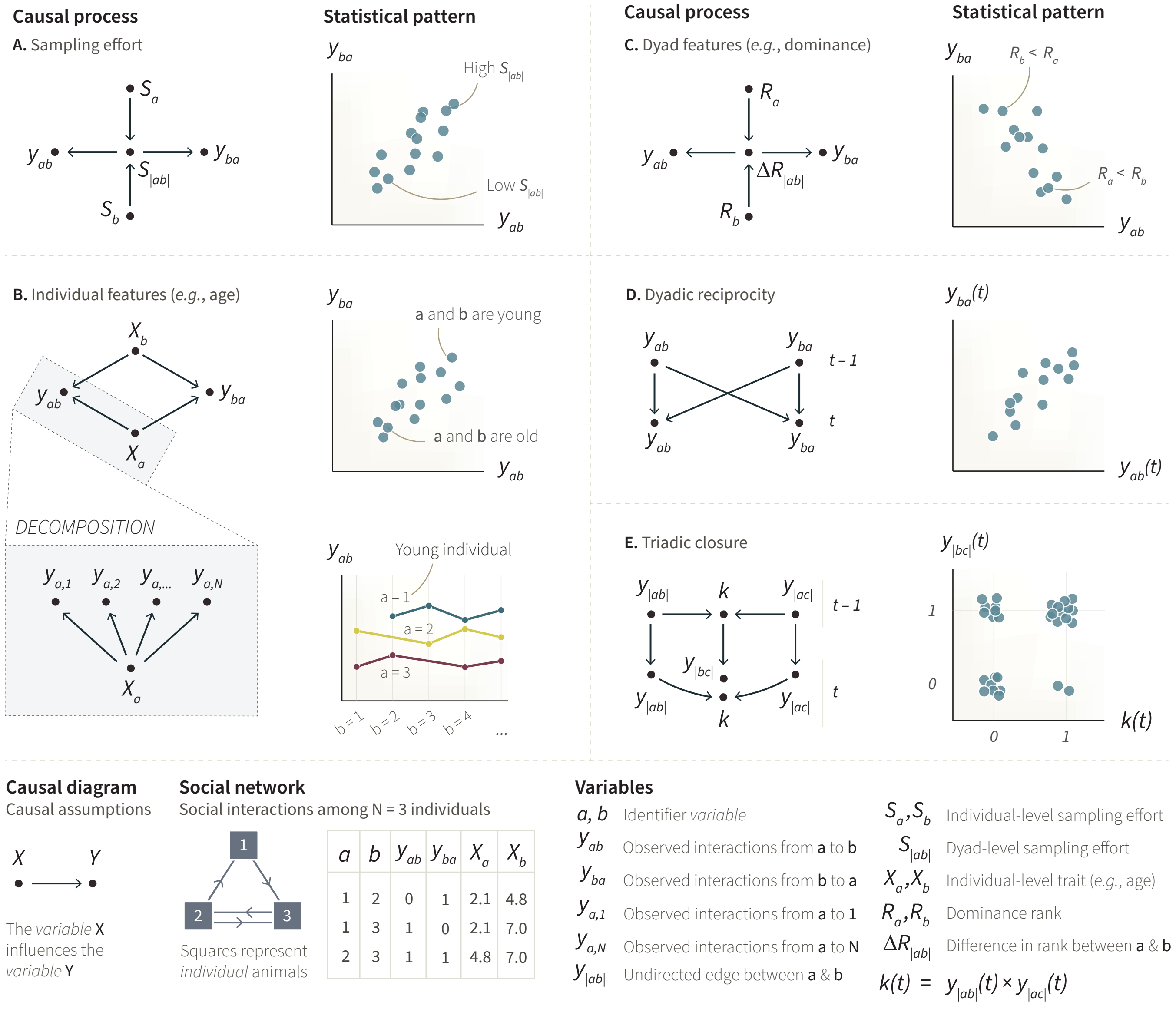
To produce a statistical estimate for the estimand, one needs to stipulate a model for the causal process generating social network data. This model should instantiate assumptions about the sampling and measurement on the one hand, and biological phenomena on the other. A data-generating process implies a specific set of statistical dependencies—which can be thought of as *symptoms* of the underlying causal structure. In Figure 1**A**-**E**, we illustrate a number of simple causal structures for social network data using causal diagrams, and show some of the dependencies that they imply.[2] In addition, we have translated these causal diagrams into generative simulations to illustrate how they create the association patterns shown in Figure 1—the simulation code can be found on our GitHub repository.

Figure 1 highlights the lack of a one-to-one relationship between observable dependencies and underlying causal structures. For instance, the causal diagrams in panels **A**, **B**, **C**, and **D**, all imply an association between y_{ab} and y_{ba}. Now, suppose that we were interested in studying dyadic reciprocity in a given population of animals. One may think that we can simply capture this process through the statistical relationship between y_{ab} and y_{ba}. However, we might observe a positive correlation between y_{ab} and y_{ba} even in the absence of dyadic reciprocity, recalling that *correlation does not imply causation*. This could result from the sampling regime (Figure 1**A**), or be because of individual-level features like age (Figure 1**B**). We might also observe a *lack* of association between y_{ab} and y_{ba} in the presence of reciprocity (*causation does not imply correlation*), if dominance rank counteracts its effect—*e.g.*, if we were to combine the causal structures of Figure 1**C** and 1**D**.

To deal with such potential confounding effects, researchers are required to combine the two ingredients above—the estimand and causal assumptions about the data-generating process—to build a custom-designed statistical model, whose role is to produce a statistical *estimate* for the target *estimand* (Kawam et al., 2024; McElreath, 2020; Pearl et al., 2016). In the simplest case, it might be a standard regression, but it can often be more complex and specific. Fortunately, social network data-generating processes share common characteristics across many data sets, and the generative network models that have been developed in the field can be used as a starting point for new analyses (Rawlings et al., 2023; Snijders, 2011). For instance, the Social Relations Model is a statistical model for dyadic interactions, both directed and undirected, that estimates the true dyadic interaction rates from noisy data, and models the unobserved causes at both the individual and dyadic levels (Kawam et al., 2024; McElreath, 2020; Redhead et al., 2023b; Ross et al., 2023). Another example is the Stochastic Blockmodel, which allows researchers to estimate the causal effects of categorical variables like kin groups, or the combination of sexes and age classes, on dyadic interactions (Jang et al., 2024; Redhead et al., 2023a).

---

[2] We have not included statistical dependencies that result from the *aggregation* of data. For instance, it is common for researchers to compute *node degree* from binary edge data, before subsequently decomposing the variation in node degree as functions of explanatory variables (Farine & Whitehead, 2015). Doing this—*i.e.* constructing individual-level indices from dyad-level observations—creates dependencies between individual-level measures. We generally recommend against using such indices as outcomes or predictors of statistical models, as doing so obscures the differential levels of uncertainty that we might have due to sampling variation, and the resulting model will not appropriately account for the dependencies in the data. Instead, we encourage researchers to model the data prior to its aggregation—analysts who are interested in studying the causes of individual-level variation with binary data might, for instance, model edge formation using a binomial model that estimates individual-level sender and receiver parameters using a varying-effect structure (Ross et al., 2023).



