Social structure as a strategy to mitigate the costs of group living: a comparison of gelada and guereza monkeys

R. I. M. Dunbar a, b, * 

a Department of Experimental Psychology, University of Oxford, Oxford, U.K.
b Department of Computer Science, Aalto University, Aalto, Finland

A key innovation in some of the more intensely social mammals, including anthropoid primates, elephants, equids and delphinids, is the capacity to form intense, typically lifelong, bonded relationships (sensu Silk, 2002; see also Curley & Keverne, 2005; Dunbar, 2001; Dunbar & Shultz, 2010; Massen, Sterck, & de Vos, 2010; Silk, Altmann, & Alberts, 2006; Silk et al., 2010a,b; Smuts, 1985). In primates, these relationships are invariably based on regular mutual grooming. Besides creating structure within social groups (primates: Kudo & Dunbar 2001; Lehmann & Dunbar, 2009; elephants: Wittenmyer, Douglas-Hamilton, & Getz, 2005; orcas: Hill, Bentley, & Dunbar, 2008), these relationships commonly provide the basis for recruiting support against third party aggression and harassment (Dunbar, 2012; Dunbar & Shultz, 2007, 2017; Harcourt, 1992), but they also create stress lines where fission will occur (baboons: King, Douglas, Huchard, Isaac, & Cowlishaw, 2008; elephants: Archie, Moss, & Alberts, 2006). Harcourt (1992) pointed out that primate coalitions are unique in that they are established long before they are needed; in contrast, the kinds of coalitions found in most other birds and mammals are typically ‘of the moment’ (i.e. due to a momentary convergence of interests), short lived and less likely to involve risky altercations. In fact, it is likely that the same claim can be made for alliances in some of the more intensely social nonprimate mammals (elephants: Archie et al., 2006; dolphins: Ford, Ellis, & Balcomb, 2000; Lusseau et al., 2003; bats: Kerth, Perony, & Schweitzer, 2011; see also Shultz & Dunbar, 2010a). As a result, the distinction is now often drawn between coalitions (casual support offered during specific altercations) and alliances (long-term relationships that form the basis of regular coalitionary support).

Long-lasting social relationships of the kind found in primates and some other mammals have demonstrable fitness implications for females: grooming partners are more likely to respond to each other’s distress and alarm calls (primates: Seyfarth & Cheney, 1984; Micheletta et al., 2012), and individuals who are socially well

* Correspondence: R. I. M. Dunbar, Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, U.K.
E-mail address: robin.dunbar@psy.ox.ac.uk.
embedded and/or have more grooming partners are harassed less often, have lower cortisol levels in response to stressful events, and greater longevity and reproductive success (primates: Archie, Tung, Clark, Altmann, & Alberts, 2014; Crockford, Wittig, Whitten, Seyfarth, & Cheney, 2008; Dunbar, 1984; Engh et al., 2006; Silk, Altmann, & Altman, 2003; Silk et al., 2009, 2010a,b; Wittig et al., 2008; equids: Cameron, Setsaas, & Linklater, 2009; dolphins: Frère et al., 2010).

These close ‘friendships’ (Silk, 2002), mediated via social grooming, inevitably give rise to structured social networks (primates: Kudo & Dunbar, 2001; Lehmann & Dunbar, 2009; other mammals: Hill et al., 2008; Wittemeyer et al., 2005) that resemble the layered networks found in humans (Arnaboldi, Passarella, Conti, & Dunbar, 2015; Dunbar, Arnaboldi, Conti, & Passarella, 2015; MacCarron, Kaski, & Dunbar, 2016; Sutcliffe, Dunbar, Binder, & Arrow, 2012). However, not all anthropoid primates exhibit this kind of structuring. While cercopithecine monkeys live in groups that typically consist of a number of small, self-contained cliques whose members mainly confine their interactions to themselves (Dunbar, 1983a; Kudo & Dunbar, 2001; Lehmann & Dunbar, 2009), other Old and New World monkeys such as colobines and many cebids typically have less highly differentiated relationships and typically exhibit less complex forms of social behaviour and communication (Byrne & Corp, 2004; Dobson, 2009, 2012; McComb & Semple, 2005).

Despite these benefits from sociality, living in groups incurs costs in a number of different respects, including competition for access to resources (van Schaik, 1989), additional travel and social time demands and disrupted time budgets (Dunbar, Korstjens, & Lehmann, 2009), all of which impose physiological stress on the animal. Irrespective of whether resources are involved or not, animals that live in groups also experience constant low to moderate levels of harassment from other group members. This can range from casual, even accidental, displacements during foraging, to direct theft of resources that an individual has just harvested, to outright attack (sometimes for no immediate reason, as when one individual persistently targets another). In addition, aggression involving third parties inevitably spills over to affect others, even though they may not be directly involved, thereby giving rise to psychological stress. Because of this, low-ranking animals often visually check the location of dominant individuals (Alberts, 1994; Caine & Marra, 1988; Hirsch, 2002).

