

31 and (iii) a combination of the two. Commonly-used methods aimed at dealing with
32 incomplete observations perform poorly because they are based on arbitrary
33 observation probabilities. We then derive complete indices that can be calibrated
34 for the different types of observation probabilities to generate accurate estimates of
35 association rates. These are provided in an R package that readily interfaces with
36 existing routines. We conclude that using calibration data is an important step when
37 constructing animal social networks, and that in their absence, researchers should
38 use a simple estimator and explicitly consider the impact of this on their findings.

39

40 KEYWORDS

41 Affiliations, Animal social networks, Interactions, Social network analysis, Social
42 structure

43

44

45 INTRODUCTION

46

47 A foundation of animal social network analysis is estimating the frequency that two
48 individuals associate or interact. Social networks are typically a description of
49 interconnections that are formed by relationships (edges) among multiple
50 individuals (nodes). Social network analysis is a set of tools that can be used to
51 describe the patterns formed by these interconnections or evaluate these against
52 hypotheses (Farine & Whitehead, 2015; Whitehead, 2008). One feature of social
53 network analysis that is perhaps unique to studies on animal populations is that
54 researchers rarely have a complete record of all interactions or all associations (but
55 see Boogert, Farine, & Spencer, 2014; Farine, Spencer, & Boogert, 2015; Strandburg-
56 Peshkin, Farine, Couzin, & Crofoot, 2015). Thus, relationships are often imperfectly
57 sampled, which can introduce uncertainty in the social network. To account for
58 variation in sampling effort and observation frequency, Cairns and Schwager (1987)
59 outlined commonly-used association indices. These indices convert the number of
60 observations of pairs of individuals seen associating or interaction into an

61 association rate, representing their propensity to associate or their probability of
62 being observed together.

63

64 Incomplete sampling of animal interactions or associations can occur due to a range
65 of different reasons. We can classify datasets as having two possible types of missing
66 data (Cairns & Schwager, 1987): (i) single or few observers can only collect data on
67 one or a few groups at a time and miss many simultaneous associations or
68 interactions occurring elsewhere, and (ii) individuals are difficult to observe or
69 identify and missed even when they are present. In type (i), while a number of pairs
70 of individuals (also known as dyads) are being observed together in one or more
71 groups, the status of other individuals in the population is unobserved. In type (ii),
72 when one or more groups are being observed they are incompletely sampled,
73 resulting in data that suggests that certain dyads were not interacting or associating
74 even when they were and could have been observed doing so. In both cases, the
75 relationships inferred from the observed data is likely to be influenced by the
76 amount of data that was missed. However, the propensity for each type of missing
77 observations to impact our estimates of association or interaction rates and social
78 network structure remains to be properly explored.

79

80 Properly controlling for missed observations is one of the most important steps in
81 social network analysis. Using simulated data, Franks, Ruxton, and James (2010)
82 identified the impact of missing observations when constructing social networks.
83 They found that missing observations between known individuals was more
84 problematic than missing individuals altogether, and concluded that social network
85 sampling should maximize the amount of data collected about known individuals
86 rather than maximizing the number of individuals sampled. One reason for this is
87 because a key component of social networks, weak edges, are often
88 disproportionately likely to be missed, and leaving these out can have profound
89 implications on the structure of the social network (Granovetter, 1973). These
90 findings are also supported by the work of M. J. Silk, Jackson, Croft, Colhoun, and
91 Bearhop (2015) who explored the effect of completely missing individuals in the

92 social network. They found that, with adequate sampling, having as few as 30% of
93 individuals known can be enough to produce informative networks for hypothesis
94 testing.

95
96 Missing observations that could have been recorded can have large impacts on the
97 social network that is generated, and these impacts are made worse when particular
98 individuals are missed more often than others. Farine and Whitehead (2015)
99 recently demonstrated how small differences in the likelihood of observing
100 individuals of different classes can introduce systematic biases in their social
101 network. They first simulated observations of individuals associating with preferred
102 and avoided associates. They then introduced a small observation bias, in this case
103 reducing the probability of observing one of two classes of individuals to 80% by
104 removing 20% of the observations of those individuals. This resulted in a significant
105 effect of class on degree (the sum of the association strengths in the nodes with
106 intact data was higher than in the nodes where data had been removed). This means
107 that the social network estimated for the individuals in this population is incorrect.

108
109 In this paper, we theoretically re-evaluate existing association indices and derive
110 new measures to deal with missing observations of groups, missing individuals in
111 groups, and the combination of these. We show that the extent that existing
112 association indices adjust estimates of association strength is entirely arbitrary, and
113 are as likely to over-correct any bias that might occur as they are to reduce it.
114 Existing association indices can also perform poorly at estimating relative
115 association strengths, which has implications for many social network studies. We
116 then derive improved association indices that enable researchers to correct
117 properly for the biases arising from group location error and individual
118 identification error, and discuss how to collect appropriate calibration data. Finally,
119 we provide an R package “assocInd” that allows researchers to calculate accurate
120 association indices for pairs of individuals from their observation data, and to
121 simulate the effects of different types of errors on estimates of associations.

122

123

124 THE SIMPLE RATIO AND THE HALF-WEIGHT INDEX

125

126 In many cases we wish to calculate an association index that estimates the
127 proportion of time any two individuals, a and b , spend associated. Association
128 indices typically range from 0 (the two individuals were never observed together)
129 to 1 (the individuals are always seen together), and the association rates are used as
130 a proxy to quantify the propensity for pairs of individuals to interact (Farine, 2015;
131 Whitehead & Dufault, 1999), although the assumption that individuals interact in
132 proportion to their association rate should be considered on a case-by-case basis
133 (Castles et al., 2014). Association data is frequently collected by repeatedly sampling
134 the population, and recording whom is observed in the same group in each sampling
135 period. For any two individuals we can then calculate:

136 x the number of sampling periods with a and b observed associated
137 y_a the number of sampling periods with just a identified
138 y_b the number of sampling periods with just b identified
139 y_{ab} the number of sampling periods with a and b identified but not associated
140 y_{Null} the number of sampling periods with neither a nor b identified

141

142 In an ideal scenario, every individual is seen and correctly identified in every
143 sampling period, such as in many captive populations, or at least we have the
144 situation where $y_{Null} = 0$. Intuitively, in the ideal scenario researchers can validly use
145 the simple ratio index (SRI), $x/(y_a + y_b + y_{ab} + x)$, as an estimate of the proportion
146 of time A and B spend together. However, when errors arise from missing
147 observations of individuals or groups, it is less clear that the simple ratio is
148 appropriate. The most commonly-used approach for correcting association indices
149 to account for missing observations is to reduce the weighting given to observations
150 of just one individual (because we have a lower confidence in these). Because
151 missing observations are widespread in behavioural research, many researchers use
152 the half-weight index (HWI): $x/(\frac{1}{2}(y_a + y_b) + y_{ab} + x)$. This index is believed to

153 correct for the biases arising from such error, in particular when individuals are
154 relatively more likely to be detected when they are apart than when they are
155 together. When investigating the performance of association indices, Cairns and
156 Schwager (1987) found that the HWI resulted in lower bias and lower error for a
157 given estimate than the simple ratio when observations were missed. However,
158 whilst this has served as useful justification for many researchers, it is also
159 important to note that Cairns and Schwager (1987) reported up to 4 times greater
160 error in the HWI than what they achieved using a maximum likelihood function (see
161 also below). Further, they noted a number limitations of association indices arising
162 from hidden assumptions.

163

164 Here we revisit some of the assumptions of the half-weight index. Notably, we show
165 that the extent to which the half weight index adjusts estimates of association is
166 entirely arbitrary, and is as likely to “overcorrect” any bias that might occur as it is
167 to reduce that bias. Note that an alternative variant to the HWI, the twice-weight
168 index (TWI) $x/(2(y_a + y_b) + y_{ab} + x)$, is a monotonic function of the HWI and thus
169 we do not investigate it in this paper. Ginsberg and Young (1992) previously raised
170 the issue that the HWI and TWI use arbitrary weightings, and predicted that
171 association indices will continue to be widely used. Indeed, the HWI is still the most
172 commonly-used index in animal social network studies.

173

174 To address the need to properly correct for biases arising from group location error
175 and individual identification error, we derive improved association indices that can
176 be calibrated independently for each study. We start by addressing the impact of
177 group location error before moving on to the effect of individual identification error,
178 and finally the combination of the two.

179

180

181

182

183

184 CORRECTING FOR GROUP LOCATION ERROR

185

186 Here we start with the assumption that if you see a in a group but do not see b in
187 that group, you know a is not with b , and vice versa. This assumption will be valid if
188 there is no individual identification error, i.e. all the individuals in a group that is
189 located by the researcher will always be identified. However, uncertainty remains
190 for all of the sampling periods in which we did not see a or b , since we do not know
191 if they were together during such periods.