a	b	y_{ab}	y_{ba}	X_a	X_b
1	2	0	1	2.1	4.8
1	3	1	0	2.1	7.0
2	3	1	1	4.8	7.0

FIGURE 1 – *Five causal processes generating social network data, and the statistical dependencies that they create.* Throughout the five panels, causal processes are represented on the left using causal diagrams, and the resulting associations are shown on the right with idealised data (see our GitHub repository for simulated data). As shown by the legend at the bottom, the variables a and b are identifiers, which can take values 1 to N, where N is the number of individuals in a social group. Thus, variables that are indexed by a or b will take on the value that corresponds to the respective individuals, or dyad. For instance, y_{ab} represents the variable $\{y_{1,2}, y_{1,3}, \dots\}$, where $y_{1,2}$ is a measurement (*e.g.* a count) of social interactions (*e.g.* grooming, aggression) from individual 1 to individual 2. Similarly, y_{ba} represents $\{y_{2,1}, y_{3,1}, \dots\}$. X_a corresponds to the variable $\{X_1, X_2, \dots, X_N\}$. X_b represents the variable $\{X_1, X_2, \dots, X_N\}$ as well, but for the other individuals across dyads (see example with $N = 3$ at the bottom of the figure). We denote symmetric dyad-level variables using vertical bars ($S_{|ab|} = S_{|ba|}$), and directed dyad-level variables without vertical bars ($y_{ab} \neq y_{ba}$). **A.** The individual-level sampling effort S_a and S_b (*e.g.*, number of hours an individual has been observed during focal protocols) determine the dyad-level sampling effort $S_{|ab|}$ (*e.g.*, number of hours the dyad has been observed). In turn, $S_{|ab|}$ affects whether interactions are observed or not, *i.e.* it causes the observations y_{ab} and y_{ba}. The scatterplot shows an association between y_{ab} and y_{ba} resulting from their common cause, $S_{|ab|}$, forming an *open path* between them on the causal graph. **B.** On the top left of the panel, we represent an individual-level feature X affecting how many interactions animals give to others ($X_a \rightarrow y_{ab}$ and $X_b \rightarrow y_{ba}$), and how many interactions they receive from others ($X_a \rightarrow y_{ba}$ and $X_b \rightarrow y_{ab}$). X might, for instance, represent age if older individuals tend to withdraw from social activity, and thus, give and receive fewer interactions. The resulting open paths between y_{ab} and y_{ba} create an association between them, as shown on the scatterplot on the top-right of the panel. At the bottom left, we decompose the arrow going from X_a to y_{ab}. We show how X_a affects $y_{a,1}$ to $y_{a,N}$, where $a \neq b$. It implies, for instance, that X_1 affects $y_{1,2}, y_{1,3}, y_{1,4}$, etc.; that X_2 affects $y_{2,1}, y_{2,3}, y_{2,4}$, and so on. This diagram shows that the variables $y_{a,1}$ to $y_{a,N}$ are influenced by a common cause X_a, and thus, we know that these variables should be associated. In other words, an individual (a) should show some level of "stability," or "consistency" in how it interacts with other individuals (b). We represent this stability across individuals at the bottom-right panel of the panel, for three individuals a (coloured lines and points), and for four individuals b (*x-axis*). See the next page for the rest of the caption.

Figure 1 – *Continued:* **C.** The dominance rank of a (R_a) and the rank of b (R_b) together determine the difference in rank between them ($\Delta R_{|ab|} = R_b - R_a$). In turn, this *dyad-level feature* affects the observed interactions y_{ab} and y_{ba}—suppose, for instance, that individual monkeys groom their dominants more than their subordinates. The open path between y_{ab} and y_{ba} creates an association between them, as shown on the right of the panel. Note that from the graph alone, we can read that y_{ab} and y_{ba} should be associated, but not whether this association should be positive or negative. **D.** The number of interactions that a gives to b at time t is affected by how b interacted with a at time $t-1$, $y_{ba}(t-1)$, because of reciprocity; and by $y_{ab}(t-1)$ because the directed ties have some level of stability—and inversely for $y_{ba}(t)$. Thus, at a given time unit t, there are two open paths between y_{ab} and y_{ba}: $y_{ab}(t) \leftarrow y_{ab}(t-1) \rightarrow y_{ba}(t)$, and $y_{ab}(t) \leftarrow y_{ba}(t-1) \rightarrow y_{ba}(t)$, causing the association represented on the right-hand side. **E.** Consider all triads (a, b, c) in a social network whose dyadic edges are non-directed and binary ($y_{|a,b|}, y_{|a,c|}, y_{|b,c|} \in \{0, 1\}$). Suppose that if two of the three possible edges in a triad exist at time $t-1$, these two edges increase the probability that a third edge is formed at time t (*e.g.*, $y_{|bc|}(t) = 1$): *"friends of friends become friends"*—a phenomenon known as triadic closure. This causal process gives rise to several dependency patterns, including the following: at any point in time t, if we know that $k(t) = y_{|ab|}(t) \cdot y_{|bc|}(t) = 1$, it is likely that the triad's third edge exists as well: $y_{|bc|}(t) = 1$. This pattern of dependence, shown on the right-hand side, can be read from the causal diagram, since there are direct paths between $k(t)$ and $y_{|bc|}(t) = 1$.

Misunderstanding 2: "Statistical models for social network structure require evenly sampled data"

As we have highlighted above (*e.g.*, in Figure 1**A**), it is common for individuals and dyads to be sampled unevenly in animal social network studies—in fact, it is probably the norm. Whether behaviour is recorded using continuous-time (*e.g.*, focal sampling) or discrete-time sampling procedures (*e.g.*, scan sampling, gambit of the group), some individuals will generally be observed more than others. As a consequence, the observed ties between animals are often uninterpretable without factoring in sampling effort, and thus, unadjusted downstream statistical estimates can be noisy or confounded (Farine, 2017; Sosa et al., 2025).

It has been argued in ASNA that statistical models for social network structure, as opposed to permutation approaches, require that all study animals and dyads are evenly sampled (Farine, 2024). This is fortunately untrue. Statistical models designed in light of the social network data-generating process can integrate variation in sampling effort—and, crucially, they can do so while maintaining the associated levels of inferential *uncertainty*. In contrast, index variables like the Simple Ratio Index, the Interaction Index, or the Composite Dyadic Sociality Index (CSI/DSI) divide out sample size in sampling effort and give sparsely-observed data-points disproportionate leverage (see Hart et al., 2023; Sosa et al., 2025).

In Figure 2, we illustrate how we can think about sampling effort as an integral part of the social network data-generating process. Imagine a hypothetical population of animals, where an individual-level trait, X_a, affects how many interactions an animal, a, gives to its conspecifics (*i.e.*, higher y_{ab} across partners b). Additionally, a high value for X_a leads a to be sampled more frequently (*i.e.*, higher dyad-level sampling effort, $S_{|ab|}$). Suppose further that we wish to estimate the causal effect of X_a on y_{ab} across dyads (a, b). From the structure of the causal graph, we see that the estimate for the effect of interest will be biased if sampling effort is not accounted for (Figure 2**A**; for an introduction to Directed Acyclic Graphs, see Pearl et al., 2016). Indeed, if individuals with high values of X are observed to interact a lot—that is, to have high values of y—, it might not be due to a biological effect, but instead because individuals with high X tend to be observed more often.

In this scenario, conditioning on $S_{|a,b|}$ is necessary to correct for the bias introduced by uneven sampling, and doing so allows us to recover the causal effect of interest (Figure 2**B**). If y_{ab} is a count, it can be modelled using a Poisson distribution—a canonical choice for counts generated under a fixed rate (Blitzstein & Hwang, 2019):

$$y_{ab} \sim \text{Poisson}(\lambda_{ab} \cdot S_{|ab|}),$$

Where λ_{ab} represents the true rate of interactions from a to b per *one* time unit. Thus, the number of observed interactions from a to b can be considered a random draw from a Poisson distribution, whose mean is simply the dyad-specific rate multiplied by the number of sampling periods that the dyad has been observed for. The estimated true rate can then be described by

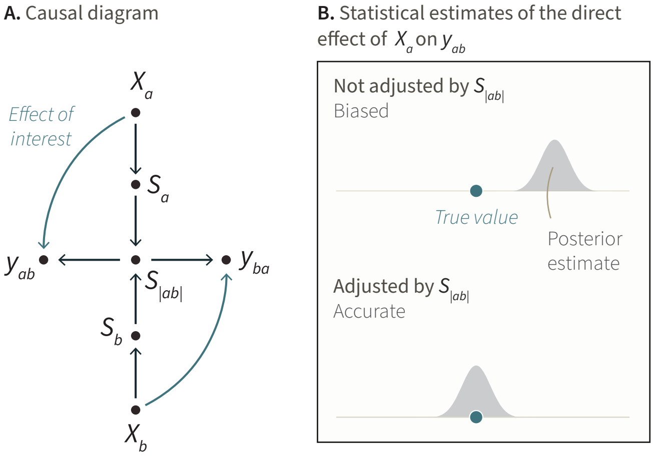


Figure 2 – *Integrating sampling effort into animal social network analysis.* **A.** The causal graph is a qualitative representation of the process generating social network data in a hypothetical population of animals. An individual level trait, X_b (*e.g.*, colouration), may affect the number of interactions that a gives to *b*. This effect is our estimand, marked in blue. X_b might also affect the individual-level sampling effort S_b (*e.g.*, how many hours that *b* is observed for), if animals with less vibrant colouration are harder for researchers to locate and observe. S_a, together with S_b (the sampling effort of the other individual in the dyad), determine the dyad-level sampling effort, $S_{|ab|}$ (*e.g.*, how many hours was one of the two individuals in the dyad observed for). In turn, $S_{|ab|}$ affects the observed number of interactions from a to *b*, y_{ab}. To identify the effect of interest, we must close the directed path going from X_a to y_{ab} that passes through $S_{|ab|}$. **B.** The panel represents idealised posterior estimates for the effect of interest, for a model that is not stratified by $S_{|ab|}$ (top), and for a model that is stratified by $S_{|ab|}$ (bottom). In the former case, the path going from $X_{[a]}$ to y_{ab} that passes through $S_{|ab|}$ is open, and thus, the estimate is biased (it does not overlap with the true value, in blue). In the latter, the path is closed, and consequently, the estimate is accurate (it perfectly overlaps the true value, in blue). We reproduce this pattern in our GitHub repository, using data-simulations and Bayesian multilevel models.

a generalised linear model:

$$\lambda_{ab} = \exp(\alpha + \beta \cdot X_a + ...),$$

Where the slope β can recover the true, unbiased, causal effect of X_a on the rate of interaction λ_{ab}, now that $S_{|a,b|}$ has been conditioned on (Figure 2**B**). The ellipsis (...) is a placeholder for any other parameters that may be included in the model. Which parameters (*e.g.,* covariate-related predictors) to include in the generalised linear model will again depend on assumptions about the data-generating process. They may, for instance, correspond to the varying-effect (or random effects) structure of the Social relations Models, as we show in the full example developed in our GitHub repository.