The reproductive endocrinology of female mammals is extremely sensitive to stress. In primates, females subject to even modest levels of social stress experience reduced levels of fertility (rodents: Geraghty et al., 2015; ungulates: von Borell, Dobson, & Prunier, 2007; Einarsson, Brandt, Lundheim, & Madej, 2008; primates: Abbott, 1984, 1987; Abbott, Keverne, Moore, & Yodvingyud, 1984, 1986; Bowman, Dilley, & Keverne, 1978; Rowell, 1970; humans: An, Sun, Li, Zhang, & Ji, 2013; Buch Louis et al., 2011; Caldwell, Immerwahr, & Ruzicka, 1982; Ji et al., 2013; Laatikainen, 1991; Pettay, Lahdenperä, Rotkirch, & Lummaa, 2016; Sanders & Bruce, 1997; Schenker, Meirow, & Schenker, 1992; Schliep et al., 2015; Seifer & Collins, 1990; Zacur, Chapans, Lake, Ziegler, & Tyson, 1976), in some cases even complete reproductive suppression (callitrichid primates: Abbott, Hedges, & George, 1988; Abbott, McNeilly, Lunn, Hulme, & Burden, 1981; Abbott et al., 1984; many canids: Moehlman & Hofer, 1997; mole-rats: Bennett, 1994; Faulkes, Abbott, & Jarvis, 1990). In some species, stress is even known to cause pregnant females to abort (sheep: Howarth & Hawk, 1968). The mechanism is now well understood, and mainly involves the way prolactin and endorphins, triggered by social or physical stress, inhibit the production of gonadotrophin releasing hormone, thereby blocking the luteinizing hormone surge needed to precipitate ovulation (Abbott, 1988; Gordon, Hodgen, & Richardon, 1992; Guidotti & Grandison, 1979; Margulis, Altmann, & Ober, 1993; McNeilly, 2001a,b; McNeilly, Forsyth, & McNeilly, 1994; Ziegler, Widowski, Larson, & Snowdon, 1990). In rats, even the stress induced by being physically constrained for short periods can be sufficient to block the cyclic release of gonadotrophins necessary for ovulation (Euker & Riegel, 1973).


Ecologists invariably assume that any within-group fertility differences are due to differences in food intake arising from scramble competition (e.g. Koenig, 2000). In fact, starvation itself triggers the endorphin system (e.g. as a response to hunger pangs or the physiological stress that hunger causes), thereby precipitating infertility in mammals (sheep: Dobson, Ferengan, Routly, & Smith, 2012; pigs: Clarke, 2014). While lack of food can certainly cause the reproductive system to shut down, this usually happens only in cases involving significant weight loss (>15%) or excessive exercise (e.g. athletes) (primates: Williams, Berga, & Cameron, 2007; humans: Boyar et al., 1974; Bullen et al., 1985; Dean, 1949; Ellison 1990; Frisch, 1987; Frisch & McArthur, 1974; Howlett et al., 1984; Kirchengast & Huber, 2001; McClure, 1968; Nattiv et al., 2007; Smith, 1947; Warren & Perillo, 2001). Even then, it is the hypothalamic pathway that regulates this, rather than nutrition per se (Kalra & Kalra, 1996; Schwartz & Seeley, 1997). In other words, the same endorphin/HPA pathway is involved in both social and ecological routes, perhaps explaining why the effects of one can easily be mistaken for the other. So even though nutrient constraints will usually be the major factor influencing gestation and lactation, the loss of fertility under adverse ecological conditions could in fact be due to the stress rather than the shortage of nutrients per se.

Irrespective of whether social or ecological stress is the ultimate cause, the fact that fertility declines with group size has significant implications for female fitness for animals that live in social groups. There will be intense selection pressure on females that are doing poorly in terms of fertility to leave in order to join smaller groups (Dunbar et al., 2008; van Horn, Buchan, Altmann, & Alberts, 2007), with an inevitable downward pressure on average group size. In effect, fertility should act as a limit on the evolution of group size in mammals. If groups are to increase in size beyond this limit in response to environmental demands such as predation risk, some behavioural mechanism will be needed to buffer females against these stresses. Failure to find solutions that allow larger numbers of individuals to live together will inevitably prevent species from occupying high-risk terrestrial or open country habitats where large groups are necessary to provide a defence against predators (Bettridge & Dunbar, 2012; Bettridge, Lehmann, and Dunbar, 2010; Dunbar et al., 2009; van Schaik, 1983; Shultz, Noe, McGraw, & Dunbar, 2004).
In this paper, I explore the possibility that some species have been able to use behavioural strategies to mitigate the fertility costs of group living in order to be able to live in larger groups. To do this, I compare network structure and fertility patterns across group size for two species of Old World monkeys, black-and-white colobus, or guerezas, *Colobus guereza*, and gelada, *Theropithecus gelada* (a close relative of the *Papio* baboons), that represent opposite extremes of the social complexity dimension.

These two species make a particularly appropriate comparison for two reasons. First, while network analysis methods have proved extremely illuminating in the study of animal behaviour, many network indices correlate negatively with group size, making it notoriously difficult to compare indices for species that live in groups of radically different size (Croft, James, & Krause, 2008; James, Croft, & Krause, 2009). Unusually, both guereza and gelada live in one-male (very occasionally multimale) reproductive units of similar size, thereby allowing a direct comparison across groups of similar size, something that would normally be difficult because most of the species that regularly form alliances ( cercopithecines, great apes, equids) live in groups that are significantly larger than those characteristic of species, such as colobines, that do not regularly form alliances (Campbell, Fuentes, Mackinnon, Panter-Brick, & Bearder, 2007; Shultz, Fitzpatrick, Dunbar, Hughes, & Atkinson, 2017). Gelada differ from guereza in that their reproductive units are embedded within higher level groupings (teams, bands), as well as themselves being highly substructured (Dunbar, 1984; Kawai, Dunbar, Ohsawa, & Mori, 1983; MacCarron & Dunbar, 2016); in contrast, guereza groups occupy individual territories with less obvious substructuring and little evidence for higher level groupings (Fashing, 2007). In addition, and perhaps uniquely, in each case data are available on the network structure of a large sample of groups from the same location, allowing us to standardize for any effect that environmental conditions might have on available social time and other aspects of social behaviour.

