1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24

Title: The Goldilocks Effect: Female geladas in mid-sized groups have higher fitness

Running head: The Goldilocks Effect

Authors: Elizabeth Tinsley Johnson^{1*}, Jacob A. Feder^{2*}, Thore J. Bergman^{3,4}, Amy Lu^{2,5}, Noah Snyder-Mackler^{6,7**}, Jacinta C. Beehner^{3,8**}

Affiliations:

¹ Department of Integrative Biology, Michigan State University, East Lansing, MI 48824, USA.

² Interdepartmental Program in Anthropological Sciences, Stony Brook University, Stony Brook, NY 11794-4364, USA.

³ Department of Psychology, University of Michigan, Ann Arbor, MI 48109-1043, USA.

⁴ Department of Ecology & Evolution, University of Michigan, Ann Arbor, MI 48019-1085, USA.

⁵ Department of Anthropology, Stony Brook University, Stony Brook, NY 11794-4364, USA.

⁶ School of Life Sciences, Arizona State University, Tempe, AZ 85287-4701, USA.

⁷ Center for Evolution and Medicine, Arizona State University, Tempe, AZ 85287-1701, USA.

⁸ Department of Anthropology, University of Michigan, Ann Arbor, MI 48109-1107, USA.

* Co-first authors

** Co-senior authors

Corresponding authors: etinsley@msu.edu; nsnyder@asu.edu; jbeehner@umich.edu

25 Abstract:

26 The cost-benefit ratio of group-living is thought to vary with group size: individuals in “optimally-sized”
 27 groups should have higher fitness than individuals in groups that are either too large or too small.
 28 However, the relationship between group size and individual fitness has been difficult to establish for
 29 long-lived species where the number of groups studied is typically quite low. Here we present evidence
 30 for optimal group size that maximizes female fitness in a population of geladas (*Theropithecus gelada*).
 31 Drawing on 14 years of demographic data, we found that females in small groups showed the highest
 32 mortality risk, while females in mid-sized groups exhibited the highest reproductive performance. This
 33 group-size effect on female reproductive performance was largely explained by variation in infant
 34 mortality (and, in particular, by infanticide from immigrant males) but not by variation in reproductive
 35 rates. Taken together, our results show that females in mid-sized groups appear to attain optimal fitness
 36 due to the combination of social pressures and, potentially, predation. Our findings provide insight into
 37 how and why group size shapes fitness in long-lived species.

38

39 Keywords: optimal group size, fitness, folivore paradox, infanticide, reproduction, survival

40 INTRODUCTION

41 Variability in group size within species reflects a balance between the costs and benefits of group living
 42 [1,2]. For example, individuals living in larger groups may experience increased intragroup foraging
 43 competition (a cost: e.g., [3–5]) but decreased predation risk (a benefit: e.g., [3,6–9]). Individuals living
 44 in optimally-sized groups, where (by definition) the benefits outweigh the costs, should have the highest
 45 lifetime reproductive success [i.e., fitness: 10] compared to others in the population [11]. Although the
 46 link between group size and fitness has long been the focus of socioecological theory, it has been
 47 challenging to demonstrate this relationship with empirical data. This is, in large part, because group
 48 size fluctuates over an individual’s lifespan, meaning it is difficult to align a specific group size with a
 49 clear fitness metric (especially in long-lived species). As a result, testing the effects of group size
 50 requires long-term data across multiple groups of variable size.

51
 52 Ecological costs and benefits at both extremes of group size support the idea that mid-sized groups are
 53 optimal. For example, the feeding costs of living in large groups are well-documented: variables that
 54 lead to greater feeding competition, such as home range size, day range length, and feeding time are
 55 often positively correlated with group size [carnivores: 12, primates: 13,14], and those that reflect the
 56 consequences of competition, such as body condition, fertility, and survival are often negatively
 57 correlated with group size [15–20]. Despite the costs of living in large groups, individuals are often
 58 found in groups significantly larger than the theoretical species optimum [21,22], likely because splitting
 59 into two smaller groups is equally, if not more, costly. This may be particularly true if small groups are
 60 more susceptible to predation [23,24] or if competitively weaker groups are pressed into inferior home
 61 ranges (carnivores: [25], primates: [26,27]). Perhaps as a consequence of costs at both extremes,
 62 individuals in both large and small groups exhibit higher glucocorticoid levels, greater ranging
 63 disadvantages, and experience more extreme energetic conditions when compared to individuals in mid-
 64 sized groups (e.g., [14,28–30], reviewed in [5]).