Misunderstanding 3: "Network permutations are sufficient to deal with dependencies in Social network data"

Current practices in ASNA almost exclusively treat dependencies within networks as an inferential threat, and typically aim to control for such dependencies when performing hypothesis tests (*e.g.,* calculating *p*-values) through two common permutation procedures. Both procedures aim to generate null distributions (*i.e.,* "null models") by breaking the link between observed data and hypothesized effects, such as the effect of a trait on social connectivity, while preserving some structural features of the data.

The first approach, "network permutation," entails the random shuffling of the rows and columns of an adjacency matrix that has been constructed by aggregating all observed ties over a given study period to represent a network (*i.e.,* through a quadratic assignment procedure (QAP); see Krackhardt, 1987). Permutations (*i.e.*, random shuffles) are repeated many times to produce a distribution of values where the identity of an individual no longer matters, aiming to break any statistical associations between, for example, a given attribute of the individual (*e.g.*, their sex) and their social ties (*e.g.*, their weighted out-strength). The QAP randomisation procedure aims to preserve dyadic dependence structure by randomly shuffling rows and columns together, but may include custom constraints—such that only a set of data points (or ties/edges) are permuted—in an attempt to represent certain social dependencies (*e.g.,* Chodrow, 2020; Fosdick et al., 2018).

Multiple regression quadratic assignment procedure (MRQAP) is an extension of QAP. MRQAP is adapted for linear regression models being applied to network data; it aims to estimate the statistical significance of regression coefficients by using permutations that attempt to preserve some of the dyadic structure of social networks. The most common MRQAP approach in ASNA is "Double-Semi-Partialing", which is a variation of a permutation test that helps to account for overlap between a main variable of interest, X, and a confounding variable, Z (Dekker et al., 2007). This procedure first removes the shared influence of Z from X, then entails permuting the residuals of X to test whether the adjusted X continues to explain an outcome variable, Y, all while keeping the effects of Z constant.

The second approach, "data-stream permutations" or "pre-network data permutations" (Bejder et al., 1998) has been developed and applied to animal network data generated by several different sampling procedures—such as "gambit of the group" data (Franks et al., 2010), focal follows (Altmann, 1974), and GPS data (He et al., 2023)—to correct for the unique dependencies that these sampling procedures create. This approach involves the random shuffling of rows or columns, but of a non-aggregated matrix that represents observed ties *at a given point of data collection*. These permutations are also constrained in a way that attempts to reflect the non-social structuring features that they aim to control for. For example, for "gambit of the group" data, individuals that are observed in a specific group are randomly reassigned to another group (*e.g.*, if individual i was originally observed in group 1, they would be re-assigned to appear in group 2). By doing this, both permutation approaches are used for hypothesis testing and attempt to produce null distributions for a given test statistic (*e.g.*, a t-value and an associated p-value) where the "null hypothesis/model" is true, all while aiming to preserve some of the important features of the network (*e.g.*, the total number of observed ties, or the number of times individuals were observed: Farine & Whitehead, 2015).

While some still consider permutation approaches to be flexible tools that account for all of the measurement and sampling issues endemic to animal social network data (Farine, 2017), they have fallen out of favour in many other fields of network analysis. This move away from permutation tests was ushered in by several fundamental flaws to the methodology, particularly where well-specified parametric models have proven more effective (Sosa et al., 2025).

NETWORK PERMUTATION METHODS CANNOT CONTROL FOR SOCIAL DEPENDENCIES

As previously noted, dependencies in social networks are often products of the social processes that are the targets of inference and should rarely be considered a nuisance that needs to be controlled away (Snijders, 2011). However, when dependencies are considered problems that need to be eliminated in order to make inferences about a phenomenon of interest, evidence suggests that some standard network permutation methods fail to appropriately correct for the biases that such dependencies create. Research highlighting the pitfalls of permutation methods for networks (in)famously brought about longstanding and heated debate among community ecologists in the 1970s (Connor & Simberloff, 1979; Diamond, 1975; Diamond & Gilpin, 1982). This led to seminal figures within the field referring to permutation methods as being "*characterized by hidden structure, inefficiency, lack of common-sense and statistical weakness*" (Gilpin & Diamond, 2014, p. 313), and stating that reading (even their own) permutation methods gave the impression of "*reading a book of magic spells*" (de Bello, 2012, p. 313). A growing body of work in community ecology has highlighted that permutation approaches do not do what they are believed to do and that non-permutation approaches exhibit superior performance in most settings (*e.g.*, Gotelli & Ulrich, 2010; Harris, 2016; Veech, 2013). Similar conclusions were drawn in human SNA in the early 2000's (Anderson & Robinson, 2001), which led to the widespread proliferation of alternative, more effective statistical approaches for analysing social network data (Back & Kenny, 2010).

A growing body of complementary evidence is emerging in ASNA. Several simulation studies have shown that network permutations (Hart et al., 2022) and data stream permutations (Weiss et al., 2021a) produce biased inferences due to the social and non-social dependence structures that they were developed and applied to "control for" (Puga-Gonzalez et al., 2021). Several issues associated with permutation methods are also caused by misspecification of the constraints of the randomisation procedure being applied, which results in permutations that do not appropriately capture the intended dependence structure (Hobson et al., 2021). Given the computational limitations of permutation methods, and the inability to specify converging models that incorporate the many types of constraints as part of the randomisation procedure that are necessary to preserve network structure, applications of common permutation methods may to produce misleading results and incorrect inferences (*e.g.,* as the null model may assume that there is no effect of X on Y and/or that such network structure or sampling feature is random; Hobson et al., 2021; Weiss et al., 2021a).

More fundamentally, common permutation methods in ASNA typically assume that any randomisation of any given data is equally likely (*i.e.,* that they are exchangeable; Good, 2013). However, in the presence of any dependence structure in a given dataset that is not included as a constraint in a randomisation procedure, the exchangeability assumptions of such permutation methods are broken (Winkler et al., 2015). This issue has been clearly demonstrated for many MRQAP methods: simulation studies show that many MRQAP approaches (particularly permutations of raw data) yield either inflated or overly conservative Type I error rates (*i.e.,* false positive significant p-values) under conditions of collinearity and network autocorrelation (Dekker et al., 2007). Across most conditions within the simulation studies, however, Double-Semi-Partialing was the most robust MRQAP approach. While the Double-Semi-Partialing MRQAP procedure is approximately valid under the assumption that the aspects of the data that cannot be explained by a statistical model (*i.e.*, the residuals) should behave the same way no matter how the data are randomly shuffled. If this assumption holds, the method can yield reliable and meaningful tests of statistical significance (*i.e.,* p-values). If, however, these residuals have patterns or structure that change when randomised (*e.g.*, in the presence of higher order features of a network, such as transitivity), or if the test is applied to highly skewed or non-normal data distributions—which are common in social interaction data (*e.g.,* where most dyads have no, or very few, connections, but a few have very strong ties)—such permutation approaches are not robust (Dekker et al., 2007).

Misunderstanding 4: "Generative network models are not appropriate for analysing animal social network data"

In this section, we outline several misunderstandings about the use of generative network modelling frameworks in animal social networks analysis. These misconceptions are grouped into three main issues and are elaborated on in reference to modelling strategies that some erroneously believe offer more robust alternatives, namely permutation methods. In ASNA, analysis approaches are often lumped into either being network "permutation approaches" or "regression-based approaches," with generative network models typically being referred to as being in the regression category. An important point for clarification is that this dichotomy is in itself a misunderstanding. Generative network models need not be typical regression models—and many are in fact not (*e.g.*, Robins et al., 2007b; Snijders, 2017)—and many permutation methods (*e.g.,* MRQAP) use regression models as part of their underlying algorithms. Here, we further highlight the pitfalls of contemporary permutation approaches for ASNA and clarify the utility of generative network models as viable solutions.