The question at issue is whether the two species differ in the way their social networks are structured, and whether this difference is correlated with both fertility rates and patterns of group fission. The species that makes more use of alliances should be able to prevent its groups fragmenting and undergoing fission over a wider range of sizes, and this should be reflected in fertility rates that are less strongly affected by group size such that the negative slope is either shallower or shifted towards larger groups. I first compare network structure between the two species to determine how they differ, and what limits this might place on group size; I then present data on agonistic rates, alliance formation and group fission rates to provide mechanistic details; finally, I examine data on fertility in order to ask (1) whether fertility is negatively related to group size, and if so (2) whether the differences in alliance formation and social structuring are in principle sufficient to explain the differences in fertility patterning.

**METHODS**

The gelada were studied at Sankaber, Simen Mountains National Park (Ethiopia), in 1971–1972 and 1974–1975, in each case for ca. 9 months. There were six bands at Sankaber, but the principal focus of study was the Main band (comprising 25 reproductive units, or harems, in 1971 and 17 harems in 1974) and the Abyss band (six harems) during 1971–1972 before it completely fissioned from the Main band (Dunbar, 1984; MacCarron & Dunbar, 2016). The guereza were studied in the Bole Valley, Ethiopia, at intervals between May 1971 and September 1974, with the main focus on two continuous 6-month periods in 1972 and 1974; a total of nine groups were censused and studied on a regular basis (Dunbar, 1988).

The data on gelada grooming patterns are based on samples of social grooming recorded during focal follows for 13 individual reproductive units (harems) sampled in 1974–1975 (Dunbar, 1984). On average, each unit was sampled for 14.7 h. The data for guereza derive from eight groups, and are based on grooming frequencies recorded during focal group follows in 1972 and 1974 (Dunbar & Dunbar, 1976; Dunbar, 1988). On average, each group was sampled for 7.0 h. All data analysed here are for reproductively mature adults only. I include males as well as females because some gelada females use males as alliance partners when they do not have female allies available (Dunbar, 1984).

Following Lehmann et al. (2009), grooming data were analysed in UciNET (Borgatti, Everett, & Freeman, 2002) to calculate three key unweighted, undirected network indices (density, 2-clan and betweenness centrality) for each individual social group. Density is the proportion of all possible first-order ties (dyadic grooming partnerships) that occur, and indexes how interconnected the animals are. The 2-clan index specifies the number of separate grooming cliques (such that every member can be reached directly or indirectly in a maximum of two tie steps: A grooms with B who grooms with C) into which the group can be divided; in effect, it indexes how many weakly connected subnetworks a group is divided into (Humphries, Gurney, & Breakwell, 2015). Betweenness centrality is a measure of the extent to which an individual is linked, directly or indirectly, to every other individual in the group. It is equal to the average number of shortest paths from one animal to every other member of the group that passes through a given individual; a high betweenness index implies that a particular individual is pivotal in maintaining the structural cohesion of the network, but averaged across group members it indicates how well connected the group is.

Data on the frequency of within-group agonistic interactions derive from these same focal-group samples (*N* = 13 gelada harems and *N* = 8 guereza groups), based on ad libitum sampling.

Although data on actual birth rates are available for both study populations, the sample for the Bole guereza (nine groups, comprising 18 reproductive females) is subject to large small-sample fluctuations (for further discussion, see Appendix). Instead, for both species, I use the number of infants aged 0–12 months in each group at the end-point of each study period as an index of fertility. The gelada data derive from the Main band at Sankaber, comprising 42 reproductive units (175 reproductive-age females) over the two study periods, thus providing a sample large enough to outweigh small-sample effects. To provide a sufficiently large sample for the guereza, I used data on number of infants (white and black infants, aged 0–12 months old) derived from 52 groups (139 reproductive-age females) at five study sites in East Africa: nine groups from the main study population in the Bole Valley and six at Lake Shalla in Ethiopia (Dunbar, 1988), 25 groups from the Budongo Forest (Suzuki, 1979) and seven at Kanyawara (Kibale Forest) (Oates, 1977) in Uganda, and five from Kakamega Forest, Kenya (Fashing, 2001). All except Lake Shalla involved studies in which group demography was monitored over an extended period; the Lake Shalla study was a complete census in very open, scrubby gallery forest where groups could be followed and censused easily. Detailed analyses of fertility for both species are given in the Appendix, and the data are given in the Supplementary Material.

Comparisons of network index and other data use standard ANOVAs or *t* tests. Regressions set through data are conventional least squares regressions. Statistical comparisons of intercept and slope parameters for the two species’ regressions follow Pollard (2009). Alpha was set at *P* = 0.05. One-tailed tests are used where a directional hypothesis is being tested; these are specified in the text. All statistical analyses were carried out in SPSS v.22.
No ethical issues arise from this study as the data on which it is based come from observational field studies.

RESULTS

Network Structure

Fig. 1 plots the respective values for the three network indices as a function of group size (indexed as number of adults) for each species. The linear regressions for each species are all significant, except for guereza betweenness centrality (Table 1). Parameter comparisons indicate that both intercept and slope differ significantly between the two species for betweenness centrality (note that the regressions for betweenness centrality in the two species are of opposite slope); the slopes are significantly different for density, and the intercepts are significantly different for the 2-clan index (Table 1).

The residuals from the common least squares regression for density are significantly more positive for gelada than for guereza (Fig. 1a: \(F_{1,21} = 4.980, P = 0.04\)), indicating that gelada have consistently higher densities (i.e. more first-order grooming contacts) than guereza for groups of a given size. The relationship seems to be strictly linear in each case; nonlinear equations yield worse fits (indexed as \(r^2\)). It is possible that the extreme right-hand point for gelada may disproportionately influence the slope: removing this data point from the analysis yields regression equations for the two species that do not differ in slope (\(t_{21} = 0.94, P = 0.355\)) but do differ significantly in intercept (\(t_{21} = 2.07, P = 0.047\)), as before. For betweenness centrality, removing the extreme data point in each sample does not change the significances of the slopes (Fig. 1c; linear models: guereza: \(F_{1,8} = 0.89, P = 0.374\); gelada, \(F_{1,21} = 6.03, P = 0.030\)).

