

Research questions should drive edge definitions in social network studies

Alecia J. Carter^{1,*}, Alexander E. G. Lee^{2,3} & Harry H. Marshall^{4,†}

¹Large Animal Research Group, Department of Zoology, University of Cambridge, Cambridge, UK

²The Institute of Zoology, Zoological Society of London, Regent's Park, London, UK

³Department of Zoology, University of Oxford, Oxford, UK

⁴Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Penryn, UK

* Author for correspondence: Large Animal Research Group, Department of Zoology, University of Cambridge, Downing St, Cambridge, UK, CB2 3EJ T: +44 12237 69277 F: +44 12233 36676 E: ac854@cam.ac.uk

†Authors are listed alphabetically

Keywords: baboon, fission-fusion, group living, social network analysis, social organisation

16 Recently we published a study (Castles et al., 2014) that compared social network metrics that were created
17 from two methods for defining connections (edges) among wild baboon (*Papio ursinus*) individuals (nodes):
18 proximity and interactions. We found that in many (but not all) cases individuals' positions in the proximity
19 networks were not predictive of their positions in the interaction networks and we cautioned researchers about
20 assuming that one is a proxy for the other, which is frequently done in social network studies (e.g. Carter,
21 Macdonald, Thomson, & Goldizen, 2009). In a recent Forum article, Farine (2015) outlines several
22 assumptions that researchers make about how to define edges among individuals that may affect the results
23 of social network studies, before presenting new empirical findings from wild thornbills (genus *Acanthiza*) that
24 he concludes contrast with ours. We are excited that our research has generated such interest, and this new
25 article adds to a growing body of empirical studies that consider sampling issues in social network studies
26 (Castles et al., 2014; Hobson, Avery, & Wright, 2013; Lehmann & Ross, 2011; Madden, Drewe, Pearce, &
27 Clutton-Brock, 2011; see H. Whitehead, 2008 for a comprehensive summary of sampling considerations). We
28 agree that the 'gold standard' in social network studies should be for researchers to incorporate multiple
29 networks using different methods to determine edges into their analyses. However, while Farine usefully
30 highlights assumptions that are important to consider when choosing how to collect and analyse one's
31 network data, several aspects of his article require further consideration before we extend the discussion to
32 broader issues in social network studies.

33 In the first case, Farine (2015) presents empirical data from mixed species flocks of thornbills,
34 collected over a 6-week period, in which there are correlations between individuals' network positions in
35 proximity and interaction networks. Farine states that this pattern was in contrast to our general conclusion,
36 and so suggests that our findings are not generalizable across species and that in some cases proximity can
37 be used as a proxy for interactions. We feel the first assertion is misplaced, and we caution against the
38 second. Our results were in agreement with those of the thornbills in some years for some social network
39 metrics, where we also found correlations between some proximity and interaction methods (see Fig. 3 and
40 supplementary material in Castles et al., 2014). However, the correlation between the two methods was not
41 found in other years. Thus, our results from two study groups over three years suggest that findings from one
42 time period may not be generalized to the same group(s) in a different time period, let alone to other groups of
43 a particular study species. Had we measured the social network in one particular year (or group) and found a
44 correlation between the methods, we may have erroneously concluded that we can use proximity as a proxy
45 for interaction in all future studies. To return to Farine's first assertion, we are not seeking to generalise

46 patterns from our study but rather the principle that consistency between groups/years should not be assumed
47 until it has been demonstrated. Thus, with respect to Farine's second assertion, we would reiterate our
48 conclusion from Castles et al. (2014): because of the dynamic nature of social networks, we recommend that
49 researchers take care when assuming that proximity can be a proxy for interactions. This is distinct from the
50 suggestions that (a) proximity can never be a proxy for interactions and (b) proximity cannot be used to create
51 social networks—generalisations that we do not advocate.