GENERATIVE NETWORK MODELS CAN ADJUST FOR BIOLOGICAL CONFOUNDS

A major goal of statistical analysis is to isolate the effect of a given variable, X, on another, dependent variable, Y. Many researchers (including ourselves) have long avoided stating whether a change in X directly *causes* a change in Y—instead using terms such as "is associated with" or "is related to" when reporting results (Grosz et al., 2020). In contrast, the analytical approach employed by most researchers, and the ensuing discussion of findings, often attempt to do just that (Rohrer, 2018).

In behavioural ecology, researchers often ask questions about how a trait may impact an aspect of sociality, typically operationalised as a centrality metric computed on a network. Researchers know a lot about their study systems, devoting months or years to conducting fieldwork, and will suspect that there are other factors (*e.g.,* the sex of an individual) that structure sociality, or confound the effect of interest and must be controlled for. In an attempt to deal with these perceived confounds, researchers may therefore try to incorporate this domain knowledge by including any additional observed variables in their statistical models—be they simple linear models with a centrality metric as the dependent variable, or better yet, dyadic models (*e.g.*, Ross et al., 2023; Van Duijn et al., 2004) or models of network structure (*e.g.*, Caimo & Friel, 2011; Robins et al., 2007a; Snijders et al., 2010). Such an approach has received criticism from some behavioural ecologists in ASNA, as it has been suggested that these generative network models assume that networks are free of *non-social* drivers of structure (Farine, 2024). This assertion is, fortunately, incorrect. Generative network models have been developed to not only model social dependencies, but also to flexibly incorporate the effects of non-social variables—*e.g.*, sampling bias and censoring processes (Sosa et al., 2025), or even purposeful misreporting (Redhead et al., 2023b)—that impact apparent social network structure.

While there has been a general acknowledgment of the potential value of generative network models in ASNA, skepticism is still perpetuated. For example, (Farine, 2024, p. 2) states that "*in theory, these [models] can be extended to also include confounding factors, [but] doing so would need to be specific to each study*". However, this point is not a valid critique of the generative network modelling approach. There is no one-size-fits-all statistical analysis that is free from making assumptions that are specific to the study, research question, and data at hand. Science depends on scientists' domain expertise, and statistical models similarly require this knowledge to be carefully integrated into model design to produce meaningful inferences from data.

Confusion persists in ASNA about how to integrate domain expertise into an analysis to deal with non-social confounds. This confusion reflects a broader misconception that has been expressed: that generative network models *assume away* confounding factors while network permutation methods inherently *avoid* them (Farine, 2024). A source of this confusion may perhaps be a mistaken belief that permutation methods resemble randomized controlled trials (RCTs), and therefore provide robust inference "out of the box."

RCTs remove confounding by breaking effects from confounders to the treatment via random assignment of the treatment, with the aim of ensuring that no confounding paths remain (Pearl et al., 2016). Note that this aim of controlling for confounding is complicated even in RCTs, as when a confounder is not independent from the treatment, the treatment effect can only then be considered a 'marginal' effect (over the confounder), which may or may not be what the researcher is aiming for (see Greene, 2003, for further outline and explanation). Although constrained permutations are often described as "controlling for confounding," (Farine, 2024) they only do so by attempting to preserve confounding in both the null distribution and the estimator. Instead of removing confounding from the estimator, the null distribution permutations are constrained to also be confounded. This means that confounding is present in effect sizes and may invalidate inferences. Additionally, this means that permutation regimes must be constrained in a way that best reflects the target confounds, which can force unnecessary stratification, inefficiency, and complexity. Isolating causal effects of interest is essential for addressing ecological questions and producing statistical estimates with biological meaning, and in this respect, permutation methods may fail where generative network models succeed (Franks et al., 2021; Sosa et al., 2025).

Generative network models provide a framework for behavioural ecologists to incorporate their hard-earned knowledge of their system, and adjust for important confounds—be they biological confounds, or aspects of the observation process that affect the recovered data distribution. This is not necessarily done through any modelling extension specific to this framework—in many cases, it is simply the broad notion of "conditioning on" (which is the same as "stratifying by" or "controlling for") a variable in a regression. What is perhaps unique to many generative network models is that the parameters that need to be estimated to control for certain biological factors operate at inherently different levels of a network: they may capture effects at the level of the individual (*e.g.,* X causes higher centrality), the dyadic level (*e.g.,* similarity in X causes individuals to interact reciprocally), or at a higher level of the network (*e.g.,* complementarity in X causes individuals to create transitive groups). In practical terms, coding models with these parameters from scratch can be quite complicated, but there are several open-source software packages that provide a platform for specifying a plethora of models containing such parameters (*e.g.,* Hart et al., 2023; Hunter et al., 2008; Ripley et al., 2025; Ross et al., 2023; Snijders, 2017). While these advances open fruitful pathways for drawing inferences from well-specified models, a great deal of care must be taken when incorporating covariates into these models. As we have outlined earlier, this requires that researchers define the quantity that they wish to estimate—*i.e.*, their estimand—and that they flesh out assumptions describing the causal process that generated their social network data (for an introduction, see Franks et al., 2025; Kawam et al., 2024; McElreath, 2020).

GENERATIVE NETWORK MODELS HAVE APPROPRIATE NULL MODELS

As outlined in the sections above, common network permutation approaches in ASNA are unlikely to appropriately deal with biological confounds. As a result, they are likely to not produce the appropriate null models (which we will refer to as "null distributions") that they are typically applied to generate. By null distributions, we refer to an expected distribution that accurately represents the hypothesis being tested, especially the absence of the effect or relationship of interest.

In frequentist hypothesis testing regimes, null models enable researchers to assess whether the observed network pattern is unlikely to have arisen under the null distribution through computation of a given test statistic and associated p-value. Misunderstandings remain about whether generative network models are able to produce appropriate null models. In a now seminal paper on ASNA, it is claimed that such "*test[s] could only determine whether the network metric differs significantly from zero—a highly unsatisfactory null hypothesis*" (Farine, 2017, p. 1310). Frequentist models can, however, have appropriate null hypotheses for social network data.

Generative network models enable researchers to model a network's null distribution by explicitly defining structural and biological mechanisms that are both theoretically sound and practically applicable. Through this approach, behavioural ecologists can delineate the key structural features of a network (*e.g.,* transitive group formation), and how such features in the data are generated when constructing a null distribution. To do this, we can draw from a long history of theoretical and methodological developments on generative network models—these include dyadic models, such as the stochastic block model and its extensions (Holland et al., 1983; Peixoto, 2019), social relations models (Pillinger et al., 2024; Snijders & Kenny, 1999), latent space models (Hoff et al., 2002), and more complex networks models, such as exponential random graph models and stochastic actor-oriented models (Robins et al., 2007a; Steglich et al., 2006). These generative models can specify null distributions that best reflect how the data are structured and generated (See Hobson et al., 2021, for a guide on doing this in behavioural ecology). This generally highlights that generative network models produce null distributions that are not only appropriate but also meaningful.

In Bayesian statistics, analysis regimes directly model the full probability distribution of parameters given the observed data. This may bypass the need for the construction of a null distribution for statistical testing, and instead produces a posterior distribution over possible parameter values. By combining Bayesian models with causal models, researchers can then make probabilistic statements about the size and direction of effects, such as estimating the probability that a phenotype of interest increases the number of social ties that an individual has by a given amount, and quantifying the uncertainty around that estimate (Gelman et al., 1995). Moreover, the flexibility of Bayesian hierarchical models also makes them well-suited to directly estimate structural dependencies inherent in network data (McElreath, 2020).

As generative network models, be they frequentist or Bayesian, focus on modelling the data-generating process and integrating domain expertise. By doing this, they can offer rich and often easily interpretable results—as opposed to relying solely on the interpretation of p-values, which network permutation approaches necessitate—for making scientific inferences about animal social networks.

You cannot estimate separate regression models for different network metrics

A critical limitation of permutation-based approaches in ASNA arises when network metrics are treated as dependent variables. The analytical pipeline used for examining networks as dependent variables typically entails separate, sequential estimation procedures for each metric (*e.g.*, Kaigaishi & Yamamoto, 2024; McMahon et al., 2024; Roatti et al., 2023). For example, to evaluate how hierarchical rank (*i.e.,* dominance) influences sociality, researchers might compute multiple individual-level metrics (*e.g.,* eigenvector centrality, in-strength) and fit a separate regression model for each. This leads to a set of isolated statistical tests, each examining a single metric with the assumption that they are statistically independent. This practice ignores the fact that many features of a network are jointly determined by the same underlying processes, and that these network metrics are thus statistically dependent (Boccaletti et al., 2006; Feld, 1991; Newman, 2018). The covariance structure among these metrics reflects latent structural regularities that are lost when metrics are analysed separately. As a result, univariate regression models yield incomplete and potentially misleading inferences.