Four of the gelada harems underwent changes in composition (three cases in which an adult female died and one case of an influx of four adult and subadult males associated with a harem takeover; see Dunbar, 1984), with grooming data available for sample periods before and after the change. Fig. 2 shows the changes in the three network indices for each case. To show how the number of grooming partners remains constant, I plot average degree (number of first-order ties for each group member) rather than density (the proportion of all possible dyads realized, as in Fig. 1a). Density declines across group size (Fig. 1a), and this is because the number of first-order ties remains stable at ca. 1.7 across both groups and phases (Fig. 2).

The data in Fig. 2 suggest that there is a phase transition in the number of 2-clans from 2c = 1 \(\rightarrow\) 2 or more at a harem size of seven adults, which fits well with the evidence in Fig. 1b suggesting that this transition point lies somewhere in the interval \(6 \leq N \leq 9\). Betweenness centrality exhibits the same pattern as in Fig. 1c, with centrality increasing with group size in gelada; this again suggests that animals make efforts to distribute their social effort more widely as group size increases in such a way as to maintain connectivity (thereby, presumably, reducing the risk of fission). However, this does not mean that they were actively grooming with everyone in the group: Figs 1a and 2 suggest that the number of primary grooming partners does not change that much.

Coalitionary Support

Hourly rates of within-group agonistic events (threats, active and passive displacements, attacks, fights) are plotted for small and large groups for the two species in Fig. 3, representing the net experience of altercations for the animals in each group whether or not they were actually involved in an incident. Subadults are included in this analysis because they are often involved in agonistic events and add measurably to the level of stress created by such events. While rates are just significantly higher in larger groups, as one would expect (\(F_{1,16} = 3.40, P = 0.042\) one-tailed), the difference between species is at best marginal (\(F_{1,16} = 4.01, P = 0.025\) one-tailed).

\[F_{2,16} = 3.40, P = 0.042\]

\[F_{2,16} = 4.01, P = 0.025\]

Figure 1. (a) Density, (b) number of 2-clans and (c) average betweenness centrality as a function of group size for individual groups of guereza (open symbols) and gelada (filled symbols). In each case, all the groups for a species were sampled from the same population. Lines are least square regressions (guereza: dashed lines; gelada: solid lines).
of both the severity of the altercation and the proportion of the
individual is threatened by a neighbouring harem is a function
that female will go to the aid of another adult female of her harem when
especially evident when females become involved in altercations
with neighbouring harems, which occur fairly frequently in the

Changes in network indices following changes in size for four gelada harems
Figure 2. Changes in network indices following changes in size for four gelada harems
(identiﬁed by the N-sufﬁx numbers on the X-axis). The data are organized on the X-axis
in order of the number of adults in the harem for each sample, rather than the
temporal direction of the change (which is indicated by the arrow between N values in
the ﬁgure). The changes in N2, N12 and N17 involved deaths (all adult females), that in
N5 an inﬂux of four adult males (an event that can be expected to give rise to harem
fission sometime within the following 6–12 months). The number of reproductive
females in each harem were: N2 = 4 (dropping to 3), N17 = 5 (dropping to 4), N12 = 7
unchanged), N5 = 9 (unchanged). Group size increased across the change in N12
because of the maturation of juveniles. Filled circles: mean number of ﬁrst-order ties
(degree); open circles: number of 2-clans; triangles: mean betweenness centrality.
‘pre’ and ‘post’ refer to the period before and after the change in harem size.

Mean (±1 SE) number/h of agonistic events for small (open symbol) and
large (ﬁlled symbol) groups of guereza and gelada. An agonistic event involves threats,
active or passive displacements, outright attack and ﬁghts (although the latter were
rare). For guereza, small groups have three to ﬁve adults and subadults, and large
groups have six to nine; for gelada, small groups have four to six adults and subadults,
and large groups have seven to 12. Subadults are included as they are often involved
in agonistic interactions. The data are based on focal group sampling of ﬁve guereza
groups in the Bole Valley in 1974 and 12 gelada harems at Sankaber (Simien Mountains)
in 1974–1975. The mean sample per group was 8.4 h for guereza and 15.9 h for gelada,
with a minimum sample set at 5 h in both cases.

Table 1
Linear regressions for network indices on group size

<table>
<thead>
<tr>
<th>Index</th>
<th>Species</th>
<th>Equation</th>
<th>Standardized β</th>
<th>t*</th>
<th>P</th>
<th>Equation differencesa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Density</td>
<td>Guereza</td>
<td>$D = -1.259 - 0.098N$</td>
<td>-0.704</td>
<td>-2.97</td>
<td>0.016</td>
<td>Intercept: $t = 0.45$, $P = 0.656$</td>
</tr>
<tr>
<td></td>
<td>Gelada</td>
<td>$D = 1.005 - 0.046N$</td>
<td>0.644</td>
<td>-3.95</td>
<td>0.001</td>
<td>Slope: $t = -2.31$, $P = 0.028$</td>
</tr>
<tr>
<td>2-clan</td>
<td>Guereza</td>
<td>$2C = -0.022 + 0.290N$</td>
<td>0.765</td>
<td>3.60</td>
<td>0.006</td>
<td>Intercept: $t = 3.1$, $P = 0.004$</td>
</tr>
<tr>
<td></td>
<td>Gelada</td>
<td>$2C = -1.262 + 0.399N$</td>
<td>0.881</td>
<td>8.70</td>
<td>&lt;0.001</td>
<td>Slope: $t = 1.8$, $P = 0.082$</td>
</tr>
<tr>
<td>Betweenness centrality</td>
<td>Guereza</td>
<td>$C = 50.394 - 3.848N$</td>
<td>-0.208</td>
<td>-0.54</td>
<td>0.591</td>
<td>Intercept: $t = -5.18$, $P &lt; 0.0001$</td>
</tr>
<tr>
<td></td>
<td>Gelada</td>
<td>$C = -10.821 + 2.721N$</td>
<td>0.874</td>
<td>6.24</td>
<td>&lt;0.001</td>
<td>Slope: $t = 3.81$, $P = 0.001$</td>
</tr>
</tbody>
</table>