192

193 Let us denote the event that a and b are together in a sampling period as ab and the
194 event that they are not together as $!ab$. The aim is to estimate the association
195 between a and b , $a_{ab} = p(ab)$. We start by developing a maximum likelihood
196 estimator (MLE) for a_{ab} . In any given sampling period, the probability we see only
197 individual a (i.e. not b), is given by:

198

$$199 \quad P(\text{see } a, \text{ not see } b) = p(\text{see } a | !ab) (1 - p(\text{see } b | !ab)) (1 - a_{ab})$$

200

201 Note that $p(!ab) = 1 - p(ab) = 1 - a_{ab}$. The probability of seeing a and b in different
202 groups is given by:

203

$$204 \quad P(\text{see } a, \text{ see } b) = p(\text{see } a | !ab) p(\text{see } b | !ab) (1 - a_{ab})$$

205

206 The probability of seeing a and b together in a group is:

207

$$208 \quad P(\text{see } ab) = p(\text{see } ab | ab) a_{ab}$$

209

210 And the probability of seeing neither a nor b is:

211

$$212 \quad P(\text{not see } a, \text{ not see } b) = (1 - p(\text{see } a | !ab)) (1 - p(\text{see } b | !ab)) (1 - a_{ab}) + (1 - p(\text{not} \\ 213 \quad \text{see } ab | ab) a_{ab}$$

214

215 From this we can derive the log-likelihood, L , for the data obtained:

$$\begin{aligned}
 216 \quad L = & y_a [\log(P(\text{see } a|ab)) + \log([1-P(\text{see } b|!ab))] + \log(1 - a_{ab})] \\
 217 \quad & + y_b [\log(P(\text{see } b|!ab)) + \log([1-P(\text{see } a|!ab))] + \log(1 - a_{ab})] \\
 218 \quad & + y_{ab} [\log(P(\text{see } a|!ab)) + \log(P(\text{see } b|!ab)) + \log(1 - a_{ab})] \\
 219 \quad & + x [\log(P(\text{see } ab|ab)) + \log(a_{ab})] \\
 220 \quad & + y_{Null} \log[(1-P(\text{see } a|!ab))(1-P(\text{see } b|!ab))(1 - a_{ab}) + (1-P(\text{see } ab|ab))a_{ab}]
 \end{aligned}$$

221

222 This simplifies to:

223

$$\begin{aligned}
 224 \quad L = & (y_a + y_b + y_{ab}) \log(1 - a_{ab}) + x \log(a_{ab}) + y_a [\log(1-A) + \log(B)] + y_b [\log(1- \\
 225 \quad & B) + \log(A)] + y_{ab} [\log(1-A) + \log(1-B)] + x \log(1-C) + y_{Null} \log[AB(1 - a_{ab}) + C a_{ab}]
 \end{aligned}$$

226

227 where $A = P(!\text{see } a|!ab)$; $B = P(!\text{see } b|!ab)$; $C = P(!\text{see } ab|ab)$. We can find the

228 maximum likelihood estimator, \widehat{a}_{ab} by solving:

229

$$\frac{dL}{da_{ab}} = -\frac{(y_a + y_b + y_{ab})}{1 - \widehat{a}_{ab}} + \frac{x}{\widehat{a}_{ab}} + y_{Null} \frac{C - AB}{C a_{ab} + AB(1 - \widehat{a}_{ab})} = 0$$

230

231 In practice A (the probability of not seeing a when a and b are not together), B (the
 232 probability of not seeing b when a and b are not together), and C (the probability of
 233 not seeing a or b when they are together) will not be known. Thus, to estimate an
 234 accurate value for the association strength between two individuals requires
 235 validation data at the level of individuals. However, progress can be made by making
 236 different assumptions about the relationship between A , B and C . First, if $C = AB$, the
 237 MLE is:

238

$$\widehat{a}_{ab} = \frac{x}{(y_a + y_b + y_{ab} + x)}$$

239

240 which is the simple ratio index. So, if the probability of failing to see a and b together
 241 is the same as the probability of failing to see both when they are apart, then the SRI

242 is valid. Note also that the simple ratio is valid as an MLE if $y_{Null} = 0$, as intuition
 243 suggests.

244

245 Alternatively, we could assume that $C = AB(1 + \omega)$, where failing to observe a
 246 group containing a and b is more ($\omega > 0$) or less ($\omega < 0$) likely than failing to
 247 observe both the group containing a and the group containing b when a and b are
 248 not together. In this case the MLE is given by the solution to:

249

$$-\frac{(y_a + y_b + y_{ab})}{1 - \widehat{a}_{ab}} + \frac{x}{\widehat{a}_{ab}} + y_{Null} \frac{\omega}{\omega \widehat{a}_{ab} + 1} = 0$$

250

251 which can be re-arranged to form a quadratic:

252

$$-\omega T \widehat{a}_{ab}^2 + (x\omega + y_{Null}\omega - T_i)\widehat{a}_{ab} + x = 0,$$

254

255 where T_i is the number of directly informative sampling periods, i.e. $T_i = y_a + y_b +$
 256 $y_{ab} + x$, and T is the total number of sampling periods, i.e. $T = I + y_{Null}$. The MLE is
 257 given by the lower root of this equation, i.e.

258

$$\widehat{a}_{ab} = \frac{-(x\omega + y_{Null}\omega - T_i) - \sqrt{(x\omega + y_{Null}\omega - T_i)^2 + 4\omega T x}}{-2\omega T}$$

259

260 We term this index the group location error corrected index (GLECI). As expected,
 261 the GLECI reduces to the simple ratio index $\widehat{a}_{ab} = x/T$ when $y_{Null} = 0$. The standard
 262 error (see Appendix for derivation) can be calculated as for a proportion,

263 $\sqrt{\widehat{a}_{ab}(1 - \widehat{a}_{ab})/N_e}$, but with an effective sample size of:

264

$$N_e = T_i + y_{Null} \frac{\omega a_{ab}(1 - a_{ab})}{(\omega a_{ab} + 1)^2}$$

265

266 assuming sampling periods are sufficiently spaced in time and/or space to assume
267 they are independent.

268

269

270 CORRECTING FOR INDIVIDUAL IDENTIFICATION ERROR

271

272 Here we assume that there is no group location error, but define the probability of
273 failing to identify an individual in a group that has been under observation, i.e. the
274 individual observation error rate, as ϵ . We suspect this scenario will be rare, as
275 individual identification error will usually be accompanied by group location error
276 (see next section). However, we consider the case in order to analyse the effect of
277 the two types of error. We get the following probabilities:

278

$$p(\text{see } a, \text{see } b) = (1 - \epsilon)^2(1 - a_{ab})$$

$$p(\text{see } ab) = (1 - \epsilon)^2 a_{ab}$$

$$p(\text{not see } a, \text{not see } b) = \epsilon^2$$

$$p(\text{see } a, \text{not see } b) = +\epsilon(1 - \epsilon)$$

279

280 Therefore the likelihood for the data can be obtained as:

$$L = y_{ab}[2 \log(1 - \epsilon) + \log(1 - a_{ab})] + x[2 \log(1 - \epsilon) + \log(a_{ab})] + y_{null}[2 \log(\epsilon)] \\ + (y_a + y_b)[\log(\epsilon) + \log(1 - \epsilon)]$$

281

282 And the MLE found as follows:

$$\frac{dL}{da_{ab}} = \frac{-y_{ab}}{1 - \widehat{a}_{ab}} + \frac{x}{\widehat{a}_{ab}} = 0$$

283

$$(-y_{ab} - x)\widehat{a}_{ab} + x = 0$$

$$\widehat{a}_{ab} = \frac{x}{y_{ab} + x}$$

284

285 In this case the MLE \widehat{a}_{ab} is a simple proportion, using only the x and y_{ab} counts
 286 taken to the informative data, and requiring no calibration data. The terms y_a and y_b
 287 are not used, in contrast to the simple ratio and half weight indices, since they are
 288 known to be unreliable in this scenario: if only a is recorded, it is possible that b was
 289 associated with a and has been missed through individual identification error. We
 290 call this index the very simple ratio index (vSRI). The standard error (see Appendix
 291 for derivation) is calculated for a proportion as usual with $\sqrt{a_{ab}(1 - a_{ab})/N_e}$, with
 292 an effective sample size of $N_e = y_{ab} + x$, assuming sampling periods are sufficiently
 293 spaced in time and/or space to assume they are independent.

294

295

296 GENERAL ERROR MODEL

297

298 In practice, both types of error are likely to occur in a given sampling procedure. To
 299 model this situation, we define a model with a more general relationship between
 300 the errors in each count. We define:

301

$$\begin{aligned}
 p(\text{see } a, \text{see } b) &= (1 - \epsilon_{a|ab})(1 - \epsilon_{b|ab})(1 - a_{ab}) \\
 p(\text{see } ab) &= (1 - \epsilon_{a|ab} - \epsilon_{b|ab} - \epsilon_{ab|ab})a_{ab} \\
 p(\text{not see } a, \text{not see } b) &= \epsilon_{ab|ab} a_{ab} + \epsilon_{a|ab}\epsilon_{b|ab} (1 - a_{ab}) \\
 p(\text{see } a, \text{not see } b) &= \epsilon_{b|ab} a_{ab} + (1 - \epsilon_{a|ab}) \epsilon_{b|ab} (1 - a_{ab})
 \end{aligned}$$

302

303 where $\epsilon_{a|ab}$ is the probability of missing a , given a is not with b ; $\epsilon_{a|ab}$ is the
 304 probability of missing a , but not b , given a is with b ; and $\epsilon_{ab|ab}$ is the probability of
 305 missing a and b given they are together.

306

307 Note that the group location error scenario is the special case with $\epsilon_{a|ab} = A$;

308 $\epsilon_{b|ab} = B$; $\epsilon_{ab|ab} = C = (1 + \omega)AB = (1 + \omega) \epsilon_{a|ab}\epsilon_{b|ab}$; $\epsilon_{a|ab} = \epsilon_{b|ab} = 0$. The

309 individual identification error scenario is given by $\epsilon_{a|!ab} = \epsilon_{b|!ab} = \epsilon$; $\epsilon_{a|ab} = \epsilon_{b|ab} =$
 310 $\epsilon(1 - \epsilon)$; $\epsilon_{ab|ab} = \epsilon^2$.