Generative network models aim at representing a network as a whole, which implies (as an implicit consequence) models for any network metric and, thus, bypasses the need for modelling network metrics through separate regression models. More specifically, generative network models can, and must, account for the dependence structure *between* network features (*e.g.*, an individual's out-strength is associated with their tendency to reciprocate ties), as well as any other important factors relating to measurement and sampling regimes that may bleed into the generation of animal network data. By integrating specialist knowledge about the system of study and careful consideration of how the measurement or sampling design may fundamentally bias estimates, researchers can construct multivariate or hierarchical models that jointly model the multiple network features that may best characterise the structure of their network of interest. This further captures the shared variance between such features, and enables more accurate inference about the biological processes that produce network structure (Kawam et al., 2024; Redhead et al., 2023b; Sosa et al., 2025). This integrated approach cannot be approximated by permutation workflows, which lack the functionality to model multivariate dependence or latent social structure.

Misunderstanding 5: "Interaction- and association-based networks are comparable"

Studying social network structure depends fundamentally on how social ties are measured and collected. Different data collection methods require different approaches to analysis, and a key distinction exists between interaction-based and association-based data. Interaction data are based on direct, observable behaviours such as grooming, aggression, mating, or food sharing (Altmann, 1974). These interactions typically occur between pairs of individuals and are the level of social structure from which social relationships or "ties" emerge (Hinde, 1976). In contrast, association data are usually gathered from patterns of spatial and temporal co-occurrence—sometimes by observers (*e.g.*, "gambit-of the-group") but also often through methods such as GPS tracking (He et al., 2023), or RFID tagging (Sabol et al., 2018). Such spatio-temporal co-occurrence (Debetencourt et al., 2024; Franks et al., 2010; Ginsberg & Young, 1992; Whitehead, 2008) may also reflect emergent social relationships. However, the link between co-occurrence and social relationships is often weaker or noisier than for directly observed interactions, due to the ambiguity surrounding the intended target of association (*e.g.,* are animals a and b together because of mutual attraction to each other or because they both want to be with animal c?) and the valence of the association (*e.g.,* are animals together to be agonistic? affiliative? sexual?).

While both association and interaction data can provide useful information about animal social networks, they entail different data generation processes, and thus, different modelling strategies must be devised to study their structure. Traditional social network models (*i.e.,* those that look at specific ties between pairs of individuals) are naturally applied to interactions between pairs of individuals because these data are inherently dyadic. These models are designed to infer the structure of relationships, together with their causes (Butts, 2008; Kawam et al., 2024; Stadtfeld & Block, 2017). However, applying the same models to association data can be problematic, as association events often involve multiple individuals at any given time and are not truly pairwise. Converting these group events into pairwise links to answer questions pertaining to dyadic social relationships may overestimate the number of social ties that individuals have, and inflate network measures like degree distributions and centrality (Danaher-Garcia et

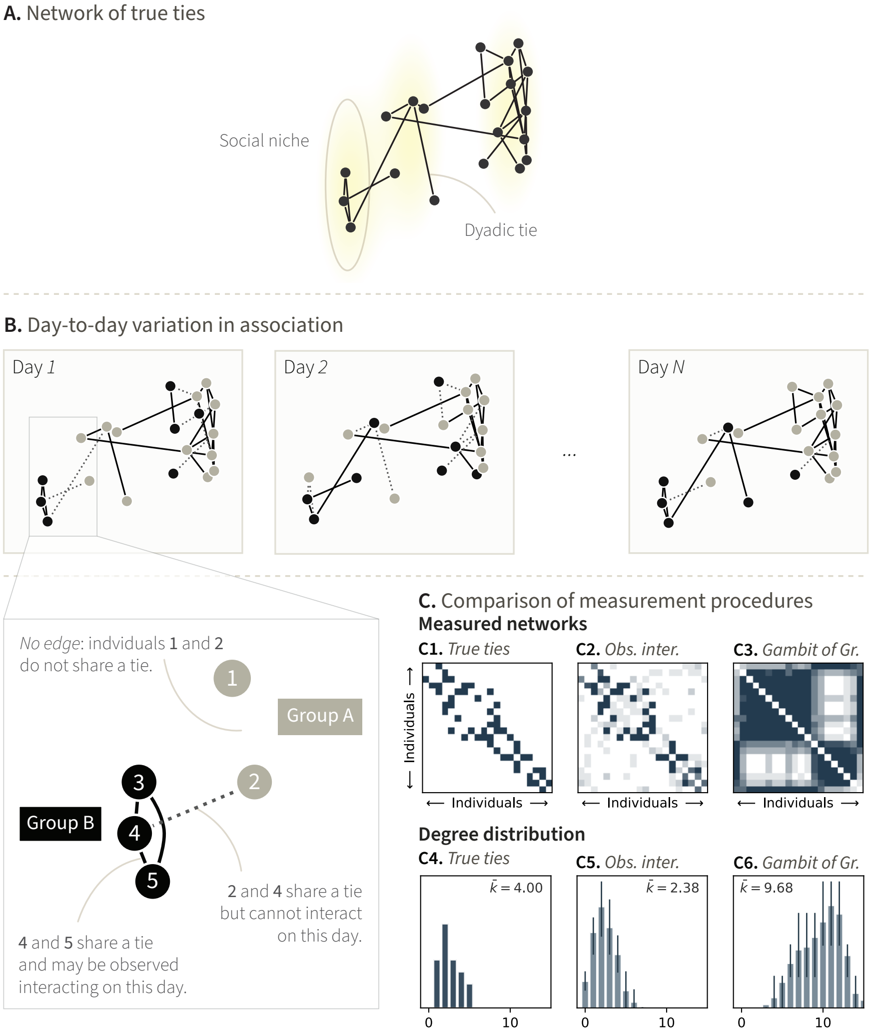


Figure 3 – *Mismatches between data generation and measurement processes can invalidate scientific conclusions.* To demonstrate how conflating interactions and associations can alter research outcomes, we developed a simulation with three components: (i) an interaction network generation process that determines the *potential* interactions we might measure, (ii) a group formation process that dictates which configurations of animals can be observed on any given day, and (iii) an imperfect measurement process that generates interaction/association data based on potential interactions and group memberships. While each component could be implemented in various ways, we made specific choices for this illustrative example. **A**. Potential interactions emerge from a latent space network model. Animals occupy random positions in a social trait space, with proximity between animals indicating stronger relational tendencies. For this example, we created three distinct social niches by sampling centroids μ_r for each niche $r = 1, 2, 3$ from a Gaussian distribution in $d = 2$ dimensions. We randomly assigned animals to niches and sampled each animal's trait space location from a Gaussian distribution centered at its niche's μ_r—the distribution is shown in yellow. We then generated potential interactions probabilistically, with interaction likelihood increasing as the distance between animals in trait space decreases.[3] **B**. To obtain associations, we randomly grouped animals while making sure that interacting pairs of animals would likely appear together. The panel displays three random instances of this process (node color corresponds to groups), created by first positioning animals in *physical space*, in close proximity to their interaction partners[4], and then forming groups of animals based on physical proximity. [5] **C**. Finally, we simulated the measurement process with association and interaction methods across seven days of data collection. To generate association measurements, we assumed that pairs of animals found in the same group were likely connected. This produced a matrix of connection probabilities $\mathbf{W}_A$, where each w_{ij} represents the proportion of days animals i and j were observed together. To generate interaction measurements, we translated the network of potential interactions into connection probabilities $\mathbf{W}_I$ and added two processes that capture experimental conditions. First, we only recorded ties between animals that were associating during measurement (the solid edges in panel B.). Second, we allowed for false positive and false negative measurements (each occurring at a rate of 0.05 per pair of individuals per day). As the results demonstrate, association-based methods misrepresent the interaction network. This misunderstanding can propagate to network characteristics. Panels **C4**-**C6** show the average number of animals having a certain degree (bars), and a 50% interval for this quantity (error bars), with mean degree $\bar{k}$ in the inset. The error bars were obtained by generating $n = 100$ networks compatible with the measurements.[6] Association data completely mischaracterise the degree distribution of the interaction network.

al., 2022; Weiss et al., 2021b). We illustrate this phenomenon in Figure 3.