52 In the second case, Farine (2014) explores some methodological considerations that were not
53 addressed in our study. We focussed on one decision a researcher could make at the data collection stage,
54 specifically, the behaviours that could be used to create edges in a social network. Yet, as we mentioned in
55 our study (Castles et al., 2014), there are many considerations after the data collection stage, as highlighted
56 by Farine (2015) and outlined in detail elsewhere (H. Whitehead, 2008). We appreciate that Farine is using
57 our study to illustrate some general points, and agree that had we analysed our data differently (e.g., by using
58 rates, rather than proportions, of dyadic grooming interactions) we may have obtained different results.
59 However, this simply further supports our conclusion that social networks measured (and analysed) using
60 different techniques are not necessarily comparable and care should be taken when generalising research
61 findings. These considerations in data collection and analysis also highlight more general issues of research
62 design which have perhaps been overlooked in the largely descriptive studies of social networks thus far (H.
63 Whitehead, 2008). The definition of an edge connecting nodes in a network should first and foremost depend
64 on the research question, and assumptions about correspondence between networks should be tested. In the
65 case of the former, for example, if the research question relates to the transfer of visual information between
66 individuals in a network, then edges based on shared proximity are likely to be most informative (but see our
67 further considerations below). But if the research question addresses the likelihood of ectoparasitic disease
68 spread between individuals, then instances of physical interaction between individuals may be more
69 appropriate. In the case of the latter, we would encourage descriptive studies to adopt richer analyses that
70 encompass multiple methods of measuring associations, as do others (Lehmann & Ross, 2011; Madden et
71 al., 2011; H. Whitehead, 2008). Furthermore, we would return again to the conclusion of our original study that
72 any researchers using proximity as a proxy for interactions (and we appreciate this is often the only available
73 source of data on dyadic associations) should be wary that proximity does not always equal interaction, and
74 *vice versa*. For example, individuals are able to interact via olfaction, vocalisations, and visual signals when
75 not in close proximity, or may be in proximity but not interacting (we develop this further below).

76 Consequently, the appropriateness of using proximity as a proxy for interactions will depend on the type of
77 interaction identified as meaningful and important for the research question in the context of the biology of
78 study system.

79 The biology of a study species is likely to influence the appropriateness of different edge definitions
80 for answering specific research questions. The definition of an edge should not solely be dictated by what is
81 *possible* for a study species, but what is *appropriate* for it with respect to the study question and species
82 biology; one should not use instances of close proximity to infer grooming when the research question is ‘does
83 social rank influence grooming equality?’, for example, unless this link has been empirically demonstrated
84 (preferably repeatedly) beforehand. Since, for some study systems, building the social network that is most
85 appropriate for a given research question can be prohibited by logistical constraints on data collection, while
86 other methods may be more practical, Farine’s question remains: can proximity networks be a proxy for
87 interaction networks? Before we expand on this in more detail, we would mention again that this question is
88 distinct from the value of proximity measures to describe social structure/organisation—we find proximity
89 measures valuable for both this task and for hypothesis testing in networks (but see Macdonald & Voelkl,
90 2015; Hal Whitehead & Dufault, 1999). As we mention above, we are in agreement with Farine (2015) that the
91 gold standard in network studies requires a multi-network framework. In our original article (Castles et al.,
92 2014), we were largely concerned with issues of comparability between studies which use different methods
93 to define an association, and raised the issue of using proximity as a proxy for interactions in the discussion of
94 our findings. Where we disagree with Farine is in his assertion that proximity edges can sometimes be used to
95 infer interaction edges or *vice versa* without prior testing of this assumption. This does not preclude the use of
96 proximity edges to determine, for example, individuals’ preferred associates (for an example, see Carter et al.,
97 2009).

98 Below, we will consider under which circumstances we might reliably expect a correspondence
99 between proximity and interaction networks in an effort to provide guidelines for researchers wishing to use
100 proximity as a predictive surrogate for interaction (see also Hal Whitehead & Dufault, 1999). This need not be
101 limited to difficult-to-observe species, but could also apply to different methods of collecting data that do not
102 involve direct but remote observation, such as the use of global positioning system collars to assign group
103 membership by some measure of proximity. We also appreciate that understanding how and why different
104 networks may or may not correspond or interrelate is an important research topic in its own right. However,

105 we have not yet imagined any case where one could assume a correspondence between networks without
106 testing for it, though our thought experiment provoked some overlooked considerations in social network
107 studies: (1) some interactions can occur *between* individuals of different subgroups, (2) proximity networks
108 describe only *opportunities* for interaction, and (3) individuals are likely to vary in both their gregariousness—
109 their propensity to be in proximity to others—and their sociability—their propensity to take the opportunity to
110 interact with others when in proximity to them. We will use the baboon system as a worked example of our
111 reasoning by way of explanation where necessary, and we assume for this exercise that the hypothetical
112 proximity network that is putatively predictive of the interaction network is well-sampled and representative of
113 the ‘true’ proximity network.