65

66 However, many species do not exhibit group sizes that conform to predictions based on ecological
67 factors like feeding competition alone [e.g., 31]. For example, although folivores are thought to
68 experience low feeding competition, females in numerous folivorous taxa live in groups that are smaller
69 than those found in frugivorous species (the "Folivore Paradox": [32], e.g. [33]). For some folivorous
70 species, feeding competition may still be high, thus limiting group size (primates: [34], elephants: [35],
71 deer: [36]). Alternatively (or additionally), in many folivorous primates, selective pressures may favor
72 smaller groups because they are less attractive to immigrant males and thus less vulnerable to infanticide
73 [37,38]. By contrast, in some species the risk of infanticide may favor larger groups that are better able
74 to prevent male immigration or defend dependent offspring [39]. As a result, conspecific threat of
75 infanticide has the potential to shape group size such that mid-sized groups are optimal, but isolating the
76 effects of such social factors from environmental factors has proven challenging.

77

78 Geladas are folivorous monkeys that feed on widely-dispersed food resources, that comprise primarily
79 graminoid (grass and sedge) leaves [40,41]. Due to this resource base, geladas should face low feeding
80 competition and should, correspondingly, live in large groups for predator protection [32]. However,
81 gelada social organization is not so simple: individuals form small reproductive groups (hereafter,
82 "units", varying in size from 1-12 adult females: [42]) that, in turn, unite to form large bands (100+
83 individuals) that travel and forage together in the same home range (forming a 'multi-level' society:
84 [43]). The large foraging band likely represents the ecological unit (i.e., by offering some degree of
85 protection from predation, [44]) and conforms to the predictions of standard socioecological theory [32].
86 In contrast, the small reproductive unit represents the social unit and, as such, is likely shaped by social
87 factors. Specifically, in this population, male immigrations (called "takeovers" because the new male
88 replaces the dominant group male) dramatically alter female reproductive patterns via infanticide and
89 the Bruce effect [45,46]. Infanticide is the leading cause of infant mortality: immigrant males can kill up

90 to half of the dependent infants in a unit [47]. Moreover, larger units experience more takeovers than
 91 smaller units [48].

92

93 Ecological factors like habitat quality are unlikely to vary significantly among units in the same band;
 94 therefore, we predict that females in larger units incur disproportionately higher fitness costs due to
 95 social factors like male takeovers compared to females in smaller units. Alternatively (or additionally),
 96 when a takeover occurs, infants in small groups may have a higher chance of being killed: new dominant
 97 males are faced with fewer fertile females as groups get smaller and may, as a result, be more likely to
 98 commit infanticide. If both large and small units incur fitness costs due to social factors, then we predict
 99 that females in mid-sized units will exhibit the highest fitness. Finally, adult female survival may also be
 100 affected by unit size: females in small units may be more vulnerable to predation (e.g. [23]), while
 101 females in larger units may be more vulnerable to disease (e.g. [49]). If both factors affect female
 102 longevity in geladas, then we predict the lowest adult female mortality in mid-sized units.

103

104 To test these predictions, we first examined the evidence for an optimal unit size for gelada females.
 105 Specifically, we tested if unit size predicted female mortality rates and/or reproductive performance (the
 106 production of surviving infants). Second, because reproductive performance could be a product of both
 107 female fertility and/or infant survival, we explored the effect of unit size on these two components.
 108 Specifically, we assessed variation in both interbirth intervals, which should track both female energetic
 109 condition (e.g., [50,51], reviewed in [7]) and takeover-related pregnancy loss [45,52], and in infant
 110 mortality, which presumably is driven by infanticide. Third, we examined how the cause of infant deaths
 111 varied by unit size, specifically considering the extent to which maternal mortality and/or infanticide
 112 explained these patterns. Taken together, our results show that females in mid-sized units appear to
 113 display optimal fitness, resulting from the combination of survival and social pressures.

114 MATERIALS AND METHODS

115 *Study site and subjects*

116 The data for this study derive from 14 years of observation (Jan 2006-Jul 2020) on a population of wild
 117 geladas living in the Simien Mountains National Park, in northern Ethiopia (13°13.5' N latitude). The
 118 Simien Mountains Gelada Research Project (SMGRP, formerly the University of Michigan Gelada
 119 Research Project) has collected behavioral, demographic, genetic, and hormonal data from individuals
 120 since Jan 2006. All gelada subjects are habituated to human observers on foot and are individually
 121 recognizable. We used longitudinal data from 200 adult females in 46 reproductive units (20 original
 122 “founding” units that later became 26 “daughter” units due to fissions).