An analytical approach that can reflect and preserve the complexity of association data is to use hypergraph models (Battiston et al., 2020) and hyper-event models for longitudinal analyses (Lerner & Lomi, 2023). Unlike standard network models, hypergraphs allow researchers to represent group-level events directly, by linking multiple individuals together in a single event without breaking it down into dyadic ties (Chodrow & Mellor, 2020; Contisciani et al., 2022; Torres et al., 2021). This means that researchers may analyse association data (or indeed interaction data for interactions that involve multiple partners) in a way that respects its original structure, and further incorporate their domain knowledge by specifying factors that they believe to be important for generating such structure. For instance, hypergraph models are especially useful for exploring questions about group dynamics, such as whether some individuals consistently appear together in the same groups or similar physical spaces (Taramasco et al., 2010), or how group size influences the spread of information or disease (Iacopini et al., 2019; Silk et al., 2022). By maintaining the group-based nature of association data, behavioural ecologists can produce more accurate and biologically meaningful insights with hypergraph modelling approaches than by applying traditional network models.

Ultimately, both interaction- and association-based data offer valuable insights into animal social behaviour, but they are not interchangeable. This is not to argue that one type of network data is better than another. Rather, the type of network data being collected should be decided upon in direct relation to its utility in answering a motivating research question, including whether that question requires the valence of network edges or the certainty of partner choice to be known, while being cognisant of the logistical barriers and high temporal and/or monetary costs associated with collecting animal social network data. The key to directly and accurately answering research questions in ASNA is to carefully match the analytical approach to the way the data were collected. This alignment is crucial for examining theories of social behaviour and drawing valid conclusions from empirical data. By developing and applying statistical models that best reflect the process through which data have been generated, behavioural ecologists can make the most of the diversity of available data and avoid the pitfalls associated with mismatched, traditional approaches.

Concluding remarks

Animal social network analysis has become a central pillar of behavioural ecology, offering theory and tools for understanding the complexity of animal societies and their ecological and evolutionary causes and consequences. In this paper, we have outlined five common misunderstandings that have hindered progress in the field. These misunderstandings have collectively contributed to a disconnect between ASNA and developments in the broader field of SNA.

Encouragingly, this is a moment for methodological transformation within the field. By revisiting foundational principles, such as defining clear estimands, modelling empirically plausible and causally explicit data-generating processes, and embracing model-based statistical approaches that derive from generative assumptions, behavioural ecologists are increasingly well-positioned to overcome these challenges. The application of methodological innovations made

---

[3] We employed an exponentially decreasing connection probability, $p_{ij} = e^{-\beta\Delta_{ij}}$, where Δ_{ij} represents the distance between animals i and j in trait space.

[4] We implemented a repulsion-attraction model where potential interactions "pull" animals together while disconnected pairs are pushed apart, using the Fruchterman-Reingold Algorithm (Fruchterman & Reingold, 1991).

[5] We applied an Affinity Propagation algorithm (Frey & Dueck, 2007) to cluster animals based on distance in physical space.

[6] Edge (i, j) was included in each posterior sample with probability w_{ij}, where w_{ij} is derived from the measurement process and treated as independent of others once conditioned on w_{ij} (Young et al., 2020).

in the wider field of SNA makes this possible, as the combination of generative network models and causal inference frameworks offers powerful alternatives to traditional permutation-based approaches. These tools not only provide greater inferential clarity but also allow researchers to more directly integrate their expert knowledge about their system into their analytical procedures.

The use of generative network models provides a platform for future research to deepen the connection between the predictions made by formal theoretical models in behavioural ecology and evolutionary biology, and the statistical estimates from empirical studies that aim to test such predictions. Many classical models, from the formation of dominance hierarchies (Chase et al., 2002) and theory on costly signalling (Gintis et al., 2001) to models of partner choice (Fu et al., 2008) and reciprocity (Efferson et al., 2024), make explicit predictions about social dynamics that can be directly translated into statistical models of network structure. By grounding empirical analyses in such theoretical frameworks, researchers can conduct sharper examinations of formal theory and further contribute to the refinement of theory itself. In addition, the increasing availability of high-resolution, long-term interaction and association data opens the door to dynamic network analyses that capture how social structure emerges over time. Rather than treating social networks as static snapshots, researchers now have the opportunity to model the processes that bring about relationship formation, maintenance, and decay (Snijders, 2009). This provides a fruitful avenue for future research to develop and better integrate dynamic network models for understanding the mechanisms underlying animal sociality.

In sum, embracing flexible generative network modelling approaches and aligning analytical strategies with those from the rest of SNA will bring ASNA and the broader interdisciplinary field into closer conversation, enabling more precise, biologically meaningful insights into the structure and function of animal societies.

Acknowledgements

The authors would like to thank Prof. Tom A. B. Snijders for thorough and extremely helpful feedback on a previous version of this manuscript, as well as Alice Hill for her assistance in designing Figure 1. Lauren Brent acknowledges support from a Consolidator Grant from the European Research Council (864461 – FriendOrigins) and grants from the National Institutes of Health (R01AG087902, R01AG084706).

References

Allee, W. C. (1931). Co-operation among animals. *American Journal of Sociology*, *37*(3), 386–398.

Altmann, J. (1974). Observational study of behavior: Sampling methods. *Behaviour*, *49*(3-4), 227–266.

Anderson, C. J., Wasserman, S., & Crouch, B. (1999). A p* primer: Logit models for social networks. *Social networks*, *21*(1), 37–66.

Anderson, M. J., & Robinson, J. (2001). Permutation tests for linear models. *Australian & New Zealand Journal of Statistics*, *43*(1), 75–88.

Back, M. D., & Kenny, D. A. (2010). The social relations model: How to understand dyadic processes. *Social and Personality Psychology Compass*, *4*(10), 855–870.

Bassett, D. S., & Sporns, O. (2017). Network neuroscience. *Nature Neuroscience*, *20*(3), 353–364.

Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., Young, J.-G., & Petri, G. (2020). Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, *874*, 1–92.

Bejder, L., Fletcher, D., & Bräger, S. (1998). A method for testing association patterns of social animals. *Animal Behaviour*, *56*(3), 719–725.

Blitzstein, J. K., & Hwang, J. (2019). *Introduction to probability*. Crc Press.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., & Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, *424*(4-5), 175–308.

Brent, L. J., Lehmann, J., & Ramos-Fernández, G. (2011). Social network analysis in the study of nonhuman primates: A historical perspective. *American Journal of Primatology*, *73*(8), 720–730.

Butts, C. T. (2008). A relational event framework for social action. *Sociological methodology*, *38*(1), 155–200.

Caimo, A., & Friel, N. (2011). Bayesian inference for exponential random graph models. *Social Networks*, *33*(1), 41–55.

Carpenter, C. R. (1945). Concepts and problems of primate sociometry. *Sociometry*, *8*(1), 56–61.

Chase, I. D., Tovey, C., Spangler-Martin, D., & Manfredonia, M. (2002). Individual differences versus social dynamics in the formation of animal dominance hierarchies. *Proceedings of the National Academy of Sciences*, *99*(8), 5744–5749.

Chepko-Sade, B. D., Reitz, K. P., & Sade, D. S. (1989). Sociometrics of Macaca mulatta IV: Network analysis of social structure of a pre-fission group. *Social Networks*, *11*(3), 293–314.

Chodrow, P., & Mellor, A. (2020). Annotated hypergraphs: Models and applications. *Applied Network Science*, *5*(1), 9.

Chodrow, P. S. (2020). Configuration models of random hypergraphs. *Journal of Complex Networks*, *8*(3), cnaa018.

Clutton-Brock, T. (2016). *Mammal societies*. John Wiley & Sons.

Connor, E. F., & Simberloff, D. (1979). The assembly of species communities: Chance or competition? *Ecology*, *60*(6), 1132–1140.

Contisciani, M., Battiston, F., & De Bacco, C. (2022). Inference of hyperedges and overlapping communities in hypergraphs. *Nature Communications*, *13*(1), 7229.

Danaher-Garcia, N., Connor, R., Fay, G., Melillo-Sweeting, K., & Dudzinski, K. M. (2022). Using social network analysis to confirm the "gambit of the group" hypothesis for a small cetacean. *Behavioural Processes*, *200*, 104694.

De Bacco, C., Contisciani, M., Cardoso-Silva, J., Safdari, H., Lima Borges, G., Baptista, D., Sweet, T., Young, J.-G., Koster, J., Ross, C. T., et al. (2023). Latent network models to account for noisy, multiply reported social network data. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *186*(3), 355–375.

de Bello, F. (2012). The quest for trait convergence and divergence in community assembly: Are null-models the magic wand? *Global Ecology and Biogeography*, *21*(3), 312–317.

Debetencourt, B., Barry, M. M., Arandjelovic, M., Stephens, C., Maldonado, N., & Boesch, C. (2024). Camera traps unveil demography, social structure, and home range of six unhabituated western chimpanzee groups in the moyen bafing national park, guinea. *American Journal of Primatology*, *86*(2), e23578.