114 Before we address these points in more detail, we should first take a brief digression to define the
115 term ‘group’ here. To this point, we have used the term to mean a set of behaviourally-connected individuals
116 in which the majority of individuals are connected to most others; this is what H. Whitehead (2008) refers to as
117 a ‘community’ and is the equivalent of a troop in baboons. From here, however, we will use the term to refer to
118 a ‘subgroup’—a subset of a group that is behaviourally connected (either by proximity or interaction) at a
119 particular point in time (Castles et al., 2014)—that is, the level of observation at which social network data are
120 collected. To return to our first consideration then, it is important to address the assumption that researchers
121 make about the proximity needed *for* interactions (Hal Whitehead & Dufault, 1999). As we mention above,
122 individuals are able to interact via olfaction, vocalisations and visual signals when they are not in close
123 proximity, but this is rarely considered as we suspect that it is implied that the interactions are physical. For
124 example, Farine (2015) considered only physical interaction between individuals in his empirical example. In
125 most cases, but not all—consider, for example, olfactory signals provided via latrines or the scent-marking of
126 surfaces—we acknowledge that individuals will have to be within a particular proximity to interact using these
127 other modalities that are of shorter temporal duration. Our point is not that proximity is not important for
128 interaction, but that the range over which visual, auditory and olfactory signals can be transmitted is often
129 beyond the range that is used to define group membership by proximity (and conversely, physical interactions
130 are often well *inside* the range considered for group membership by proximity). This is not a semantic point,
131 but a conceptual one about how we define edges and thus groups by proximity, and how this will limit
132 comparability of networks. To illustrate by an example, baboons can interact via visual signals (using ‘come
133 hither’ faces and lip-smacking) over tens of meters and via vocalisations over hundreds of meters; often these
134 interaction distances are well beyond what we consider as group membership by our proximity rules. As such,

135 individuals can readily and frequently interact *between* groups: conceptually, individuals could have an
136 association index of zero but a non-zero interaction index. Of course, physical interaction requires group co-
137 membership (however spatially defined) and here again the research question should drive the types of
138 interactions that are reasonable to consider; we mean only to highlight an unconsidered assumption that may
139 lead to a mismatch between edge definitions that may lower comparability between networks and studies.

140 Regarding the second consideration, association matrices represent only opportunities for interaction:
141 they describe who *can* interact, but not who *does* interact. While this statement seems obvious, the use of
142 proximity as proxy of interaction is predicated on the implicit assumption that the relationship between
143 proximity strength and interaction rate is probabilistic (and also assumes, as we will do for the rest of this line
144 of argument, that the interaction occurs over a short distance that necessarily places interacting individuals in
145 the same group as defined by proximity; see our point above). This raises a problem with zero edges in the
146 association network. It is logical to assume that individuals who are never in close proximity will never interact:
147 proximity edges valued zero must be coupled with interaction edges valued zero. However, following this
148 logic, the presence of zero-zero proximity and interaction edges will ‘tether’ any linear model that investigates
149 the correlation between these values to the origin (see Fig. 1 in: Farine, 2015); in fact, these models *must*
150 logically pass through the origin. Combined with the impossibility of negative rates of association, the
151 presence of zero-zero values should increase the probability of at least a weakly positive correlation between
152 proximity and interaction edges as soon as there are *any* non-zero interaction edges, and tells us only that
153 individuals interact with those with whom they have an opportunity to interact (and suggests that proximity
154 edges valued zero should be removed for this kind of analysis as they bias the relationship towards the
155 origin). The only logical argument that holds is that individuals that are never in proximity do not interact.
156 However, the assumption that proximity edge weights will provide (detailed) predictive data on differential
157 rates of interactions between those individuals that are connected cannot be made. Consider, for example,
158 Fig. 1 in Farine (2015): none of the of dyads exhibiting an (above average) proximity edge weight of 0.5 were
159 observed interacting over the six-week study. Thus, proximity networks rather show who is connected and
160 who is not, and therefore who can interact (at some unknown rate, which may include 0) and who cannot.