a Signiﬁcance of regression slope: $df = 9$ for guereza and $df = 22$ for gelada.
b Difference between intercepts and slopes for the two species’ equations: $df = 2$.

Fertility

Fig. 5 plots the regressions for the mean number of infants per female against number of adults in the group for the two species
The gelada is displaced to the right by a significant amount. Although the graphs are very similar in shape, that for guereza is maximized at two to three reproductive females, and declines rapidly (Fig. 1c). In practice, however, it makes little difference: the main point is that guereza fertility is maximized at two to three reproductive females, and declines rapidly (Fig. 1c). Whereas the equivalent for gelada is ca. nine females. In effect, harem fertility falls to zero in groups with ca. six females, whereas the equivalent for gelada is ca. nine females. In effect, gelada are able to defer the effect of social stress on fertility by a significant amount. For gelada, the curve is identical to that obtained if we use actual birth rates, suggesting that our measure is a reasonable approximation for true fertility. For both species, number of infants has an inverted U-shaped relationship with the number of adults in the group. The relationship for guereza is independent of the environmental differences between the five habitats involved; for gelada, the curve is identical but higher in the more ecologically benign genus (see Appendix). Although the graphs are very similar in shape, that for the gelada is displaced to the right by a significant amount. For obvious reasons, we cannot be sure that the guereza curve continues to decline on the lefthand side, and hence it is possible that their curve is a linear relationship. In practice, however, it makes little difference: the main point is that guereza fertility is maximized at two to three reproductive females, and declines rapidly thereafter, whereas gelada fertility reaches its maximum at ca. five reproductive females before declining at approximately the same rate. In guereza, fertility falls to zero in groups with ca. six females, whereas the equivalent for gelada is ca. nine females. In effect, gelada are able to defer the effect of social stress on fertility by a significant factor.

Group Fission

Of a total of 46 gelada harems from three well-studied Sankaber bands (Main and Abyss bands in 1971–1972 and Main band in 1974–1975), four harems underwent fission (when they had seven, eight and nine females, respectively: mean = 7.75), for an average fission risk of 0.11 per harem per year; in addition, two females (out of 211) moved between harems (one only very briefly). Three of nine guereza groups at Bole underwent fission over the 41-month sample period (when they had four, four and five females, respectively: mean = 4.32), for an average fission risk of 0.096 per year; in addition, a female moved from one group to its daughter group after fission had occurred. Fig. 5 suggests that, for both species, the mean size at fission approximates very closely the point at which fertility drops below demographic replacement, calculated as two surviving offspring in a lifetime (i.e. four live births in a 12–15-year reproductive life span, allowing for ca. 50% mortality to puberty, as is typical at least for baboons: Altmann, Altmann, Hausfater, & McCuskey, 1977; Sigg, Stolba, Abegglen, & Dasser, 1982).

Note that, in gelada harems, structural fragmentation (as indexed by the phase transition from one to two 2-clans) seems to occur once there are seven adults in the group (Fig. 1c), whereas actual fission occurs when there are ca. 10 adults (ca. eight females plus the necessary two adult males; Fig. 5). In other words, harems start to show structural fragmentation before they undergo fission, suggesting that the first may be necessary for, or a prelude to, the second (in addition, of course, to the presence of a second male: Dunbar, 1984).

DISCUSSION

These two catarrhine species differ in the size of their grooming cliques and the extent to which their groups are sub-structured, in the way coalitionary alliances are used to buffer females during altercations, and in their patterns of fertility with respect to social group size. For both species, the fertility curves appear to be independent of environmental (or at least climatic) effects otherwise known to influence fertility. Papio baboons also exhibit an inverted U-shaped relationship between birth rates and number of females (Dunbar, MacCarron, & Robertson, 2017; Hill et al., 2000), and this holds even when adjusting for habitat quality, suggesting that this pattern may be widespread at least among the catarrhine primates. Although the sample of fissions is small, the data for both species suggest that group fission is increasingly likely as fertility rates approach the demographic replacement rate, but that this point occurs at very different group sizes in the two species.

Gelada are able to maintain more densely interconnected groups than guereza as group size increases (Fig. 1b and c); as a result, their groups are less likely to become fragmented into distinct subgroups that do not interact at smaller group sizes, and hence are less likely to fission, and this is despite the fact that their groups are more likely to be subdivided into distinct grooming cliques. Since rates of aggression do not differ significantly between the two species, it is likely that the effect of grooming-based alliances is to provide a greater sense of security for females during potentially stressful events, hence reducing their stress levels (Crockford et al., 2008; Wittig et al., 2008). While coalitionary support involving grooming-based alliances are relatively common among the gelada, none were observed in the guerezas. Only one study of any Colobus species has reported coalitionary behaviour (Sicotte, Teichroeb, & Saj, 2007; Teichroeb & Sicotte, 2008; Wikberg et al., 2014); however, in this case, coalitions were extremely infrequent, and almost entirely directed at infantid males rather than other females (perhaps suggesting collective action rather than alliances as such). Indeed, the genus is invariably listed in compilations as not forming coalitions (e.g. Shultz et al., 2017; Sterck, Watts, & van Schaik, 1997).