311

312 Let us assume that the probability of missing both a and b when they are in the same
 313 group is $(1 + \omega)$ x the probability of missing a and b when they are not together
 314 (where $\omega > -1$). The probability both a and b will be missed when they are
 315 together is therefore:

316

$$317 \quad \epsilon_{ab|ab} = (1 + \omega) \epsilon_{a|!ab} \epsilon_{b|!ab}$$

318

319 The probability of at least one of a or b being missed is $\acute{\epsilon} = 1 - p(\text{see } ab|ab)$. Let us
 320 set $\epsilon_{ab|ab} = \phi \acute{\epsilon}$, where $0 < \phi \leq 1$. Since $\acute{\epsilon} = \epsilon_{a|ab} + \epsilon_{b|ab} + \epsilon_{ab|ab}$

321

$$322 \quad \epsilon_{a|ab} + \epsilon_{b|ab} = (1 - \phi) \acute{\epsilon}$$

$$\epsilon_{a|ab} + \epsilon_{b|ab} = \frac{(1 - \phi)(1 + \omega) \epsilon_{a|!ab} \epsilon_{b|!ab}}{\phi}$$

323

324 If we assume that $\epsilon_{a|ab} / (\epsilon_{a|ab} + \epsilon_{b|ab}) = \epsilon_{a|!ab} / (\epsilon_{a|!ab} + \epsilon_{b|!ab})$, this gives us:

325

$$\epsilon_{a|ab} = \frac{(1 - \phi)(1 + \omega) \epsilon_{a|!ab}^2 \epsilon_{b|!ab}}{\phi(\epsilon_{a|!ab} + \epsilon_{b|!ab})}$$

326

327 where ϕ determines the relative importance of group location error relative to
 328 individual identification error, with the group location error model given when
 329 $\phi = 1$. We can now refine the probabilities given above:

330

$$p(\text{see } a, \text{ see } b) = (1 - \epsilon_{a|!ab})(1 - \epsilon_{b|!ab})(1 - a_{ab})$$

$$p(\text{see } ab) = (1 - (1 + \omega) \epsilon_{a|!ab} \epsilon_{b|!ab} / \phi) a_{ab}$$

$$p(\text{not see } a, \text{ not see } b) = (1 + \omega) \epsilon_{a|!ab} \epsilon_{b|!ab} a_{ab} + \epsilon_{a|!ab} \epsilon_{b|!ab} (1 - a_{ab})$$

$p(\text{see } a, \text{ not see } b)$

$$= \epsilon_{b|ab} \left(\frac{(1-\phi)(1+\omega) \epsilon_{b|ab} \epsilon_{a|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})} a_{ab} + (1 - \epsilon_{a|ab})(1 - a_{ab}) \right)$$

331

332 Giving a log likelihood of:

333

$$\begin{aligned} L = & y_a \left[\log(\epsilon_{b|ab}) \right. \\ & \left. + \log \left(\frac{(1-\phi)(1+\omega) \epsilon_{b|ab} \epsilon_{a|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})} a_{ab} + (1 - \epsilon_{a|ab})(1 - a_{ab}) \right) \right] \\ & + y_b \left[\log(\epsilon_{a|ab}) \right. \\ & \left. + \log \left(\frac{(1-\phi)(1+\omega) \epsilon_{b|ab} \epsilon_{a|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})} a_{ab} + (1 - \epsilon_{b|ab})(1 - a_{ab}) \right) \right] \\ & + y_{ab} [\log(1 - \epsilon_{a|ab}) + \log(1 - \epsilon_{b|ab}) + \log(1 - a_{ab})] \\ & + x [\log(1 - (1 + \omega) \epsilon_{a|ab} \epsilon_{b|ab}) + \log(a_{ab})] \\ & + y_{null} \log[(1 + \omega) \epsilon_{a|ab} \epsilon_{b|ab} a_{ab} + \epsilon_{a|ab} \epsilon_{b|ab} (1 - a_{ab})] \end{aligned}$$

334

335 To obtain the MLE, \widehat{a}_{ab} , we need to solve the equation:

$$\begin{aligned} \frac{dL}{da_{ab}} = & \frac{y_a \left(\frac{(1-\phi)(1+\omega) \epsilon_{b|ab} \epsilon_{a|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})} - (1 - \epsilon_{a|ab}) \right)}{\frac{(1-\phi)(1+\omega) \epsilon_{b|ab} \epsilon_{a|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})} a_{ab} + (1 - \epsilon_{a|ab})(1 - a_{ab})} \\ & + \frac{y_b \left(\frac{(1-\phi)(1+\omega) \epsilon_{b|ab} \epsilon_{a|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})} - (1 - \epsilon_{b|ab}) \right)}{\frac{(1-\phi)(1+\omega) \epsilon_{b|ab} \epsilon_{a|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})} a_{ab} + (1 - \epsilon_{b|ab})(1 - a_{ab})} - \frac{y_{ab}}{1 - \widehat{a}_{ab}} \\ & + \frac{x}{\widehat{a}_{ab}} + \frac{y_{null} \left((1 + \omega) \epsilon_{a|ab} \epsilon_{b|ab} - \epsilon_{a|ab} \epsilon_{b|ab} \right)}{(1 + \omega) \epsilon_{a|ab} \epsilon_{b|ab} \widehat{a}_{ab} + \epsilon_{a|ab} \epsilon_{b|ab} (1 - \widehat{a}_{ab})} = 0 \end{aligned}$$

336

337 To generate an estimate of association using this function firstly requires values for
338 ω (the group location error term) and ϕ (the error components importance term).

339 These could reasonably be calibrated at the population level, i.e. we could assume
340 that these quantities are constant across all dyads. However, the estimate also

341 requires estimates for $\epsilon_{a|ab}$ and $\epsilon_{b|ab}$ which require calibration data at the level of
342 individuals, making this approach infeasible in most cases. Nonetheless, we might

343 obtain an approximate solution, \widetilde{a}_{ab} if we substitute a population averaged estimate
344 (averaged across all dyads, or those dyads for which data is available) $\epsilon = \epsilon_{a|ab} =$

345 $\epsilon_{b|ab}$:

346

$$(y_a + y_b) \frac{((1 - \phi)/\phi)(1 + \omega) \epsilon - 2(1 - \epsilon)}{((1 - \phi)/\phi)(1 + \omega) \epsilon \widetilde{a}_{ab} + 2(1 - \epsilon)(1 - \widetilde{a}_{ab})} - \frac{y_{ab}}{1 - \widetilde{a}_{ab}} + \frac{x}{\widetilde{a}_{ab}} + \frac{y_{null}\omega}{\omega \widetilde{a}_{ab} + 1} = 0$$

347

348 The calibration measures required to solve this equation are ω , ϕ and ϵ , and can be
349 solved using a non-linear equation solver. We call the solution to this equation, \widetilde{a}_{ab} ,

350 the combined errors index (CEI). In our R package, we provide a function that

351 calculates the CEI in the R statistical environment (R Development Core Team,

352 2015), using the uniroot function in the rootSolve package (Soetaert & Herman,

353 2009). Note that by setting $\phi = 1$, we reduce the model to the group location error

354 model, giving \widetilde{a}_{ab} as the GLECI. We can also reduce the model to the individual

355 identification error model by setting $\epsilon_{a|ab} = \epsilon_{b|ab} = \epsilon(1 - \epsilon)$ and $\epsilon_{ab|ab} = \epsilon^2$, giving

356 us $\phi = \epsilon/(2 - \epsilon)$ and $\omega = 0$. Thus \widetilde{a}_{ab} reduces to the vSRI.

357

358

359 COMPARISON OF INDEX PERFORMANCE

360

361 In this section we examine how the SRI, HWI, GLECI, vSRI and CECI perform under

362 scenarios where there is group location error, individual identification error, and a

363 combination of the two. In simple cases we do this by first deriving expressions for
364 the expected value of each index, and then dividing by the value it is intended to
365 estimate, $E[I_{ab}]/a_{ab}$, thus showing us the circumstances under which I_{ab} is biased
366 upwards or downwards. However, we also recognise that in many circumstances
367 only the relative sizes of a_{ab} within the social network may be required, e.g. if
368 estimating scale free node-based or network metrics. Consequently, we also derive
369 $(E[I_{ab}]/E[I_{uv}])/(a_{ab}/a_{uv})$ to determine the circumstances under which each index
370 tends to overestimate or underestimate ratios of association values. Here u and v
371 denote a different dyad, so $(E[I_{ab}]/E[I_{uv}])/(a_{ab}/a_{uv})$ measures the bias when
372 index I_{ab} is used to estimate the relative strength of two associations.

373 In each case we also use simulations to illustrate the performance of the indices for
374 a given set of values, investigate bias where we were unable to do so analytically,
375 and examine the performance of Wald 95% confidence intervals calculated from the
376 standard errors presented above. For each scenario, we simulated 10,000 datasets
377 consisting of 1000 independent sampling periods for two individuals a and b , in the
378 R statistical environment (R Development Core Team, 2015). We start by allocating
379 the probability $a_{ab} = 0.5$ that a and b were associating in a given sampling period.
380 We then repeated all simulations with $a_{ab} = 0.25$ and $a_{ab} = 0.75$, and the results
381 we qualitatively similar, so here we present the results for $a_{ab} = 0.5$ only. We then
382 simulated the observation process according to the models described above, to yield
383 values for y_a, y_b, y_{ab}, x and y_{null} which we used to calculate the value of each of the
384 target association indices. For each scenario, we repeated the simulation for a range
385 of values of group location and individual identification errors. For the scenario
386 including only group location error, we ran simulations for a range of values of
387 $\omega = \{-0.9, -0.8, \dots, 2.5\}$ with $A = B = 0.5$, where $A = P(\text{!see } a|\text{!}ab)$ and $B = P(\text{!see } b|\text{!}ab)$. For the scenario including only individual identification error, we ran
389 simulations for a range of values of $\epsilon = \{0, 0.05, \dots, 0.95\}$. For the scenario with both
390 types of error, we varied $\omega = \{-0.9, -0.8, \dots, 0.9\}$, $\epsilon = \{0.1, 0.3, 0.5\}$ and
391 $\phi = \{0.1, 0.3, 0.5\}$, excluding impossible cases where $(1 + \omega) \epsilon^2 / \phi > 1$, since this
392 would mean there is a negative probability of observing a and b together. In each

393 case we recorded the mean value for each association index, in order to detect bias,
394 and the proportion of times the Wald 95% confidence intervals (calculated as
395 $\pm 1.96xSE$) contain the true value for a_{ab} .