161 We feel that it is at this point that disagreements may arise about the usefulness of proximity as a
162 proxy for interaction, and raises our third consideration. We argue that assumptions regarding the patterns of
163 interactions between connected individuals should not be made, since individuals can vary not just in their

164 gregariousness (the propensity to be in proximity to others), but also their sociability (the propensity to interact
165 with others to whom they are in proximity). Furthermore, these propensities need not be positively correlated,
166 and may be influenced by a range of social factors. This may lead to relationships between proximity and
167 interaction that deviate from a neutral probabilistic model (i.e. increasing probability of interaction with
168 increasing time spent in proximity), and—depending on patterns of within- and between-individual variation in
169 these two traits—may result in the correspondence between proximity and interaction differing for different
170 dyads' edges: specifically, individuals exhibiting similar association edge weights, and so similar
171 gregariousness, may have different interaction edge weights if they differ in their sociability. While this is
172 similar to Farine's (2015) fourth point about calculating rates of interaction while controlling for time in
173 proximity as opposed to calculating the proportion of an individual's interactions directed to other individuals,
174 we mean to highlight here the individual variation that may make proximity edge weights a poor predictor of
175 interaction probability.

176 For example, we consider a hypothetical population (Fig. 1) in which dyads interact on average on
177 half the occasions that they occur in the same group as defined by proximity (we assume that the probability
178 that dyads interact, or $P(\text{interact})$, is $0.5 * P(\text{co-occur})$). The dashed line in the graph, therefore, describes the
179 average relationship between shared proximity and interaction rate for this population. This relationship is
180 likely to differ between species and may not necessarily be linear. In this hypothetical example, we have
181 plotted three dyads—A, B and C—which co-occur with a probability 0.5. Dyad B interacts at the average rate
182 for the population (near 0.5) and sits close to the line. However, dyads A and C interact more and less than
183 expected than the average for the population, respectively, and consequently sit in darker parts of the plot. All
184 three dyads are equally gregarious (to be more accurate, the result of the combination of the individuals'
185 gregariousness in the dyads makes them equally gregarious); however, dyads A and C are more and less
186 sociable than expected for their gregariousness, respectively. If researchers are not interested in this variation
187 but are simply interested in determining those individuals who are likely to interact, then using proximity
188 networks as a proxy for interaction *probability* (which requires individuals to be in close proximity) *might* be
189 reasonable. However, if researchers are interested in this variation then information on who can and cannot
190 interact clearly does not provide detailed insight into social interactions between individuals, since *a priori*
191 assumptions cannot be made about the relationship between time in proximity and interaction rates. In this
192 case we feel that researchers should (in order of decreasing preference): (1) collect and use data on
193 individual interactions; (2) test this assumption in their study system, perhaps on a smaller subset of the

194 network with more intensely collected data, before proceeding with the use of proximity data; and (3) use
195 proximity as a proxy for interaction (probability) with caution, understanding that this assumption may not
196 necessarily hold.

197 Next, we would like to address two other conceptual issues raised by Farine (2015). We will first
198 consider the potential confusion that is introduced in social network analyses by making a distinction between
199 fission-fusion societies and stable social groups. There is an argument that a particular edge definition will be
200 more informative for species of a particular social organisation (Farine, 2015). As we mention above, we
201 made no judgement on the value of proximity and interaction edges as being more or less accurate
202 representations of the 'real' social network in our original paper (Castles et al., 2014). We suggest only that
203 the different methods provide a different aspect of an individual's social environment, both of which we believe
204 are important and both of which should be collected and compared when possible. Furthermore, we are
205 certainly in agreement that species biology should determine the rules used to define edges in networks for a
206 particular method. However, we think it misleading to make assumptions about how *informative* a particular
207 method is for species of particular social organisations for two reasons. First, it is impossible to categorise all
208 species into particular social organisations, let alone categorise unequivocal types of social organisation.
209 Second, there is substantial variation within categories of social organisation such as those suggested by
210 Farine (2015). As this variation is continuous, categorisation is arbitrary and generalisations at the level of
211 social organisation are impractical.

212 Using the category of fission-fusion species as an example, there is variation among species in the
213 extent of fluidity of individuals among groups, prohibiting the assumption that group co-membership is more
214 informative than interaction in all fission-fusion species. Group membership in fission-fusion species can be
215 highly fluid, where individuals in a local population form one community of connected individuals, such as in
216 guppies (*Poecilia reticulata*) (Darren P. Croft, Krause, & James, 2004). It can also be arranged in a
217 segregated community structure, where association between individuals from the same community is
218 common but association between individuals from different communities is rare, such as in chimpanzees (*Pan*
219 *trogodytes*) (Symington, 1990) and eastern grey kangaroos (*Macropus giganteus*) (Best, Seddon, Dwyer, &
220 Goldizen, 2013). It can also be based around multilevel societies, where there are tiers of closely-connected
221 individuals nested within 'higher' levels of clustered lower tiers, such as in African elephants (*Loxodonta*
222 *africana*) (Wittemyer, Douglas-Hamilton, & Getz, 2005) and hamadryas baboons (*Papio hamadryas*)