Dekker, D., Krackhardt, D., & Snijders, T. A. B. (2007). Sensitivity of MRQAP tests to collinearity and autocorrelation conditions. *Psychometrika*, *72*, 563–581.

Diamond, J. M. (1975). The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological conservation*, *7*(2), 129–146.

Diamond, J. M., & Gilpin, M. E. (1982). Examination of the "null" model of Connor and Simberloff for species co-occurrences on islands. *Oecologia*, *52*, 64–74.

Dunne, J. A., Williams, R. J., & Martinez, N. D. (2002). Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences*, *99*(20), 12917–12922.

Earl, A. D., Carter, G. G., Berlinger, A. G., Korir, E., Shah, S. S., Watetu, W. N., & Rubenstein, D. R. (2025). A cryptic role for reciprocal helping in a cooperatively breeding bird. *Nature*.

Efferson, C., Bernhard, H., Fischbacher, U., & Fehr, E. (2024). Super-additive cooperation. *Nature*, *626*(8001), 1034–1041.

Epstein, J. M. (2012). *Generative social science: Studies in agent-based computational modeling*. Princeton University Press.

Falkenberg, M., Zollo, F., Quattrociocchi, W., Pfeffer, J., & Baronchelli, A. (2024). Patterns of partisan toxicity and engagement reveal the common structure of online political communication across countries. *Nature Communications*, *15*(1), 9560.

Farine, D. R. (2017). A guide to null models for animal social network analysis. *Methods in Ecology and Evolution*, *8*(10), 1309–1320.

Farine, D. R. (2024). Modelling animal social networks: New solutions and future directions. *Journal of Animal Ecology*, *93*(3), 250–253.

Farine, D. R., & Whitehead, H. (2015). Constructing, conducting and interpreting animal social network analysis. *Journal of Animal Ecology*, *84*(5), 1144–1163.

Feld, S. L. (1991). Why your friends have more friends than you do. *American Journal of Sociology*, *96*(6), 1464–1477.

Formica, V. A., Wood, C., Larsen, W., Butterfield, R., Augat, M., Hougen, H., & Brodie Iii, E. (2012). Fitness consequences of social network position in a wild population of forked fungus beetles (bolitotherus cornutus). *Journal of Evolutionary Biology*, *25*(1), 130–137.

Fosdick, B. K., Larremore, D. B., Nishimura, J., & Ugander, J. (2018). Configuring random graph models with fixed degree sequences. *SIAM Review*, *60*(2), 315–355.

Franks, D. W., Ruxton, G. D., & James, R. (2010). Sampling animal association networks with the gambit of the group. *Behavioral Ecology and Sociobiology*, *64*, 493–503.

Franks, D. W., Ruxton, G. D., & Sherratt, T. (2025). Ecology needs a causal overhaul. *Biological reviews of the Cambridge Philosophical Society*.

Franks, D. W., Weiss, M. N., Silk, M. J., Perryman, R. J., & Croft, D. P. (2021). Calculating effect sizes in animal social network analysis. *Methods in Ecology and Evolution*, *12*(1), 33–41.

Frey, B. J., & Dueck, D. (2007). Clustering by passing messages between data points. *Science*, *315*(5814), 972–976.

Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and experience*, *21*(11), 1129–1164.

Fu, F., Hauert, C., Nowak, M. A., & Wang, L. (2008). Reputation-based partner choice promotes cooperation in social networks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, *78*(2), 026117.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (1995). *Bayesian data analysis*. Chapman; Hall/CRC.

Gilpin, M. E., & Diamond, J. M. (2014). 17. are species co-occurrences on islands non-random, and are null hypotheses useful in community ecology? In *Ecological communities* (pp. 297–315). Princeton University Press.

Ginsberg, J. R., & Young, T. P. (1992). Measuring association between individuals or groups in behavioural studies. *Animal Behaviour*, *44*(1), 377–379.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology*, *213*(1), 103–119.

Good, P. (2013). *Permutation tests: A practical guide to resampling methods for testing hypotheses*. Springer Science & Business Media.

Gotelli, N. J., & Ulrich, W. (2010). The empirical Bayes approach as a tool to identify nonrandom species associations. *Oecologia*, *162*, 463–477.

Greene, W. H. (2003). Econometric analysis. *Pretence Hall*.

Grosz, M. P., Rohrer, J. M., & Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspectives on Psychological Science*, *15*(5), 1243–1255.

Haraway, D. (2023). When species meet. In *The routledge international handbook of more-than-human studies* (pp. 42–78). Routledge.

Harris, D. J. (2016). Inferring species interactions from co-occurrence data with Markov networks. *Ecology*, *97*(12), 3308–3314.

Hart, J., Weiss, M. N., Brent, L. J., & Franks, D. W. (2022). Common permutation methods in animal social network analysis do not control for non-independence. *Behavioral Ecology and Sociobiology*, *76*(11), 151.

Hart, J., Weiss, M. N., Franks, D., & Brent, L. (2023). Bison: A Bayesian framework for inference of social networks. *Methods in Ecology and Evolution*, *14*(9), 2411–2420.

He, P., Klarevas-Irby, J. A., Papageorgiou, D., Christensen, C., Strauss, E. D., & Farine, D. R. (2023). A guide to sampling design for gps-based studies of animal societies. *Methods in Ecology and Evolution*, *14*(8), 1887–1905.

Hidalgo, C. A. (2016). Disconnected, fragmented, or united? a trans-disciplinary review of network science. *Applied Network Science*, *1*, 1–19.

Hinde, R. A. (1976). Interactions, relationships and social structure. *Man*, 1–17.

Hobson, E. A., Silk, M. J., Fefferman, N. H., Larremore, D. B., Rombach, P., Shai, S., & Pinter-Wollman, N. (2021). A guide to choosing and implementing reference models for social network analysis. *Biological Reviews*, *96*(6), 2716–2734.

Hoff, P. D., Raftery, A. E., & Handcock, M. S. (2002). Latent space approaches to social network analysis. *Journal of the American Statistical Association*, *97*(460), 1090–1098.

Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, *5*(2), 109–137.

Hunter, D. R., Handcock, M. S., Butts, C. T., Goodreau, S. M., & Morris, M. (2008). ergm: A package to fit, simulate and diagnose exponential-family models for networks. *Journal of Statistical Software*, *24*, 1–29.

Iacopini, I., Petri, G., Barrat, A., & Latora, V. (2019). Simplicial models of social contagion. *Nature Communications*, *10*(1), 2485.

Jang, H., Ross, C. T., Boyette, A. H., Janmaat, K. R., Kandza, V., & Redhead, D. (2024). Women's subsistence networks scaffold cultural transmission among BaYaka foragers in the Congo Basin. *Science Advances*, *10*(2), eadj2543.

Kaigaishi, Y., & Yamamoto, S. (2024). Higher eigenvector centrality in grooming network is linked to better inhibitory control task performance but not other cognitive tasks in free-ranging japanese macaques. *Scientific Reports*, *14*(1), 26804.

Kandler, A., & Powell, A. (2018). Generative inference for cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *373*(1743), 20170056.

Kawam, B., Ostner, J., McElreath, R., Schuelke, O., & Redhead, D. (2024). A causal framework for the drivers of animal social network structure. *bioRxiv*, 2024–06.

Kenny, D. A., & Judd, C. M. (1986). Consequences of violating the independence assumption in analysis of variance. *Psychological Bulletin*, *99*(3), 422.

Kilduff, M., & Brass, D. J. (2010). Organizational social network research: Core ideas and key debates. *Academy of Management Annals*, *4*(1), 317–357.

Krackhardt, D. (1987). QAP partialling as a test of spuriousness. *Social Networks*, *9*(2), 171–186.

Krackhardt, D. (1988). Predicting with networks: Nonparametric multiple regression analysis of dyadic data. *Social networks*, *10*(4), 359–381.

Krackhardt, D. (1992). A caveat on the use of the quadratic assignment procedure. *Journal of quantitative anthropology*, *3*(4), 279–296.

Krause, J., Croft, D. P., & James, R. (2007). Social network theory in the behavioural sciences: Potential applications. *Behavioral Ecology and Sociobiology*, *62*, 15–27.

Krause, S., Mattner, L., James, R., Guttridge, T., Corcoran, M. J., Gruber, S. H., & Krause, J. (2009). Social network analysis and valid markov chain monte carlo tests of null models. *Behavioral Ecology and Sociobiology*, *63*, 1089–1096.

Lambert, B. (2018). A student's guide to Bayesian statistics.

Lerner, J., & Lomi, A. (2023). Relational hyperevent models for polyadic interaction networks. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *186*(3), 577–600.