396 However, recall that the CECI relies on an approximation, $\epsilon = \epsilon_{a||ab} = \epsilon_{b||ab}$, which
397 replaces the individual specific error rates with population level ones. The
398 simulations described above only test the performance of the CE when this
399 approximation holds in the data, i.e. when error rates are the same across all
400 individuals. Consequently, we re-ran simulations to test the performance of the CECI
401 when there was individual variation in error rate. In each case we set the population
402 mean error, ϵ , arbitrarily to 0.5, but drew individual errors from a normal
403 distribution with standard deviation $\sigma = \{0, 0.2, \dots, 2.0\}$, discarding and resampling
404 values that were < 0 or > 1 , and likewise for $\epsilon_{b||ab}$. We then conducted the
405 simulations as described above with $\phi = 0.5$.

406 *Group location error only*

407 We find that the GLECI is an unbiased estimator of a_{ab} (see Table 1 and Fig. 1)
408 across a range of group location errors ω . By contrast, the simple ratio is biased
409 upwards when $\omega < 0$ (i.e. when a and b are less likely to be missed when associated
410 than both are to be missed when apart) and biased downwards when $\omega > 0$. The
411 commonly used HWI shifts the estimate of a_{ab} upwards, such that it is biased
412 upwards when $AB(1 + \omega) < (A + B)/2$ and biased downwards when $AB(1 + \omega) >$
413 $(A + B)/2$. Consequently the HWI is only unbiased when the probability of seeing a
414 and b together is equal to the average of the probability of seeing each of them
415 apart. In our terminology, this is denoted $(1 - C) = ((1 - A) + (1 - B))/2$,
416 giving $C = (A + B)/2$. Thus under this scenario, the HWI assumes that the
417 probability of missing a and b when they are together is equal to half the probability
418 of missing either a or b when they are apart. This seems to us to be an arbitrary a
419 priori assumption, without the functionality to adjust the assumption using
420 supporting calibration data.

421 The vSRI performs very poorly when only group location error is present,
422 overestimating a_{ab} whilst $\omega < (A + B - 2AB)/AB$. This is because it excludes y_a and
423 y_b from the denominator on the assumption that these data are uninformative,
424 whereas in this scenario these are cases where we know that a and b were not
425 associating. As ω increases, an increasing number of cases where a and b were
426 associating are erroneously assigned to y_{null} . Thus, exclusion of y_{null} from the index
427 eventually offsets the positive bias (when $\omega = (A + B - 2AB)/AB$) resulting from
428 exclusion of y_a and y_b from the denominator. The vSRI is therefore not a useful
429 index as it contains assumptions that are unlikely to be met in the majority of
430 studies.

431 The GLECI is generally unbiased because y_{null} is included in the index in such a way
432 that excludes this positive bias. Importantly, the 95% confidence intervals for the
433 GLECI contained the true value of a_{ab} in close to 95% of cases, showing they perform
434 validly in this scenario (see Fig 1b). In contrast, the 95% confidence intervals
435 associated with the simple ratio, half weight index and IIEC index only performed
436 acceptably for a very narrow range of values of ω . Furthermore, the GLECI (and CEI
437 index with $\phi = 1$) is the only index of those considered that is unbiased when
438 estimating the ratio of two associations (see Table 1). This suggests that the SRI,
439 HWI and IIECI are not suitable for estimating either the relative or absolute strength
440 of associations when group location error is believed to be present, and that, ideally
441 the GLECI should be used if calibration data can be obtained.

442 Note that the simulations assumed that the researcher has an accurate estimate of ω
443 with which to calculate the GLECI (see below). In reality the better the estimate of ω
444 is, the better the estimate of a_{ab} will be, but even a rough estimate of ω will be
445 preferable to no calibration at all. Furthermore, we assume that the probability of
446 missing a and b when they are together will be $(1 + \omega)$ x that of missing both a and
447 b when they are separate ($C = (1 + \omega)AB$). Further work may conclude that this
448 relationship does not generally hold, in which case the GLECI might be suitably
449 modified to use a different calibration statistic. Nonetheless, the relationship posited
450 here requires weaker *a priori* assumptions to be made about the data than the

451 commonly used half weight index, which assumes that $C = (A+B)/2$. We show that
452 when this assumption is even slightly wrong, the half weight index will be a poor
453 estimate of a_{ab} .

454 One possible option to resolve the half weight index is to generalize it to be an M
455 weight index (MWI): $x/(M(y_a + y_b) + y_{ab} + x)$. The M weight index assumes that
456 the probability of seeing a and b together is equal to M x of the sum of the
457 probability of seeing each of them each apart. Equivalently, missing a and b when
458 they are together is M x more likely than missing them both when they are apart.
459 Thus the MWI could be calibrated to the data analogously to the GLECI, but each
460 index assumes a different relationship among the observation errors in the
461 population.

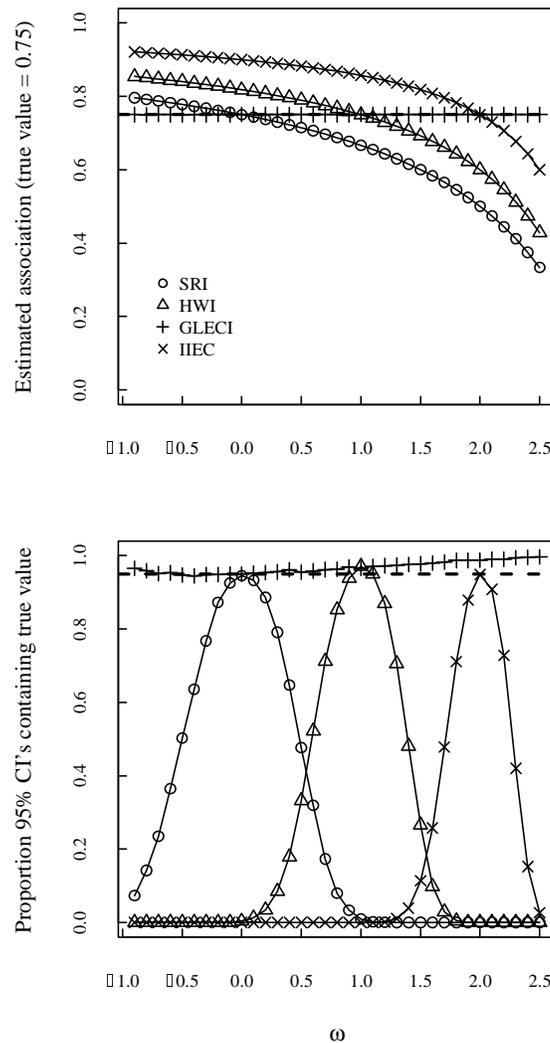
462 When calibration data cannot be obtained, there is a strong case for preferring the
463 simple ratio index. Use of the SRI results in biases that are more likely to be
464 qualitatively predictable in the absence of calibration data, than biases resulting
465 from the HWI or IIECI (see Table 1). For instance, when a researcher suspects that,
466 in general, missing two individuals when they are together is more likely than
467 missing both when they are apart, they can expect i) SRI values to be
468 underestimates of a_{ab} ; ii) for this underestimation to be more pronounced for
469 smaller values of a_{ab} and for less commonly seen individuals; and iii) for bigger
470 ratios between real associations to be overestimated relative to smaller ratios. Thus
471 a researcher can assess whether these inaccuracies are likely have any great bearing
472 on their conclusions in their specific case. In contrast, using the HWI we cannot
473 easily make such qualitative predictions unless we are in a position to judge the
474 relative size of $AB(1 + \omega)$ versus $(A + B)/2$; and, for the IIECI, ω versus
475 $(A + B - 2AB)/AB$. Making such judgments empirically is likely to be at least as
476 challenging as acquiring the calibration data required calculating the GLECI or a
477 calibrated MWI. Consequently, if only group location error is present and calibration
478 data cannot be obtained, we recommend use of the SRI, with careful consideration
479 of how the biases identified above might affect the interpretation of the study.

480 **Table 1:** Biases in different association measures arising from group location error.
 481 *signifies no bias inferred from simulations.

I_{ab}	$E[I_{ab}]/a_{ab}$	$E[I_{ab}/I_{uv}]/(a_{ab}/a_{uv})$
SRI	$\frac{(1 - AB(1 + \omega))}{1 - AB(1 + a_{ab}\omega)}$	$\frac{(1 - AB(1 + \omega))(1 - CD(1 + a_{cd}\omega))}{(1 - CD(1 + \omega))(1 - AB(1 + a_{ab}\omega))}$
	<ul style="list-style-type: none"> • Unbiased when $\omega = 0$ • Underestimates a_{ab} when $\omega > 0$: effect more pronounced for smaller values of a_{ab} and less commonly seen individuals • Overestimates a_{ab} when $\omega < 0$: effect more pronounced for smaller values of a_{ab} and less commonly seen individuals 	<ul style="list-style-type: none"> • When $\omega > 0$ bigger ratios exaggerated and biased in favour of the a_{ab} for pairs of individuals more commonly seen • When $\omega < 0$ smaller ratios exaggerated and biased in favour of the a_{ab} for pairs of individuals less commonly seen
HWI	$\frac{(1 - AB(1 + \omega))}{(1 - a_{ab})(1 - (A + B)/2) + (1 - AB(1 + \omega))a_{ab}}$	$\left[\frac{(1 - AB(1 + \omega))}{\left((1 - a_{uv})(1 - (U + V)/2) + (1 - UV(1 + \omega))a_{uv} \right)} \right]$ $\left[\frac{(1 - UV(1 + \omega))}{\left((1 - a_{ab})(1 - (A + B)/2) + (1 - AB(1 + \omega))a_{ab} \right)} \right]$
	<ul style="list-style-type: none"> • Unbiased when $AB(1 + \omega) = (A + B)/2$ • Underestimates a_{ab} when $AB(1 + \omega) > (A + B)/2$: effect is more pronounced for smaller values of a_{ab} • Overestimates a_{ab} when $AB(1 + \omega) < (A + B)/2$: effect is more pronounced for smaller values of a_{ab} • No straightforward relationship between bias and the frequency with which a and b are seen 	<ul style="list-style-type: none"> • Unbiased when $AB(1 + \omega) = (A + B)/2$ • Otherwise difficult to predict the pattern of bias in the data
HECI	$\frac{(1 - AB(1 + \omega))}{(1 - A)(1 - B) + a_{ab}(A + B - 2AB - AB\omega)}$	$\left[\frac{(1 - AB(1 + \omega))}{\left((1 - U)(1 - V) + a_{uv}(U + V - 2UV - UV\omega) \right)} \right]$ $\left[\frac{(1 - UV(1 + \omega))}{\left((1 - A)(1 - B) + a_{ab}(A + B - 2AB - AB\omega) \right)} \right]$
	<ul style="list-style-type: none"> • Unbiased when $\omega = (A + B - 2AB)/AB$, i.e. the ratio of the probability of missing only one of a or b to the probability of missing both, when a and b are apart. • Underestimates a_{ab} when 	<ul style="list-style-type: none"> • Unbiased when $\omega = (A + B - 2AB)/AB$

	$\omega > (A + B - 2AB)/AB$		• Otherwise difficult to predict the pattern of bias in the data
	• Overestimates a_{ab} when		
	$\omega < (A + B - 2AB)/AB$		
GLECI /	1	1	
CECI			
	• Unbiased for this error model*		• Unbiased for this error model*

482



483

484 **Figure 1:** a) Bias in different association indices as a function of group location
 485 error (ω) when applied to simulated data; b) performance of 95% Wald confidence
 486 intervals as a function of ω . Similar results were obtained for a true association
 487 value of 0.25 and 0.75.