223 (Kummer, 1984). We note that these descriptions of the fission-fusion social organisations of these species
224 were all made using proximity (group co-occurrence) methods, demonstrating the usefulness of the proximity
225 method for describing differences in social organisation. However, the assumption that group co-membership
226 in chimpanzees is more informative than grooming equality should, returning to our earlier point, depend on
227 the question that the research is trying inform, not on the fact that they have a fission-fusion social
228 organisation. While this particular example may be hyperbolic, we mean only to highlight that *a priori*
229 assumptions about the *meaningfulness* of one method for all species of a particular social organisation is
230 misguided, based in part on the complications associated with categorising species and variation within
231 categories. We would go so far as to argue that valuing one method above another is equally detrimental to
232 social network studies and should be avoided, not least because we as human researchers are unaware of
233 which distances or timings of co-occurrence, and proportions, counts or durations of interactions that we
234 measure are actually meaningful to the species we study. Furthermore, both proximity and interaction
235 measures are likely to be important and informative for particular biological processes, and we would prefer to
236 see researchers moving towards more holistic frameworks in social network studies that use competing
237 networks to test *a priori* hypotheses about the importance of social networks for animals.

238 Finally, three inter-related questions resulting from our consideration of these methodological issues
239 remain to be discussed: what makes a network, how should sample sizes be considered in social network
240 studies, and at which level should data be pooled? These questions relate to Farine's (2015) idea of social
241 scale and are generally beyond the scope of this reply to address in detail (being relevant research questions
242 in their own right in many systems). One small consideration of note, however, relates to our point regarding
243 the importance of research questions in determining edge definitions. We defined community above as a set
244 of behaviourally-connected individuals in which the majority of individuals are connected to most others. In
245 baboons, a community (troop) is easy to define because connections between troops are so rare (Cowlshaw,
246 1995) and connections within troops are common (Castles et al., 2014). For species with higher fission-fusion
247 organisation, where communities are more transient and home ranges can overlap substantially (e.g. eastern
248 grey kangaroos: Best et al., 2013), identifying communities and community membership is less
249 straightforward, and may influence the results of social network analyses. Once community structure has
250 been identified, we must ask which individuals should be included in the 'social network' for a given study.
251 Should all individuals in the local population be included, even if the majority never have a connection to
252 others (see our point above about zero-weighted edges)? Or should the communities be considered

253 separately, even if there are some (sometimes many) between-community connections? While at the node
254 level larger communities will result in larger sample sizes, a limit to the generalisability of network studies'
255 results is not how large the communities are but how many communities are assessed for a particular
256 research question (D. P. Croft, James, & Krause, 2008). For example, if a researcher is interested in the
257 transfer of information among individuals, the relevant unit of analysis is not the number of individuals in the
258 community but the number of communities in which the results can be replicated; the size of the community is
259 irrelevant (unless one is interested in the transfer of information in communities of different sizes, of course).
260 In our baboon system, in most cases we would rarely pool in a common network all of the individuals from
261 both of the communities we study because of the zero-weighted edges that would be generated, but after this
262 stage we may pool individuals (and control statistically for troop membership), as ever, depending on the
263 research question (as we did in Castles et al., 2014). However, we have no prescriptive advice for this
264 problem in other systems with more between-community connections; once again, we merely intend to
265 highlight an issue that is infrequently considered in social network studies which requires the careful attention
266 of researchers.

267 In conclusion, we reiterate that we do not argue that proximity data cannot or should not be used in
268 social network studies, nor that proximity data are not informative, and we appreciate that in many systems
269 proximity is the only readily available measure of association between individuals. We only caution against
270 assuming that proximity is necessarily a proxy for interactions, and encourage that this assumption is tested
271 should it be used. We also advocate that the research question and study species biology should drive the
272 definition of edges (and nodes) in networks as well as the social scales at which these are measured.

273

274 Acknowledgements

275 We thank Dr Guy Cowlshaw for comments on an earlier version of this manuscript, and two anonymous
276 reviewers. AJC is supported by a Junior Research Fellowship from Churchill College, University of
277 Cambridge. This is a publication of ZSL Institute of Zoology's *Tsaobis Baboon Project*.