Studies of wild baboons (Papio spp.) indicate that females who have more grooming partners experience lower cortisol titres (a marker of physiological stress) when exposed to naturally stressful social events (e.g. infantid males: Crockford et al., 2008; Engh et al., 2006; Wittig et al., 2008) as well as being more reproduc- tively successful and longer-lived (Silk et al., 2003, 2009, 2010a, b). Alliances based on grooming partnerships may have both active and passive effects: allies may come to one’s aid when one is attacked, but their presence alone may also deter others from attacking (Datta, 1983), while at the same time the close presence of an ally may make one feel less stressed by an agonistic incident.

![Graph showing regression for number of infants (0–12 months) per adult female against number of reproductive females in the group for guereza (dashed line) and gelada (solid line; for calculations, see Appendix). The dotted line marks the minimum fertility required for demographic replacement and a stationary population (ca. four offspring in a ca. 12-year reproductive life, allowing a 50% mortality rate to adulthood, based on data for Papio baboons: Altmann et al., 1977; Sigg et al., 1982). The vertical dotted lines mark the mean size of groups at the time they underwent fission (three guereza groups: dotted line; four gelada harems: short dashed line).](image-url)
Indeed, by triggering the endorphin system (Keperne, Martensz, & Tuite, 1989; Nummennmaa et al., 2016), grooming itself has analgesic properties that relax recipients (Schino, Scucchi, Maestripieri, & Turillazzi, 1988). The importance of the endorphin system for social bonding has been highlighted by recent genetic studies in humans (Pearce et al., 2017).

Because the strength of a relationship, and hence its reliability as an alliance, seems to be a function of the amount of time invested in it for both monkeys (Fig. 4) and humans (Roberts & Dunbar, 2011, 2015; Suctcliffe et al., 2012), and this time commitment is extremely costly (Dunbar, 2007; Miritello et al., 2013), there is inevitably a limit to the number of relationships of a given intensity/quality that an individual can maintain. In effect, once a social group exceeds a certain size, it becomes impossible for any one individual to groom with all other members due to constraints of time. This should mean that primate groups will always fission after a certain size, and this ought to happen at more or less the same group size in all species. Yet it is obvious that some species (e.g., baboons, macaques, chimpanzees) are able to maintain much larger groups as coherent, stable entities than other species (e.g., New World cebids, Old World colobines; Campbell et al., 2007). The present results suggest that the explanation may lie in the striking difference in the way social groups are structured, and in particular the formation of close-bonded alliance relationships that buffer the females against the high levels of stress induced by living in close proximity.

In respect of the density index (Fig. 1a), the slopes are essentially the same for the two species, but the intercepts differ, suggesting that gelada are able to maintain a higher proportion of dyadic ties than guereza across all group sizes. In part, this is because gelada can afford to devote more time to social grooming than guereza (Dunbar, 1992a; Korstjens & Dunbar, 2007). Guereza are tightly constrained in terms of their time budget by the need to rest for long periods to allow bacterial fermentation of a heavily leaf-based diet (Korstjens & Dunbar, 2007). The regression equations of Table 1 suggest that guereza would reach \( d = 0 \) (effectively no grooming dyads) at groups of size 12.8, whereas the gelada would not do so until groups of 218 individuals. At this point, social groups would be effectively unbonded and socially nonfunctional. These values correspond well to the relative difference in the upper limits on group size for these species: the largest observed C. guereza groups are around 12–15 (including immatures), with less than 1% of all groups larger than 15 (N = 127 groups: Fashing, 2007), whereas in gelada the upper limit on harem size is around 20–24 individuals (only 1% of groups >25, N = 161 groups: MacCarron & Dunbar, 2016).

Social group size in any animal species is a trade-off between costs and benefits. The most important benefit for most mammalian and avian species is reducing predation risk, but this has to be offset against the costs incurred (e.g., increased day journey length, more limited food acquisition and a disrupted time budget; Dunbar et al., 2009). The fertility effect identified here represents a further cost that has not been widely recognized (for a rare exception, see van Schaik, 1983). The extensive field and laboratory evidence for a relationship between fertility and group size suggests that this is a mammal-wide effect, and all else equal it will set a natural upper limit on the size of all mammalian social groups. Animals can afford to trade losses in fertility against gains in predation risk, thereby adjusting the equilibrium point to allow occupation of a particular habitat, but only up to a certain point, since no species can afford to allow fertility to drop below that required for replacement (two offspring that survive to reproduce in a lifetime). That guereza and gelada have different equilibrium points suggests that gelada have been able to find some way of mitigating the fertility decline to make that possible.

The differences in network structure seem to suggest that gelada have achieved this by creating grooming-based alliances between females. It seems, at least from the evidence for these two species, that these alliances do not necessarily reduce the rate of aggression, but rather may buffer females against the stressful effects of becoming involved in altercations. That relationships of this kind really do have significant fitness consequences is evident from long-term field studies of baboons and wild horses which demonstrate significant stress levels as well as fertility and longevity benefits from having close social partners (Cameron et al., 2008; Silk et al., 2003, 2009, 2010a,b).