123

124 *Unit size*

125 The identities of all individuals in a unit were recorded each day the unit was seen. For each month of
 126 the study period, we recorded the total number of adult females in each unit; where changes in unit size
 127 occurred (i.e., due to adult female deaths or subadult female maturations), we used the maximum
 128 number of adult females in a unit in a given month. We focus on adult females because we have
 129 longitudinal records of the number of adults in each unit for the entire study period but only started
 130 recording the total number of individuals (including juveniles and infants) in 2012. However, when we
 131 compared the number of adult females in a unit to the total number of individuals in a unit for the subset
 132 of data where we have both (2012-2020), we found that both values were highly correlated (Pearson
 133 correlation coefficient = 0.76, p -value = 2.2×10^{-16}).

134

135 Gelada units contain one reproductively dominant leader male as well as 0-3 follower males that receive
 136 few reproductive opportunities but can deter male takeovers [42]. Thus, we did not include males when
 137 calculating group size, as the number of females and the number of males could have contrasting
 138 impacts on female fitness components. Furthermore, the number of males was weakly correlated with

139 the number of females (Pearson correlation coefficient = 0.18, p -value = 2.2×10^{-16}). For these reasons,
140 we consider males separately in all relevant analyses.

141

142 Changes in the number of adult females in a unit were primarily due to female maturations and deaths.

143 Maturations were recorded as the first observation of a sex skin swelling (details are outlined in [53]).

144 Deaths were recorded as the first day an adult female was no longer observed with a unit for more than

145 three consecutive encounters with that unit (and not observed in a different unit, as in the case of

146 transfers or fissions).

147

148 Dates of fissions, fusions, and female transfers were calculated as the first day the unit females were no

149 longer observed together and subsequently observed either in a separate daughter unit with a new leader

150 male (for fissions), together with non-unit females and a new leader male (for fusions), or associating

151 with a different unit and new leader male (for the rare cases of female dispersal). In all cases, we

152 immediately identified known females in daughter units or new units following their disappearance from

153 their natal unit.

154

155 All models included female unit size as a continuous predictor variable. However, for visualization

156 purposes, we also categorized units into small, medium (i.e., mid-sized), and large units based on the

157 observed range in variation in sizes. Specifically, cut-offs were determined by calculating the tertiles of

158 the observed monthly distribution of sizes: “small” indicates units of less than or equal to 4 adult

159 females; “mid-sized” indicates units from 5-7 adult females; and “large” indicates units of 8 or more

160 adult females.

161

162 *Adult female mortality*

163 For this analysis, we included data on 200 adult female geladas. All adult females had known or
 164 estimated birth dates from which we calculated age. Dates of birth were known for 56 adult females
 165 born during the study, while the remaining females' birth dates were estimated using their size at the
 166 start of observation (n=44), back-calculated using life history milestones (i.e., maturation or first birth,
 167 n=40), or extrapolated from their total number of infants (n=60). The mean age at the mid-point of the
 168 study (2013) for all females was 11.86 +/- 5.14 SD years (range = 3.54 - 28.90 years). There was no
 169 association between unit size and female age (estimated or known; females estimate = -2.01 +/- 2.35 SE,
 170 p -value = 0.394).

171
 172 To assess whether the rate at which adult females died varied by unit size, we calculated each females'
 173 average unit size during each year of observation. Here, the unit of observation was the female-year.
 174 From this, we constructed a binomial generalized linear mixed model (GLMM) using the *lme4* package
 175 (version 1.1.20: [54]) in R (version 3.6.0: [55]). The dependent binary variable indicated whether the
 176 female died during the study year, which was modeled as a function of the following predictors: unit
 177 size (the average number of females in the unit, including both the linear and the quadratic term), the
 178 females' age at the start of the year (linear and quadratic), and the number of males (i.e., the total
 179 number of adult follower males plus one leader male). We controlled for the repeated measures of
 180 female, unit, and year as random effects.

181

182 ***Reproductive performance***

183 For the majority of infants born during the study period (n = 352 out of 394 total), the date of birth was
 184 known within a day or two. For those infants where exact date of births were unavailable (n = 42), we
 185 were able to assign a birth date within 1 month of the actual birth date based on established
 186 morphological and behavioral criteria (i.e., the size of the infant, infant motor skills, presence/absence of
 187 the umbilical cord; for more details, see [53]). Conception dates were retrospectively assigned by

188 subtracting the mean gestation length ($n = 183$ days; see [53]) in the population from the observed or
189 estimated date of birth.