278

279

280 References

281

282 Best, E. C., Seddon, J. M., Dwyer, R. G., & Goldizen, A. W. (2013). Social preference influences female
283 community structure in a population of wild eastern grey kangaroos. *Animal Behaviour*, 86(5), 1031-
284 1040. doi: <http://dx.doi.org/10.1016/j.anbehav.2013.09.008>

285 Carter, A. J., Macdonald, S. L., Thomson, V. A., & Goldizen, A. W. (2009). Structured association patterns and
 286 their energetic benefits in female eastern grey kangaroos, *Macropus giganteus*. *Animal Behaviour*,
 287 77(4), 839-846. doi: DOI 10.1016/j.anbehav.2008.12.007

288 Castles, M., Heinsohn, R., Marshall, H. H., Lee, A. E. G., Cowlshaw, G., & Carter, A. J. (2014). Social networks
 289 created with different techniques are not comparable. *Animal Behaviour*, 96(0), 59-67. doi:
 290 <http://dx.doi.org/10.1016/j.anbehav.2014.07.023>

291 Cowlshaw, G. (1995). Behavioural Patterns in Baboon Group Encounters: The Role of Resource Competition
 292 and Male Reproductive Strategies. *Behaviour*, 132(1/2), 75-86. doi: 10.2307/4535250

293 Croft, D. P., James, R., & Krause, J. (2008). *Exploring Animal Social Networks*. Princeton: Princeton University
 294 Press.

295 Croft, D. P., Krause, J., & James, R. (2004). Social networks in the guppy (*Poecilia reticulata*). [Journal Article].
 296 *Biology Letters*, 271(Suppl 6), S516-S519. doi: 10.1098/rsbl.2004.0206

297 Farine, D. R. (2015). Proximity as a proxy for interactions: issues of scale in social network analysis. *Animal*
 298 *Behaviour*, in press.

299 Hobson, E. A., Avery, M. L., & Wright, T. F. (2013). An analytical framework for quantifying and testing
 300 patterns of temporal dynamics in social networks. *Animal Behaviour*, 85(1), 83-96. doi:
 301 <http://dx.doi.org/10.1016/j.anbehav.2012.10.010>

302 Kummer, H. (1984). From laboratory to desert and back: A social system of hamadryas baboons. *Animal*
 303 *Behaviour*, 32(4), 965-971. doi: [http://dx.doi.org/10.1016/S0003-3472\(84\)80208-0](http://dx.doi.org/10.1016/S0003-3472(84)80208-0)

304 Lehmann, J., & Ross, C. (2011). Baboon (*Papio anubis*) social complexity - a network approach. *American*
 305 *Journal of Primatology*, 73(8), 775-789. doi: Doi 10.1002/Ajp.20967

306 Macdonald, S., & Voelkl, B. (2015). Primate social networks. In J. Krause, R. James & D. W. Franks (Eds.),
 307 *Animal Social Networks* (pp. 125-138). Oxford: Oxford University Press.

308 Madden, J., Drewe, J., Pearce, G., & Clutton-Brock, T. (2011). The social network structure of a wild meerkat
 309 population: 3. Position of individuals within networks. *Behavioral Ecology and Sociobiology*, 65(10),
 310 1857-1871. doi: 10.1007/s00265-011-1194-2

311 Symington, M. M. (1990). Fission-fusion social organization in *Ateles* and *Pan*. *International Journal of*
 312 *Primatology*, 11(1), 47-61. doi: 10.1007/bf02193695

313 Whitehead, H. (2008). *Analysing Animal Societies*. Chicago: University of Chicago Press.

314 Whitehead, H., & Dufault, S. (1999). Techniques for Analyzing Vertebrate Social Structure Using Identified
 315 Individuals: Review and Recommendations. *Advances in the Study of Behavior*, 28, 33-74. doi:
 316 [http://dx.doi.org/10.1016/S0065-3454\(08\)60215-6](http://dx.doi.org/10.1016/S0065-3454(08)60215-6)

317 Wittemyer, G., Douglas-Hamilton, I., & Getz, W. M. (2005). The socioecology of elephants: analysis of the
 318 processes creating multitiered social structures. *Animal Behaviour*, 69(6), 1357-1371. doi:
 319 <http://dx.doi.org/10.1016/j.anbehav.2004.08.018>