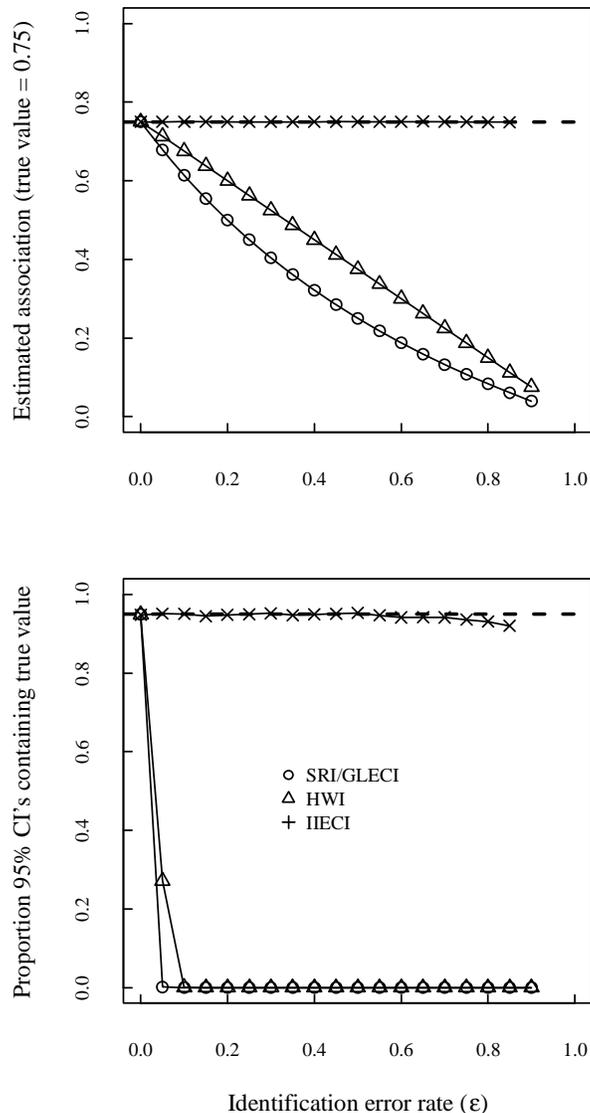
488 *Individual identification error only*

489 When only individual identification error is present we find that the IIECI is an
490 unbiased estimator of a_{ab} , whereas the simple ratio and half weight indices are
491 biased downwards whenever $\epsilon > 0$ (see Table 2 and Figure 2). This is because as ϵ
492 increases, more and more cases where a and b are associating are erroneously
493 attributed to y_a or y_b . Since y_a and y_b are included in the denominator for the SRI
494 and HWI, this results in an under-estimation of a_{ab} . The effect is reduced in the HWI
495 since y_a and y_b have a reduced weighting in this index. However, this is not
496 sufficient to ensure that the HWI is valid for even small individual identification
497 error rates. The GLECI was not included separately in these simulations since in this
498 scenario $(1 + \omega)\epsilon^2 = \epsilon^2$, so $\omega = 0$, meaning the GLECI reduces to the simple ratio.
499 Consequently, we can see that the GLECI also performs badly at estimating absolute
500 association values when there is individual identification error but no group
501 location error.

502 The 95% confidence intervals for the vSRI contained the true value of a_{ab} in close to
503 95% of cases, showing they perform validly under a scenario of only individual
504 identification error (see Fig 2b). This dropped slightly with very high error rates, as
505 a result of the effective sample size decreasing as most data is attributed to y_a , y_b or
506 y_{null} (not that with small sample sizes Wald confidence intervals are likely to be
507 anti-conservative: too narrow). In contrast, the 95% confidence intervals associated
508 with the SRI (and hence the GLECI) and HWI performed very badly with even small
509 individual identification error rates. Consequently, if a researcher is interested in
510 estimating the absolute values of a_{ab} and only individual identification error is likely
511 to be present (we anticipate this scenario to be rare), we recommend use of the
512 vSRI, which requires no calibration data.

513 In contrast to the group location error only, all association indices gave unbiased
514 estimates of the relative size of associations between pairs of individuals (see Table
515 2). This means that if the research aims are purely in the scale free properties of a
516 system, such as the relative position of individuals in a social network (Aplin, Firth,

517 et al., 2015; Wilson, Krause, Dingemanse, & Krause, 2013), any of the indices
518 considered will be sufficient. However, if there is also a risk of group location error,
519 and appropriate calibration data cannot be obtained (see next section), we
520 recommend use of the SRI to estimate relative associations due to the advantages in
521 interpreting this index in the presence of group location error (see above).



522
523 **Figure 2:** a) Bias in different association indices as a function of individual
524 identification error (ω) when applied to simulated data; b) performance of 95%
525 Wald confidence intervals as a function of ω . Similar results were obtained for a true
526 association value of 0.25 and 0.75.

527 **Table 2:** Biases in different association measures arising from individual
 528 identification error

I_{ab}	$E[I_{ab}]/a_{ab}$	$E[I_{ab}/I_{uv}]/(a_{ab}/a_{uv})$
SRI	$1 + \epsilon(2 - \epsilon)$	1
HWI	$1 + \epsilon$	1
IIECI/CECI	1	1
GLECI	$1 + \epsilon(2 - \epsilon)$	1

529

530

531 *Combined errors*

532 We find that the CECI is an unbiased estimator of a_{ab} across the possible range of
 533 values for ϵ , ω and ϕ when we assumed $\epsilon_{a|ab} = \epsilon$ for all individuals (see Fig S1 in
 534 ESM). Furthermore, the CECI was also an unbiased estimator of a_{ab} when $\epsilon_{a|ab}$ was
 535 allowed to vary across the population regardless of the magnitude of variation in
 536 $\epsilon_{a|ab}$ (see Fig S2 in ESM). In contrast, the SRI, HWI, IIECI and GLECI were biased in
 537 a manner that was dependent on the combination of values for ϵ , ω and ϕ . When ϕ
 538 was close to 1, the pattern of bias was similar to when only group location error was
 539 present. In other words, when it is unlikely that only one of a and b will be missed
 540 when they are in the same group, bias is similar to when we have only group
 541 location error. As ϕ became smaller (more likely that only one of a and b will be
 542 missed when they are in the same group) all four indices start to underestimate a_{ab}
 543 at a lower value of ω . The effects of both ϕ and ω are magnified more as ϵ gets
 544 larger. Consequently, if individual identification errors are likely to be common in
 545 addition to group location error, we suggest calibration data is acquired to estimate
 546 ϵ , ω and ϕ and the CECI is used. The pattern of bias in the other indices will be
 547 difficult to predict qualitatively unless the risk of individual identification error is
 548 known to be small. Consequently, if calibration data cannot be obtained under such
 549 circumstances we suggest extra efforts are made to minimise individual

550 identification error, and the SRI be used with the understanding that it will provide
551 noisy estimates of a_{ab} .

552 When we assume observation error is homogeneous across the population, the 95%
553 confidence intervals for the CECI tend to contain the true value of a_{ab} in >95% of
554 cases (see Fig S3). This suggests that, at large sample sizes at least, the Wald
555 confidence intervals are slightly too wide. However, given that Wald confidence
556 intervals are always an approximation and are widely used in statistics, this is a
557 minor concern. However, when observation error varied greatly across the
558 population, the 95% confidence intervals became far too narrow (see Fig S4), as a
559 result of the extra uncertainty that is unaccounted for in the derivation of the
560 standard error. Correcting the standard errors for this uncertainty does not seem
561 straightforward, though further work could address this if the CECI proves to be
562 useful and becomes widely adopted. Therefore, we suggest that the standard errors
563 and confidence intervals for the CECI be trusted as approximately valid if the
564 variation in observation rate is believed to be small, and not be trusted if that
565 variation is believed to be large.

566 Our recommendations for the choice of association index are shown as a flowchart
567 in Fig. 3. The indices and their standard errors are shown in Table 3.