Datta (1983) noted that rhesus macaques, Macaca mulatta, refrain from attacking another individual when they know that it has powerful allies, even when those allies are not physically present. Acting on knowledge that another individual has allies depends on having more sophisticated cognitive abilities in that animals must be able to factor the impact of absent third parties into the current situation (Dunbar, 1998). Neuroimaging studies of humans suggest that working in a partially virtual social world may be cognitively much more demanding than working in the real world within which one is physically embedded (Lewis, Birch, Hall, & Dunbar, 2017). It may therefore be no accident that these differences in social structure and limits on group size correlate with differences in both the typical size of grooming cliques (Kudo & Dunbar, 2001) and brain, and especially neocortex frontal lobe, volume (Dunbar, 2011; Joffe & Dunbar, 1997).

It may be significant that, across primates, brain size correlates with the number of adult females in the group better than it does with either the number of males or total group size (Dunbar, 1992b; Keperne, Martel, & Nevison, 1999; Lindenfors, 2005). Genomic imprinting studies suggest that neocortex volume (the main brain region associated with sociality) is determined mainly by maternally inherited genes, with the paternal alleles being silenced (Gregg et al., 2010; Keperne, 2013; Keperne et al., 1999). Both these findings reinforce the implication of the results reported here that, in primates, social evolution is driven mainly by the demands of the females. Fig. 5 suggests that those demands reflect the need to overcome fertility costs in order to be able to live in larger social groups (for whatever ecological reasons large groups may be required).

Shultz and Dunbar (2010a) have shown that, across mammals, the encephalization rate through evolutionary time characteristic of an order correlates with the proportion of its species that live in bonded social groups. Since, across primates, absolute brain (or brain region) volume correlates with executive function cognitive abilities (Deane, Isler, Burkart, & van Schaik, 2007; Shultz & Dunbar, 2010b), the contrasts in brain volume between guereza and gelada may reflect differences in psychological abilities that allow multiple relationships to be maintained even in the absence of significant investment in grooming. So far, the only taxon for which we currently have any evidence for species differences in social cognitive abilities are the apes (Herrmann, Call, Hernandez-Lloreda, Hare, & Tomasello, 2007), and these are in the predicted direction (species living in larger, socially more complex groups are more skilled in terms of social cognition). Similar experiments on monkeys, as well as other mammals, would illuminate the cognitive demands that underpin the differences in social structure identified here.

Finally, it is important to note that the way stress affects fertility is simply a consequence of the way mammalian reproductive endocrinology is organized. We should therefore expect to detect the consequences of group living on fertility and the effect of this on optimal and maximum group sizes in all mammalian taxa. I have focused here on primates, because this is one taxon where data are
available to explore these relationships. As noted above, many mammals (including equids, canids, hyaenids, marmots, suricates, rodents) exhibit a negative relationship between fertility and group size. Like the less social primates, many of these mammals live in small groups or have fission–fusion type social systems (the default way to defuse social stresses), but a few (elephants, dolphins, equids) have primate-like bonded social systems and some of these are multilevel societies like those of the gelada and chimpanzees (Hill et al., 2008). Whether these use bonded relationships and coalitions to buffer themselves against this effect remains to be determined.

Acknowledgments
Aleks Ruggens helped run some of the analyses. This research was supported by a European Research Council Advanced Investigator grant.

Supplementary material
Supplementary material related to this article can be found at https://doi.org/10.1016/j.anbehav.2017.12.005.

References


Appendix

Our primary interest is in the way fertility varies with the number of adult females in a group, independently of any effects due to environmental and other factors. While it would be possible to use actual birth rates for the gelada, the number of births in the guereza sample is small and hence likely to generate large random fluctuations when the typical birth interval is close to 24 months. I therefore used numbers of infants aged under 12 months recorded in a group at the end of each study period, and averaged these across the number of adult (i.e. reproductive) females in the group; I then averaged these across groups with the same number of females within a single population; finally, I averaged these averages across samples for each species. To avoid small sample bias, I required that at least three females be sampled at each group size.

This approach is likely to slightly underestimate actual birth rates, but the data for gelada suggest that number of infants counted in this way correlates with the birth rate (see below); so it is unlikely that the underestimate will be large or significant. The most likely source of the difference will be stillbirths (themselves a likely consequence of stress), poor mothering, accidental deaths and infanticide. It is unlikely that any of these will be correlated with the number of adults in the group, unless it is mediated by stress. Indeed, postnatal deaths, including infanticide, will normally have the effect of bringing the female back into breeding condition immediately, and hence lead to an increase rather than a decrease in the birth rate with respect to group size.

Note that, since I am testing a specific hypothesis (fertility is negatively related to the number of females in the group), all statistical tests of this relationship are necessarily one-tailed. If there is a significant positive relationship between fertility and number of females, this would be just as much evidence against the hypothesis as no correlation at all.

Gelada fertility

The gelada fertility data derive from a single band sampled in two separate years (1971–1972 and 1974–1975; see Supplementary Material, Table S2). In total, there were 42 harems in the sample, comprising 175 breeding females. It has previously been noted that these 2 years had strikingly different birth rates; in part, this was a demographic effect arising from a combination of reproductive synchrony among the females and a natural cycle in the growth and fission in harem size, as well as a difference in demographics. Mean harem size was 3.4 females in the first year and 4.9 females in the second, with birth rates being ca. 20% higher in the first (Dunbar, 1984; Ohsawa & Dunbar, 1984). For present purposes, the data are based on the number of infants aged <12 months recorded in individual harems at the end of each study period. This index of fertility (mean infants per female) correlates significantly with the annual birth rate per female for these harems (r = 0.469, N = 42 individual harems, P = 0.0009 one-tailed).

The data for each harem in the two time periods are shown in Fig. A1, distinguished by study year. Quadratic regressions provide much better fits (r² = 0.154 and r² = 0.169, for the two samples) than linear regressions (r² ≤ 0.027).