190

191 We used a binary scoring system to assign female reproductive performance on a monthly basis.
192 Females were “successful” in a given month if they conceived an infant that survived to 1.5 years of age,
193 which is the mean age of weaning in this population (regardless of actual weaning date). Females were
194 considered “unsuccessful” if they conceived an offspring that died before reaching 1.5 years of age or if
195 they did not conceive during the calendar month. We excluded all months and births after 2018, as the
196 survival outcomes of these infants were not resolved by the end of the study period.

197

198 To assess the effects of unit size on female reproductive performance, we constructed a binomial
199 GLMM. The dependent binary variable was the monthly success variable for each female-month (273
200 successful female-months, 996.5 female-years total, 188 females). We modeled this outcome variable as
201 a function of the following predictors: female age (both the linear and quadratic term, to control for the
202 known effects of female age on reproductive output, reviewed in [56]), number of females, and the
203 number of males. We controlled for the repeated measures of individual identity, unit, and year as
204 random effects. We additionally modeled reproductive performance using 6-month bins (i.e., Jan-Jun,
205 Jul-Dec), using the same binary outcome listed above and including the average number of females and
206 males as fixed effects. The results from this separate analysis (Supplementary Table S6) were congruent
207 with those generated from the models using monthly data.

208

209 Although we do not have the data necessary to calculate lifetime reproductive success, we coarsely
210 estimated this value by using annual mortality rates and successful reproductive rates at the three unit
211 size categories (i.e., small, medium, large). Specifically, we summed the total number of deaths and
212 successful offspring within each unit size category and divided by the total number of female-years

respectively. Using these annual mortality rates, we then calculated median reproductive lifespans, which were then multiplied with corresponding reproductive rates to approximate lifetime reproductive success at each unit size category.

Interbirth intervals

Using the birth dates of all infants in this population, we calculated interbirth intervals (IBIs) between females' successive offspring. We constructed two linear mixed models (LMM) using the *lme4* package. For both models, the dependent variable was the length of the IBI in days, which was modeled as a function of unit size at the first infant's birth (both linear and quadratic), the mother's age, and the infant's sex. In the first model, we utilized only IBIs following surviving offspring ($n = 187$), as infant mortality substantially shortens IBIs in this population and elsewhere [47]. However, because IBIs may capture instances of takeover-related fetal loss (i.e., the 'Bruce effect:' [45]), which can lengthen IBIs [52], in the second model we only included the subset of IBIs that did not contain takeovers or infant deaths ($n = 82$). Rather than reflecting pregnancy loss, variation in these "takeover-independent" IBIs should reflect the ecological advantages (e.g., higher food intake) that some females have over others in their ability to wean offspring and conceive again.

Infant mortality

The disappearance of any infant prior to the average age at weaning in this population (1.5 years of age: [53]) was assumed to be a case of infant mortality. To assess infant mortality and its causes across unit sizes, we constructed a mixed-effects Cox proportional hazards model using the R package *coxme* [57]. For this model, we used all 394 infants born during the study period. Infants entered the dataset at birth and were censored at death ($n = 90$) or if their social units stopped being observed ($n = 17$) and exited the dataset when they survived to 1.5 years of age ($n = 287$). We used the number of females present during the month of the infant's birth as a fixed effect, including both linear and quadratic terms. We

238 included birth-year, unit, and maternal ID as random effects. Schoenfeld residuals showed no significant
239 deviations from the assumption of proportional hazards.

240

241 *Takeover rates*

242 We recorded the dates of all observed male takeovers ($n = 80$) of known reproductive units (following
243 [46]) as well as the number of mature females in the unit at the time of takeover. We calculated the
244 length of time each unit was seen at a given size for each year the unit was observed (244 total unit-
245 years) and counted the number of takeovers that occurred during these periods. To calculate the
246 influence of unit size (i.e., number of adult females) on the likelihood of takeover, we modeled the
247 number of takeovers a unit experienced as the dependent variable in a Poisson GLMM, offset by the
248 total observation time the unit was seen at a given size. To control for repeated measures across units
249 and time, we included unit and year as random intercepts. Finally, we included two fixed effect
250 variables: the number of adult females in the unit and the average number of males.