568

569 OBTAINING CALIBRATION DATA

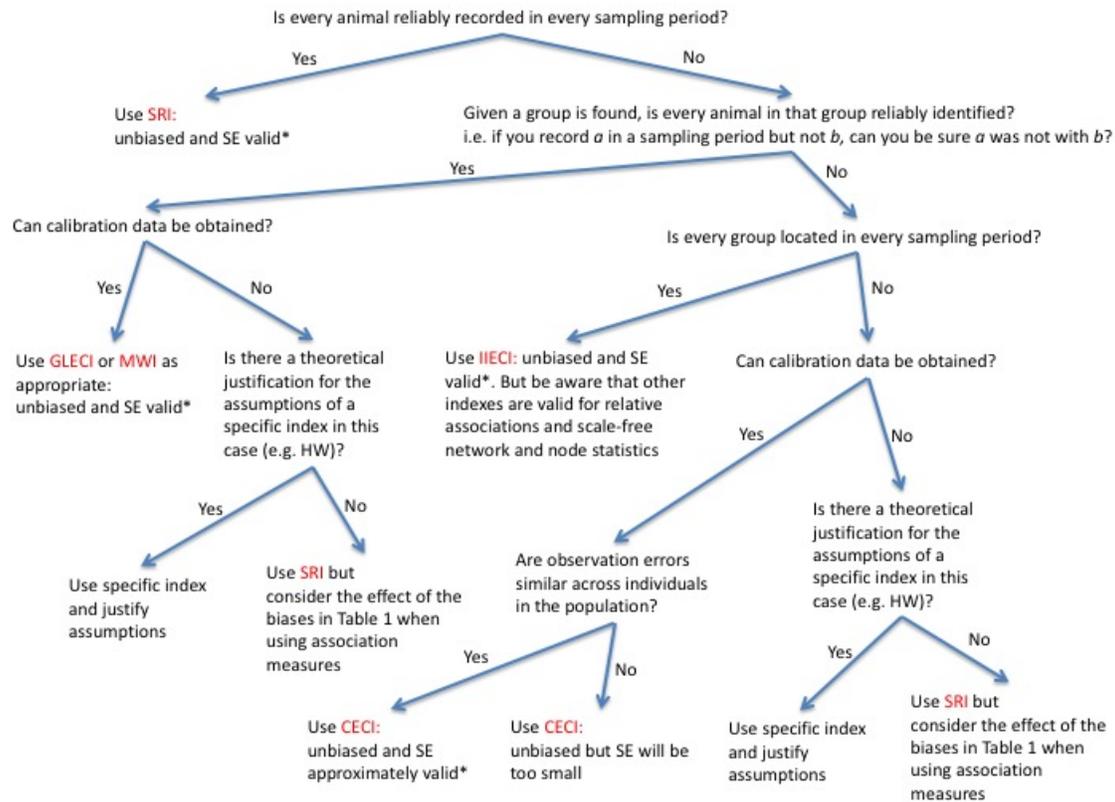
570 Here we suggest some initial ideas for obtaining calibration data that can be used to
571 estimate the calibration parameters derived above. These suggestions can almost
572 certainly be improved upon, by taking account of what data collection protocols are
573 feasible in specific circumstances perhaps by deriving maximum likelihood
574 estimates of calibration parameters given the data yielded by each such protocol.
575 Here we limit ourselves to providing relatively simple intuitive ways of estimating
576 calibration parameters. Whilst not optimal, our analysis above suggests that these
577 methods are nonetheless likely to be an improvement on the unsupported use of a
578 non-calibrated index such as the HWI.

579 **Table 3:** Summary of the indices considered in this paper. *Standard errors
 580 presented assume sampling periods are sufficiently spaced to be considered
 581 approximately independent, and that any calibration measures are known without
 582 error. See Supplementary Information for derivations of standard errors.

Index name	Formula for I_{ab}	Standard error*	Effective sample size (N_e)	Calibration measures required
Simple ratio index (SRI)	$x/(y_a + y_b + y_{ab} + x)$		$y_a + y_b + y_{ab} + x$	None
Half weight index (HWI)	$x/(\frac{1}{2}(y_a + y_b) + y_{ab} + x)$		$y_a + y_b + y_{ab} + x$	None
M weight index (MRI)	$x/(M(y_a + y_b) + y_{ab} + x)$		$y_a + y_b + y_{ab} + x$	M
Group location error corrected index (GLECI)	$\left(\frac{-x\omega + y_{Null}\omega - T_i}{-\sqrt{(x\omega + y_{Null}\omega - T_i)^2 + 4\omega T x}} \right) \sqrt{\frac{I_{ab}(1 - I_{ab})}{N_e}}$		$y_a + y_b + y_{ab} + x$ $+ y_{Null} \frac{\omega I_{ab}(1 - I_{ab})}{(\omega I_{ab} + 1)^2}$	ω
Very simple ratio index (vSRI)	$x/(y_{ab} + x)$		$y_{ab} + x$	None
Combined error corrected index (CECI)	Not available in closed form (see assocInd package for R code to generate this index and standard error)			ϵ, ω and ϕ

583

584



* Validity of SE assumes sampling periods are sufficiently spaced

585

586 **Figure 3:** Flowchart with our suggested strategy for selecting an association index.

587

588 One way a researcher might obtain estimates of calibration parameters is to collect
 589 data that can be assumed to be approximately error free for a subset of individuals,
 590 whilst simultaneously collecting association data using their standard protocol. This
 591 could be done by focal follows of a sample of individuals conducted by one
 592 researcher, whilst another collects data using the association data collection
 593 protocol. Alternatively some individuals could be tagged with GPS or proximity
 594 loggers (Kays, Crofoot, Jetz, & Wikelski, 2015; Krause et al., 2013) able to record
 595 encounters between individuals with more precision.

596

597 First we suggest that a researcher assess whether or not individual identification
 598 error is present and important. This could be done by calculating the proportion of
 599 sampling periods in which each individual included in the error free dataset was

600 present in a group that was located using the standard protocol but not recorded as
601 being present. If this proportion is 0 or close to 0 for most individuals, we suggest
602 individual identification error be ignored, and the GLECI or MWI can be used.
603 Otherwise, the CECI should be used (unless group location error is believed not to
604 be present, in which case the vSRI should be used, which does not require
605 calibration).

606

607 If individual identification error is not present or negligible, the researcher needs to
608 choose between the GLECI and MWI and then estimate the relevant calibration
609 parameter (ω or m). For any two individuals a and b in the error free sample, we
610 know in which sampling periods they were together and in which they were not
611 together. This enables the researcher to calculate $y_{null|ab}$, the number of times a and
612 b were not recorded by the association protocol during the calibration data
613 collection period, when a and b were known to be together. C , the probability a and
614 b will be missed when they are together, can be estimated as $y_{null|ab}/N_{ab}$ where N_{ab}
615 is the number of sampling periods that a and b were known to be together. A can
616 then be estimated as the proportion of sampling periods in which a was not
617 recorded by the association protocol and known not to be with b . B can be estimated
618 in an analogous manner.

619

620 This process can be repeated for every combination of two individuals in the error
621 free sample. The researcher can then use plots of these data to choose between the
622 GLECI and MWI. If the assumptions of the GLECI hold, we would expect C to have a
623 linear relationship with AB , with a slope of $(1 + \omega)$, whereas if the assumptions of
624 the MWI hold, we would expect C to have a linear relationship with $(A+B)$ with a
625 slope of $1/m$. We suggest the researcher make each plot to decide which assumption
626 is most realistic, and thus chose between the GLECI and MWI. If the GLECI is chosen,
627 they can then fit a linear regression (constrained to pass through the origin) of C
628 against AB and take $\omega = \text{slope}-1$. If the MWI is chosen, they can then fit a linear

629 regression (constrained to pass through the origin) of C against A+B and take m =
630 1/slope.

631

632 The CECI requires estimates of ω , ϕ and ϵ . ϵ can be estimated as the population
633 average of $\epsilon_{a|ab}$. To this end, we suggest for each dyad of individuals a and b , where
634 a is an individual for which we have error free data, we calculate $M_{a|ab}$ the number
635 of sampling periods in which the standard sampling protocol missed individual a
636 and we know (from the error free sample) that a and b were not together. We can
637 then estimate $\epsilon_{a|ab} = M_{a|ab}/N_{ab}$, where N_{ab} is the number of sampling periods
638 that a and b were known not to be together. ϵ can then be estimated as the average
639 of both $\epsilon_{a|ab}$ across all dyads containing at least one individual in the error free
640 dataset. Next the researcher can estimate ω , using the relationship $\epsilon_{ab|ab} =$
641 $(1 + \omega) \epsilon_{a|ab} \epsilon_{b|ab}$ in the same manner as suggested for estimating ω for the
642 GLECI above.

643

644 To estimate ϕ a researcher can use the relationship:

$$\epsilon_{a|ab} = \frac{(1 - \phi)(1 + \omega) \epsilon_{a|ab}^2 \epsilon_{b|ab}}{\phi(\epsilon_{a|ab} + \epsilon_{b|ab})}$$

645

646 For each dyad of individuals for which we have error-free data a researcher can
647 obtain an estimate of $\epsilon_{a|ab} = M_{a|ab}/N_{ab}$, where $M_{a|ab}$ is the number of sampling
648 periods in which the standard sampling protocol missed individual a and we know
649 (from the error free sample) that a and b were together, and where N_{ab} is the
650 number of sampling periods that a and b were known to be together. Using the
651 estimates of $\epsilon_{a|ab}$ and $\epsilon_{b|ab}$ obtained above, the researcher can then obtain an
652 estimate of $\epsilon_{a|ab}^2 \epsilon_{b|ab} / (\epsilon_{a|ab} + \epsilon_{b|ab})$. $\epsilon_{b|ab}$ can be likewise be estimated as
653 $M_{b|ab}/N_{ab}$, and $\epsilon_{b|ab}^2 \epsilon_{a|ab} / (\epsilon_{a|ab} + \epsilon_{b|ab})$ estimated as for individual a . The
654 researcher can then fit a linear regression (constrained to go through the origin)
655 with the set of $\epsilon_{a|ab}$ and $\epsilon_{b|ab}$ as the dependent variable and the set of
656 $\epsilon_{a|ab}^2 \epsilon_{b|ab} / (\epsilon_{a|ab} + \epsilon_{b|ab})$ and $\epsilon_{b|ab}^2 \epsilon_{a|ab} / (\epsilon_{a|ab} + \epsilon_{b|ab})$ as the independent

657 variable. Since the slope of the regression will estimate $(1 - \phi)(1 + \omega)/\phi$, we can
658 estimate $\phi = (1 + \omega)/(slope + 1 + \omega)$.