With study year as a binary variable (1972 = 0, 1975 = 1), the common best fit least squares regression line combining both distributions is:

\[
\text{Infants per female} = -0.112 - 0.143 \times \text{Study year} + 0.271 \times \text{Females} - 0.027 \times \text{Females}^2
\]

\[(r^2 = 0.419, F_{3,37} = 2.63, P = 0.033).

Smoothing the data by averaging within group sizes within years yields an equation with very similar parameters:

\[
\text{Infants per female} = -0.045 - 0.152 \times \text{Study year} + 0.230 \times \text{Females} - 0.022 \times \text{Females}^2
\]

\[(r^2 = 0.785, F_{3,10} = 5.34, P = 0.001).

For purposes of Fig. 5 (where we are only interested in the relationship to number of females), I use the common best fit regression line combining both distributions without distinguishing study year:

\[
\text{Infants per female} = -0.092 + 0.233 \times \text{Females} - 0.024 \times \text{Females}^2
\]

Gueræza fertility

Ideally, we would want to determine guereza fertility rates from the same population from which the network data were sampled (Bole Valley). However, the number of females in the sample is small and hence subject to erratic variance. To provide a large enough sample, I combine data on number of infants aged under ca. 12 months (white babies and small infants) for 53 groups from five populations of C. guereza in eastern Africa where data are given for multiple groups studied during extended field studies (see Supplementary Material, Table S2).

Fig. A2 plots the raw data. All regression slopes are negative. Although, individually, none of the regressions are significant (not surprisingly given the small samples in each case), the chances of getting five negative slopes out of five is significant (sign test: \(P = 0.032\) one-tailed).

There is some evidence that between-habitat differences in fertility are influenced by habitat quality: the intercepts for the individual linear regression equations in Fig. A2 are quadratically related to habitat mean ambient temperature (Fig. A3; \(r^2 = 0.788\)). This is reminiscent of the relationship observed in Papio baboons, where birth rates are independently determined by number of adult females (but not males) in the group and by the ambient temperature of the habitat, in both cases with quadratic relationships (Hill et al., 2000). The relationship with temperature is not surprising: birth rates will inevitably be lower in habitats at the edges of the species’ biogeographical distribution where the environment will be more challenging for the animals. Mean number of females in the group is not significantly correlated with mean ambient temperature (\(r = 0.676, N = 5, P = 0.210\)), so habitat quality alone cannot predict the fertility pattern in Fig. 5.

We can check on the generality of this result via a different route. Fashing (2007) compiled data on the mean number of juveniles (of all ages) per female for different populations of all five species of the genus Colobus. These values are, of course, likely to be subject to mortality effects that might vary between habitats. Nevertheless, plotting this index of fertility against mean number of adult females in a group for these populations yields a significant negative relationship (Fig. A4: \(b = -0.0165, t_{6} = 2.06, P = 0.028\), even without any correction for habitat quality.

For the data in Fig. A2, the best fit least squares linear and quadratic regressions are:

\[
\text{Infants/female} = 0.508 - 0.051 \times \text{Females}
\]

\[
\text{Infants/female} = 0.383 + 0.047 \times \text{Females} - 0.016 \times \text{Females}^2
\]

Averaging numbers of infants per female from Fig. A2 across groups with the same number of females yields a more obviously quadratic relationship (Fig. A5). The best fit quadratic equation is:

\[
\text{Infants/female} = 0.291 + 0.083 \times \text{Females} - 0.020 \times \text{Females}^2
\]
The guereza relationship could be interpreted as a simple linear negative slope (a linear regression set through all the data in Fig. A2 gives a slightly better fit), but, because the data are necessarily censored at $N = 1$ female on the left side of the graph, the difference is marginal.

Note how similar the slope parameters are for the gelada and guereza quadratic equations. Compared to that for the guereza, the gelada equation is simply displaced to the right (i.e. has a much lower intercept), yielding a peak at five females rather than two females as in the guereza case.

For consistency, I plot both guereza and gelada as simple quadratic relationships in Fig. 5. However, the interpretation of the results remains the same whether guereza are plotted as a linear or a quadratic function: either way, guereza fertility declines with number of females from an early peak, whereas that for gelada has its peak shifted to the right.

---

**Figure A1.** Mean number of infants aged <12 months per female in the Sankaber Main band of gelada in 1972 (open symbols, dashed line) and 1975 (filled symbols, solid line).

**Figure A2.** Mean number of infants per adult female in five populations of *Colobus guereza* in eastern Africa. Solid circles, solid line: Bole Valley, Ethiopia, 12 groups (Dunbar, 1988); open circles, long dash line: Lake Shalla, Ethiopia, six groups (Dunbar, 1989); squares, dotted line: Kakamega Forest, Kenya, five groups (Fashing, 2001); solid triangles, medium dash line: Kanyawara, Kibale Forest, Uganda, seven groups (Oates, 1977); open triangles, short dash line: Budongo Forest, Uganda, 25 groups (Suzuki 1979). Lines are least squares linear regressions.

**Figure A3.** Intercept for the linear regression equations relating number of infants per female to number of females in the group for the individual study sites in Fig. A2, plotted against habitat mean annual temperature (as an index of habitat quality). Dashed line: best fit quadratic regression equation.

**Figure A4.** Mean number of immatures (all ages) per female plotted against mean number of females per group for individual populations of the five species of the genus *Colobus*. *C. guereza*: filled circles; *C. angolensis*: open circles; *C. polykomos*: filled triangles; *C. sitanus*: open triangles; *C. vellerosus*: square. Dashed line: best fit linear regression equation. Source: Fashing (2007).

**Figure A5.** Mean number of infants in guereza groups with different numbers of females, averaged across all five sites.