320

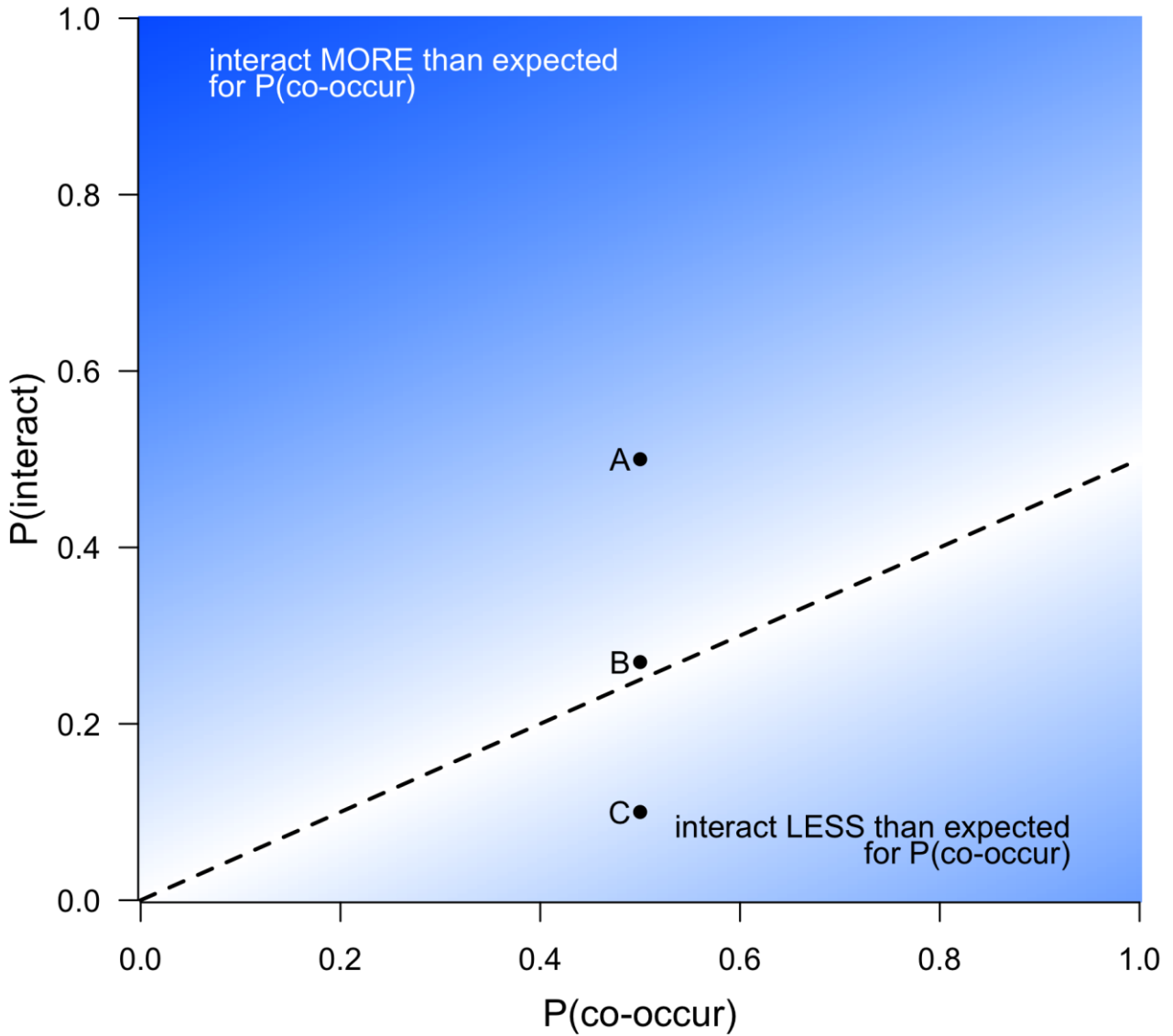
321

322

323

324 Figure 1

325 The relationship between the probability of interacting ($P(\text{interact})$) for a given probability that a dyad will co-
326 occur in the same group ($P(\text{co-occur})$). The dashed line represents the average interaction rate for the
327 population. The blue shading represents whether individuals are more or less likely to interact than expected
328 for the average of the population, with lighter (white) shading showing that dyads interact at the average rate.
329 Three hypothetical dyads (A, B and C) are shown (see text for details).



330