659

660 A second way we suggest a researcher might obtain estimates of calibration
661 parameters is to have two researchers or research teams independently collecting
662 association data using their standard protocol, for a portion of the data collection
663 period. For example, this might be done by having the second researcher (denoted
664 Y) collect data a short time after the first (denoted X), on a short enough time scale
665 that group composition is unlikely to have changed. Here we suggest a procedure
666 for obtaining calibration statistics for the CECI since this reduces down to the IIECI
667 or GLECI when the calibration data reveals the relevant component of error to be
668 absent.

669

670 We suggest that researchers first obtain estimates of $\epsilon_{a|ab}$. Ideally we wish to
671 estimate $M_{Xa|Y!ab}/N_{Y!ab}$, the proportion of events that X missed a given a and b were
672 together. We suggest researchers do this by calculating the proportion of sampling
673 periods X missed a given that Y recorded a and b in different groups (which we
674 denote $Y!ab$), i.e. $M_{Xa|Y!ab}/N_{Y!ab}$. We can repeat this procedure, reversing the role of
675 X and Y to obtain $M_{Ya|X!ab}/N_{X!ab}$. We can then take the mean as our estimate, i.e.
676 $\epsilon_{a|ab} = M_{Xa|Y!ab}/2N_{Y!ab} + M_{Ya|X!ab}/2N_{X!ab}$. ϵ can then be estimated as the
677 population average of $\epsilon_{a|ab}$ as above.

678

679 Researchers can then obtain estimates of $\epsilon_{a|ab}$, the probability that individual a is
680 missed when it is in a group with individual b . Again, we have potentially two
681 estimates of this for each combination of a and b . First we have $M_{Xa|Yab}/N_{Yab}$, the
682 proportion of sampling periods in which researcher X missed individual a given
683 researcher Y recorded a and b together, and conversely we have $M_{Ya|Xab}/N_{Xab}$. We
684 suggest $\epsilon_{a|ab}$ be estimated as the average of these, i.e. $\epsilon_{a|ab} = M_{Xa|Yab}/2N_{Yab} +$
685 $M_{Ya|Xab}/2N_{Xab}$. We suggest this be done for all combinations of a and b for which a
686 and b were frequently seen together, in order to obtain estimates of ω and ϕ . ω can

687 be estimated using the relationship $\epsilon_{ab|ab} = (1 + \omega) \epsilon_{a|ab} \epsilon_{b|ab}$ in the same
688 manner as suggested for estimating ω for the GLECI above. ϕ can be estimated as
689 above by fitting a linear regression (constrained to go through the origin) with the
690 set of $\epsilon_{a|ab}$ and $\epsilon_{b|ab}$ as the dependent variable and the set of
691 $\epsilon_{a|ab}^2 \epsilon_{b|ab} / (\epsilon_{a|ab} + \epsilon_{b|ab})$ and $\epsilon_{b|ab}^2 \epsilon_{a|ab} / (\epsilon_{a|ab} + \epsilon_{b|ab})$ as the independent
692 variable yielding the estimate $\phi = (1 + \omega) / (\text{slope} + 1 + \omega)$.

693

694 DISCUSSION

695

696 Studies of animal social networks have shed new light on many ecological and
697 evolutionary processes. For example, the structure of the social environment can
698 shape how information (Aplin, Farine, et al., 2015; Aplin, Farine, Morand-Ferron, &
699 Sheldon, 2012; Farine, Aplin, Sheldon, & Hoppitt, 2015) and diseases (K.L.
700 VanderWaal, Atwill, Isbell, & McCowan, 2013; K. L. VanderWaal et al., in press)
701 spread in wild populations. Further, they have provided important insights into the
702 role of the social environment on shaping selection (Farine & Sheldon, 2015;
703 Formica et al., 2011; McDonald, 2007; Oh & Badyaev, 2010; J.B. Silk, Alberts, &
704 Altmann, 2003; J. B. Silk et al., 2010; Wey, Burger, Ebensperger, & Hayes, 2013).
705 However, studies have used varying approaches to quantify the relationships among
706 individuals. Whilst care is generally taken to ensure that the chosen approach has
707 biological relevance, the underlying assumptions behind the approach used are
708 almost never explicitly considered. The results of our study into different
709 association indices suggests that many commonly-used approaches should be
710 avoided as they do not accurately estimate the (absolute or relative) strengths of
711 social bonds, which has implications on estimates of social structure and social
712 processes occurring through social networks.

713

714 Ideally, studies of animal social networks would capture information about all
715 individuals in the study population at once. Realistically, this is unlikely to be
716 possible in all but a very select number of studies. Thus, before constructing a social

717 network from a given set of data, we suggest that the following questions should be
718 addressed:

719

720 (1) How much data has been collected on each dyad?

721 (2) Are all individuals sampled equally?

722 (3) What proportion of the population is observed in each sample?

723 (4) Are there any mistakes in the observations?

724

725 The issues surrounding question 1 have now been relatively well outlined in the
726 literature (Farine & Strandburg-Peshkin, 2015; Franks et al., 2010; Lusseau,
727 Whitehead, & Gero, 2008; M. J. Silk et al., 2015; Whitehead, 2008). In general, these
728 studies have found that collecting enough data on each dyad (at least 20
729 observations per dyad) is important for accurately estimating global social network
730 structure. In the current study, we address issues arising from questions 2 and 3,
731 and how better indices can reduce the potential impact that missing observations
732 can have on both the absolute and relative estimates of association strengths among
733 individuals. Question 4 represents an area requiring some further investigation.

734

735 Our study suggests that a critical step in the study of animal social networks will be
736 the collection of calibration data. Currently-used association indices are all based on
737 arbitrary rates of missing observations. For example, the half-weight index assumes
738 that the probability of missing individuals a and b when they are together is exactly
739 half the probability of missing either individual when they are apart. Importantly,
740 we have shown that when this is not true, the HWI does not result in a 'better
741 approximation' of the real association rate when compared to the simple ratio index.
742 Thus, we recommend avoiding the use of the HWI, and instead using the SRI when
743 no calibration data is available (see Figure 3). In reality, it is likely that the rates of
744 observation could be estimated from parameters of the observation data, such as
745 the average group size, the average number of individuals observed in a sampling
746 period, and the average number of groups sampled. Whether these can be used to
747 parameterize the M-weighted index warrants further investigation.

748

749 Several extensions of association indices have been proposed to deal with other
750 issues arising when sampling populations. Godde, Humbert, Cote, Reale, and
751 Whitehead (2013) suggest a method to correct for the fact that individuals that
752 prefer large groups are more likely to be observed together. This involves
753 normalizing the association index values by the two individual's combined
754 gregariousness (the sum of their association indices to others). To deal with other
755 potentially confounding influences on association patterns (such as home-range
756 overlap) when attempting to estimate true association rates, Whitehead and James
757 (2015) propose regressing association indices against other input parameters. Our
758 proposed indices work equally well with both of these approaches as they are
759 simply new ways of defining the association value for pairs of individuals. The above
760 two studies highlight how patterns of affiliation can be affected by a range of
761 different factors. Thus, even if good calibration data can be obtained to estimate
762 accurate relationship strengths, it will always be important to use null models when
763 conducting hypothesis testing with animal social networks (Farine *in review*).

764

765 We encourage further investigation into methods for collecting informative
766 calibration data alongside the social network data. One potential avenue could be to
767 use mark-recapture techniques that explicitly investigate detection probabilities,
768 and these could be conditioned on having observed one or more particular
769 individuals. There are also increasing numbers of studies that are collecting
770 complete datasets from groups or populations of animals, and these could provide
771 very useful data for testing different approaches to collect calibration data. A
772 particular challenge that will arise is that social network analysis has proved
773 particularly useful in species or communities that exhibit fission-fusion dynamics
774 (Aureli et al., 2008; Couzin, 2006; M. J. Silk, Croft, Tregenza, & Bearhop, 2014). Here,
775 the rate of turnover in group membership can be very rapid (for example group
776 membership in great tits, *Parus major*, can be close to random after just 10 minutes,
777 Farine, Firth, et al., 2015). Similarly, when using focal observations—following a
778 single individual and recording its interactions with others—the concept of a

779 'group' is unclear, and we do not yet have a definition for out to estimate the group
780 location error. This is because all individuals can be observed, but only interactions
781 among a subset of edges (those connected to the focal) are recorded. Thus, in such
782 systems, more research is required to find robust approaches for collecting
783 calibration data.

784

785 In all of our simulations, and more generally in the assumptions of how well any
786 analysis captures reality, our estimates of accuracy are a 'best-case scenario'. In
787 reality, most datasets will also contain erroneous observations. At best, these are
788 simply individuals wrongly assigned into a group in which they did not occur. If
789 such individuals (and their erroneous associates) are observed many times, then the
790 resulting association strength will be low, and the error will be reasonably well
791 dealt with by using weighted social networks. However, the impact of incorrectly
792 assigning one individual identity for another could be significantly greater,
793 especially as such errors are unlikely to be randomly distributed throughout the
794 dataset (i.e. certain pairs of individuals are more likely to be confounded than
795 others). No association index will be able to correct for such errors. However, the
796 relative effect that incorrect assignments of identity have on different approaches to
797 estimate social network structure warrants further investigation.

798

799 Our paper makes it clear that we should avoid blindly using association indices
800 without proper consideration of the assumptions that they entail. We also
801 recommend discontinuing the use of the HWI, which Cairns and Schwager (1987)
802 already 30 years ago identified as being problematic because of the many
803 assumptions it makes. Instead, we recommend using properly calibrated association
804 indices wherever possible, and using the SRI if no appropriate calibration data or
805 estimates of rates of detectability are available. Whichever approach is used, we
806 hope that our paper will at least encourage researchers to carefully and explicitly
807 consider their choice of approach for estimating association strengths among
808 individuals in their study population.