Lundberg, I., Johnson, R., & Stewart, B. M. (2021). What is your estimand? Defining the target quantity connects statistical evidence to theory. *American Sociological Review*, *86*(3), 532–565.

Maryanski, A. R. (1987). African ape social structure: Is there strength in weak ties? *Social Networks*, *9*(3), 191–215.

McElreath, R. (2020). *Statistical rethinking: A Bayesian course with examples in R and Stan* (Second). Chapman; Hall/CRC.

McMahon, K., Marples, N. M., Spurgin, L. G., Rowland, H. M., Sheldon, B. C., & Firth, J. A. (2024). Social network centrality predicts dietary decisions in a wild bird population. *Iscience*, *27*(5).

Moreno, J. L. (1934). Who shall survive? A new approach to the problem of human interrelations.

Newman, M. (2018). *Networks*. Oxford university press.

Pearl, J., Glymour, M., & Jewell, N. P. (2016). *Causal inference in statistics: A primer*. John Wiley & Sons.

Peixoto, T. P. (2019). Bayesian stochastic blockmodeling. *Advances in network clustering and blockmodeling*, 289–332.

Pillinger, R., Steele, F., Leckie, G., & Jenkins, J. (2024). A dynamic social relations model for clustered longitudinal dyadic data with continuous or ordinal responses. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *187*(2), 338–357.

Puga-Gonzalez, I., Sueur, C., & Sosa, S. (2021). Null models for animal social network analysis and data collected via focal sampling: Pre-network or node network permutation? *Methods in Ecology and Evolution*, *12*(1), 22–32.

Rawlings, C. M., Smith, J. A., Moody, J., & McFarland, D. A. (2023). *Network analysis: Integrating social network theory, method, and application with R*. Cambridge University Press.

Redhead, D. (2024). Social structure and the evolutionary ecology of inequality. *Trends in Cognitive Sciences*.

Redhead, D., Dalla Ragione, A., & Ross, C. T. (2023a). Friendship and partner choice in rural colombia. *Evolution and Human Behavior*, *44*(5), 430–441.

Redhead, D., McElreath, R., & Ross, C. T. (2023b). Reliable network inference from unreliable data: A tutorial on latent network modeling using STRAND. *Psychological Methods*.

Ripley, R. M., Snijders, T., Boda, Z., Vörös, A., & Preciado, P. (2025). Manual for rsiena.

Roatti, V., Cowlishaw, G., Huchard, E., & Carter, A. (2023). Social network inheritance and differentiation in wild baboons. *Royal Society Open Science*, *10*(5), 230219.

Robins, G., Pattison, P., Kalish, Y., & Lusher, D. (2007a). An introduction to exponential random graph (p*) models for social networks. *Social networks*, *29*(2), 173–191.

Robins, G., Snijders, T. A. B., Wang, P., Handcock, M., & Pattison, P. (2007b). Recent developments in exponential random graph (p*) models for social networks. *Social networks*, *29*(2), 192–215.

Robinson, W. S. (1955). Ecological correlations and the behavior of individuals. *International Journal of Epidemiology*, *38*(2), 337–341.

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42.

Ross, C. T., McElreath, R., & Redhead, D. (2023). Modelling animal network data in R using STRAND. *Journal of Animal Ecology*.

Sabol, A. C., Solomon, N. G., & Dantzer, B. (2018). How to study socially monogamous behavior in secretive animals? Using social network analyses and automated tracking systems to study the social behavior of prairie voles. *Frontiers in Ecology and Evolution*, *Volume 6 - 2018*. https://doi.org/10.3389/fevo.2018.00178

Sade, D. S. (1972). Sociometrics of Macaca mulatta. *Folia primatologica*, *18*(3-4), 196–223.

Sade, D. S. (1989). Sociometrics of Macaca mulatta III: N-path centrality in grooming networks. *Social Networks*, *11*(3), 273–292.

Samuni, L., & Surbeck, M. (2023). Cooperation across social borders in bonobos. *Science*, *382*(6672), 805–809.

Seyfarth, R. M. (1976). Social relationships among adult female baboons. *Animal Behaviour*, *24*(4), 917–938.

Silk, M. J., Wilber, M. Q., & Fefferman, N. H. (2022). Capturing complex interactions in disease ecology with simplicial sets. *Ecology Letters*, *25*(10), 2217–2231.

Simonds, P. E. (1974). Sex differences in bonnet macaque networks and social structure. *Archives of Sexual Behavior*, *3*(2), 151–166.

Snijders, T. A. B. (2011). Statistical models for social networks. *Annual Review of Sociology*, *37*(1), 131–153.

Snijders, T. A. B. (2017). Stochastic actor-oriented models for network dynamics. *Annual review of statistics and its application*, *4*(1), 343–363.

Snijders, T. A. B., & Kenny, D. A. (1999). The social relations model for family data: A multilevel approach. *Personal Relationships*, *6*(4), 471–486.

Snijders, T. A. B., Van de Bunt, G. G., & Steglich, C. E. (2010). Introduction to stochastic actor-based models for network dynamics. *Social Networks*, *32*(1), 44–60.

Snijders, T. A. (1996). Stochastic actor-oriented models for network change. *Journal of mathematical sociology*, *21*(1-2), 149–172.

Snijders, T. A. (2009). Longitudinal methods of network analysis. *Encyclopedia of complexity and system science*, 5998–6013.

Sosa, S., McElreath, M. B., Perry, S., Redhead, D., & Ross, C. (2025). Robust bayesian analysis of animal networks subject to biases in sampling intensity and censoring. *Methods in Ecology & Evolution*, 2024–06.

Stadtfeld, C., & Block, P. (2017). Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, *4*, 318–352.

Steglich, C. E., Snijders, T. A. B., & West, P. (2006). Applying SIENA. *Methodology*, *2*(1), 48–56. https://doi.org/10.1027/1614-2241.2.1.48

Steglich, C. E., & Snijders, T. A. (2022). Stochastic network modeling as generative social science. In *Handbook of sociological science* (pp. 73–99). Edward Elgar Publishing.

Taramasco, C., Cointet, J.-P., & Roth, C. (2010). Academic team formation as evolving hypergraphs. *Scientometrics*, *85*(3), 721–740.

Testard, C., Shergold, C., Acevedo-Ithier, A., Hart, J., Bernau, A., Negron-Del Valle, J., Phillips, D., Watowich, M., Sanguinetti-Scheck, J., Montague, M., et al. (2024). Ecological disturbance alters the adaptive benefits of social ties. *Science*, *384*(6702), 1330–1335.

Tinbergen, N. (1939). On the analysis of social organization among vertebrates, with special reference to birds. *American Midland Naturalist*, 210–234.

Torres, L., Blevins, A. S., Bassett, D., & Eliassi-Rad, T. (2021). The why, how, and when of representations for complex systems. *SIAM Review*, *63*(3), 435–485.

Van Duijn, M. A., Snijders, T. A. B., & Zijlstra, B. J. (2004). P2: A random effects model with covariates for directed graphs. *Statistica Neerlandica*, *58*(2), 234–254.

Veech, J. A. (2013). A probabilistic model for analysing species co-occurrence. *Global Ecology and Biogeography*, *22*(2), 252–260.

Wasserman, S., & Faust, K. (1994). Social network analysis: Methods and applications.

Weiss, M. N., Franks, D. W., Brent, L. J., Ellis, S., Silk, M. J., & Croft, D. P. (2021a). Common datastream permutations of animal social network data are not appropriate for hypothesis testing using regression models. *Methods in Ecology and Evolution*, *12*(2), 255–265.

Weiss, M. N., Franks, D. W., Giles, D. A., Youngstrom, S., Wasser, S. K., Balcomb, K. C., Ellifrit, D. K., Domenici, P., Cant, M. A., Ellis, S., et al. (2021b). Age and sex influence social interactions, but not associations, within a killer whale pod. *Proceedings of the Royal Society B*, *288*(1953), 20210617.

Wey, T., Blumstein, D. T., Shen, W., & Jordán, F. (2008). Social network analysis of animal behaviour: A promising tool for the study of sociality. *Animal behaviour*, *75*(2), 333–344.

Whitehead, H. (1999). Testing association patterns of social animals. *Animal Behaviour*, *57*(6), F26–F29.

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

Wilson, E. O. (2000). *Sociobiology: The new synthesis*. Harvard University Press.

Winkler, A. M., Webster, M. A., Vidaurre, D., Nichols, T. E., & Smith, S. M. (2015). Multi-level block permutation. *Neuroimage*, *123*, 253–268.

Young, J.-G., Cantwell, G. T., & Newman, M. (2020). Bayesian inference of network structure from unreliable data. *Journal of Complex Networks*, *8*(6), cnaa046.