809

810

811 ACKNOWLEDGEMENTS

812

813 DRF received additional support from the BBSRC (BB/L006081/1 to B.C. Sheldon).

814

815

816 REFERENCES

817

818 Aplin, L. M., Farine, D. R., Morand-Ferron, J., Cockburn, A., Thornton, A., & Sheldon, B.

819 C. (2015). Experimentally induced innovations lead to persistent culture via
820 conformity in wild birds. *Nature*, *518*, 538-541. doi:10.1038/nature13998

821 Aplin, L. M., Farine, D. R., Morand-Ferron, J., & Sheldon, B. C. (2012). Social networks
822 predict patch discovery in a wild population of songbirds. *Proceedings of the*
823 *Royal Society B*, *279*(1745), 4199-4205. doi:10.1098/rspb.2012.1591

824 Aplin, L. M., Firth, J. A., Farine, D. R., Voelkl, B., Crates, R. A., Culina, A., . . . Sheldon, B.
825 C. (2015). Consistent individual differences in the social phenotypes of wild
826 great tits (*Parus major*). *Animal Behaviour*, *108*, 117-127.
827 doi:10.1016/j.anbehav.2015.07.016

828 Aureli, F., Schaffner, C. M., Boesch, C., Bearder, S. K., Call, J., Chapman, C. A., . . . Van
829 Schaik, C. P. (2008). Fission-Fusion Dynamics New Research Frameworks.
830 *Current Anthropology*, *49*(4), 627-654. doi:10.1086/586708

831 Boogert, N. J., Farine, D. R., & Spencer, K. A. (2014). Developmental stress predicts
832 social network position. *Biology Letters*, *10*(10), 20140561.
833 doi:10.1098/rsbl.2014.0561

834 Cairns, S. J., & Schwager, S. J. (1987). A Comparison of Association Indexes. *Animal*
835 *Behaviour*, *35*, 1454-1469. doi:10.1016/S0003-3472(87)80018-0

836 Castles, M., Heinsohn, R., Marshall, H. H., Lee, A. E. G., Cowlshaw, G., & Carter, A. J.
837 (2014). Social networks created with different techniques are not
838 comparable. *Animal Behaviour*, *96*, 59-67.

839 Couzin, I. D. (2006). Behavioral ecology: Social organization in fission-fusion
840 societies. *Current Biology*, *16*(5), R169-R171. doi:10.1016/J.Cub.2006.02.042

841 Farine, D. R. (2015). Proximity as a proxy for interactions: issues of scale in social
842 network analysis. *Animal Behaviour*, *104*, e1-e5.
843 doi:10.1016/j.anbehav.2014.11.019

844 Farine, D. R., Aplin, L. M., Sheldon, B. C., & Hoppitt, W. (2015). Interspecific social
845 networks promote information transmission in wild songbirds. *Proceedings*
846 *of the Royal Society B*, *282*, 20142804. doi:10.1098/rspb.2014.2804

847 Farine, D. R., Firth, J. A., Aplin, L. M., Crates, R. A., Culina, A., Garroway, C. J., . . .
848 Sheldon, B. C. (2015). The role of social and ecological processes in
849 structuring animal populations: a case study from automated tracking of wild
850 birds. *Royal Society Open Science*, *2*, 150057. doi:10.1098/rsos.150057

- 851 Farine, D. R., & Sheldon, B. C. (2015). Selection for territory acquisition is modulated
852 by social network structure in a wild songbird. *Journal of Evolutionary*
853 *Biology*, 28(3), 547-556. doi:10.1111/jeb.12587
- 854 Farine, D. R., Spencer, K. A., & Boogert, N. J. (2015). Early-Life Stress Triggers
855 Juvenile Zebra Finches to Switch Social Learning Strategies. *Current Biology*,
856 25(16), 2184-2188. doi:10.1016/j.cub.2015.06.071
- 857 Farine, D. R., & Strandburg-Peshkin, A. (2015). Estimating uncertainty and reliability
858 of social network data using Bayesian inference. *Royal Society Open Science*,
859 2, 150367.
- 860 Farine, D. R., & Whitehead, H. (2015). Constructing, conducting, and interpreting
861 animal social network analysis. *Journal of Animal Ecology*, 84(5), 1144-1163.
862 doi:10.1111/1365-2656.12418
- 863 Formica, V. A., McGlothlin, J. W., Wood, C. W., Augat, M. E., Butterfield, R. E., Barnard,
864 M. E., & Brodie, E. D. (2011). Phenotypic Assortment Mediates the Effect of
865 Social Selection in a Wild Beetle Population. *Evolution*, 65(10), 2771-2781.
866 doi:10.1111/J.1558-5646.2011.01340.X
- 867 Franks, D. W., Ruxton, G. D., & James, R. (2010). Sampling animal association
868 networks with the gambit of the group. *Behavioral Ecology and Sociobiology*,
869 64(3), 493-503. doi:10.1007/s00265-009-0865-8
- 870 Ginsberg, J. R., & Young, T. P. (1992). Measuring Association between Individuals or
871 Groups in Behavioral-Studies. *Animal Behaviour*, 44(2), 377-379. doi:10.1016/0003-3472(92)90042-8
- 872 10.1016/0003-3472(92)90042-8
- 873 Godde, S., Humbert, L., Cote, S. D., Reale, D., & Whitehead, H. (2013). Correcting for
874 the impact of gregariousness in social network analyses. *Animal Behaviour*,
875 85(3), 553-558.
- 876 Granovetter, M. (1973). The Strength of Weak Ties. *American Journal of Sociology*,
877 78(6), 1360-1380. doi:10.1086/225469
- 878 Kays, R., Crofoot, M. C., Jetz, W., & Wikelski, M. (2015). Terrestrial animal tracking as
879 an eye on life and planet. *Science*, 348(6240), aaa2478.
- 880 Krause, J., Krause, S., Arlinghaus, R., Psorakis, I., Roberts, S., & Rutz, C. (2013). Reality
881 mining of animal social systems. *Trends in Ecology & Evolution*, 28(9), 541-
882 551. doi:10.1016/j.tree.2013.06.002
- 883 Lusseau, D., Whitehead, H., & Gero, S. (2008). Incorporating uncertainty into the
884 study of animal social networks. *Animal Behaviour*, 75, 1809-1815.
885 doi:10.1016/J.Anbehav.2007.10.029
- 886 McDonald, D. B. (2007). Predicting fate from early connectivity in a social network.
887 *Proceedings of the National Academy of Sciences of the United States of*
888 *America*, 104(26), 10910-10914. doi:10.1073/Pnas.0701159104
- 889 Oh, K. P., & Badyaev, A. V. (2010). Structure of Social Networks in a Passerine Bird:
890 Consequences for Sexual Selection and the Evolution of Mating Strategies.
891 *American Naturalist*, 176(3), E80-E89. doi:10.1086/655216
- 892 R Development Core Team. (2015). R: A Language and Environment for Statistical
893 Computing (Version R Foundation for Statistical Computing). Vienna, Austria:
894 R Foundation for Statistical Computing. Retrieved from [http://www.R-](http://www.R-project.org)
895 [project.org](http://www.R-project.org)

- 896 Silk, J. B., Alberts, S. C., & Altmann, J. (2003). Social bonds of female baboons enhance
897 infant survival. *Science*, *302*, 1231-1234. doi:10.1126/science.1088580
- 898 Silk, J. B., Beehner, J. C., Bergman, T. J., Crockford, C., Engh, A. L., Moscovice, L. R., ...
899 Cheney, D. L. (2010). Strong and Consistent Social Bonds Enhance the
900 Longevity of Female Baboons. *Current Biology*, *20*(15), 1359-1361.
901 doi:10.1016/J.Cub.2010.05.067
- 902 Silk, M. J., Croft, D. P., Tregenza, T., & Bearhop, S. (2014). The importance of fission-
903 fusion social group dynamics in birds. *Ibis*, *156*, 701-715.
904 doi:10.1111/ibi.12191
- 905 Silk, M. J., Jackson, A. L., Croft, D. P., Colhoun, K., & Bearhop, S. (2015). The
906 consequences of unidentifiable individuals for the analysis of an animal
907 social network. *Animal Behaviour*, *104*, 1-11.
908 doi:10.1016/j.anbehav.2015.03.005
- 909 Soetaert, K., & Herman, P. M. J. (2009). *A Practical Guide to Ecological Modelling.*
910 *Using R as a Simulation Platform*: Springer.
- 911 Strandburg-Peshkin, A., Farine, D. R., Couzin, I. D., & Crofoot, M. C. (2015). Shared
912 decision-making drives collective movement in wild baboons. *Science*,
913 *348*(6241), 1358-1361.
- 914 VanderWaal, K. L., Atwill, E. R., Isbell, L. A., & McCowan, B. (2013). Linking social and
915 pathogen transmission networks using microbial genetics in giraffe (*Giraffa*
916 *camelopardalis*). *Journal of Animal Ecology*, *83*(2), 406-414.
917 doi:10.1111/1365-2656.12137
- 918 VanderWaal, K. L., Obanda, V., Omondi, G. P., McCowan, B., Wang, H., Fushing, H., &
919 Isbell, L. A. (in press). The “strength of weak ties” and helminth parasitism in
920 giraffe social networks. *Behavioural Ecology*.
- 921 Wey, T. W., Burger, J. R., Ebensperger, L. A., & Hayes, L. D. (2013). Reproductive
922 correlates of social network variation in plurally breeding degus (*Octodon*
923 *degus*). *Animal Behaviour*, *85*(6), 1407-1414.
924 doi:10.1016/j.anbehav.2013.03.035
- 925 Whitehead, H. (2008). *Analyzing animal societies*. Chicago, USA: University of
926 Chicago Press.
- 927 Whitehead, H., & Dufault, S. (1999). Techniques for analyzing vertebrate social
928 structure using identified individuals: Review and recommendations.
929 *Advances in the Study of Behavior*, *Vol. 28*, *28*, 33-74.
- 930 Whitehead, H., & James, R. (2015). Generalized affiliation indices extract affiliations
931 from social network data. *Methods in Ecology and Evolution*, *6*(7), 836-844.
932 doi:10.1111/2041-210X.12383
- 933 Wilson, A. D. M., Krause, S., Dingemanse, N. J., & Krause, J. (2013). Network position:
934 a key component in the characterization of social personality types.
935 *Behavioral Ecology and Sociobiology*, *67*(1), 163-173.
936
937