251

252 *Causes of infant mortality*

253 To examine whether the probable causes of infant mortality varied with unit size, we assigned the cause
254 of mortality based on the following characteristics: if the infant's death occurred within 6 months of a
255 takeover, the cause of death was recorded as "infanticide" ($n = 31$) [46,47]. If the infant died within 3
256 months of their mother's death or disappearance, the cause of death was recorded as "maternal death" (n
257 $= 27$). If maternal and infant disappearances co-occurred within the 6 months following a male takeover,
258 however, the cause of death was still flagged as "maternal death" ($n = 4$). All other causes of infant
259 deaths were recorded as "unknown" ($n = 32$). For each of these mortality outcomes, we constructed a
260 binomial GLMM using the *lme4* package. Here, we modeled whether the infant died of each respective
261 mortality cause (i.e., maternal death, infanticide, and unknown) as a binary variable, using unit size
262 (both linear and quadratic) as a fixed effect and unit as a random effect. For these three analyses, we

used only infants whose survival outcomes were known (i.e., no right-censored offspring; $n = 377$). All figures were constructed using *ggplot2* [58], and prediction intervals were extracted from mixed-models using the *effects* package [59].

RESULTS

Adult female death rates

Of the 200 adult females included, 90 died before the end of the study period (median age at death = 15.99 years). Unsurprisingly, the odds of dying increased substantially with age (age estimate = 26.69 $\pm$ 4.41 SE, p -value = 1.4×10^{-9}). Females in small units had the highest mortality rates (females estimate = -8.33 $\pm$ 4.41 SE, p -value = 0.043, figure 1a). Specifically, females in small units had an annual mortality rate of 9.8%, while females in medium and large units had an annual mortality rate of 6.8% (figure 1b). Using these annual mortality odds, we estimated that the median reproductive lifespan in small units would last 7.18 years, while the reproductive lifespan in mid-sized and large units would last 10.83 years.

Reproductive performance

Although units in this population can range in size from 1-12 adult females, females in units in the middle of this range (i.e., between 5-7 adult females, hereafter “mid-sized” units) had the highest reproductive performance. Females in mid-sized units were most likely to conceive successful offspring (females² estimate = -17.86 $\pm$ 6.95 SE, p -value = 0.0102, figure 1c) and were 17.6% more successful than females in small units and 39.8% more successful than females in large units (figure 1d). Given the reproductive lifespans estimated above, females are expected to produce approximately 3.46 infants in mid-sized units, compared to 1.94 infants in small units, and 2.47 infants in large units.

288

289
290
291
292
293
294
295
296
297
298
299
300
301
302
303

Figure 1. Female fitness is optimal in mid-sized units. (a) The association between unit size (number of females in a unit) on the annual probability of death, calculated as the number of deaths observed at each unit size divided by the total number of female-years. Confidence bands show the lower and upper limits of the model predictions, with colors following the unit size categories described in *Methods*. The size of each point is proportional to the number of female-years observed at each unit size. (b) The annual death-rate by unit size category. The dotted horizontal line = 0.08 (the population mean). Unit size categories were used for visualization purposes only (see *Methods*). (c) The effect of unit size on the annual likelihood that females conceived an infant that survived to weaning. Points indicate successful conception rates at each unit size, generated from raw data. The size of each point is proportional to the number of female-years observed at each unit size. Confidence bands show the lower and upper limits of the model predictions. Although the model output specified monthly probabilities, we converted these to annual probabilities for the sake of visualization. (d) The rate of producing surviving offspring by unit size category. The dotted horizontal line represents the population mean (0.274). Unit size categories were used for visualization purposes only (see *Methods*).

304 *Interbirth intervals*

305 IBIs tended to be longer in larger units (females estimate = 493.9 +/- 242.6 SE, p -value = 0.043;
 306 females² estimate = 388.7 +/- 226.4 SE, p -value = 0.087; figure 2a). However, this effect was no longer
 307 significant when only IBIs that were uninterrupted by takeovers were included for analysis (females
 308 estimate = 39.18 +/- 230.7 SE, p -value = 0.868; females² estimate = 377.7 +/- 225.1 SE, p -value =
 309 0.107; figure 2b).

310

311 *Infant mortality*

312 Out of 394 infants, 90 (22.8%) died before reaching 1.5 years. Infants born into small and large units
 313 were more likely to die before this age than infants born into medium-sized units (HR=1.06 +/- 0.017
 314 SE, p -value = 0.0018; figure 2c). Specifically, 28.1% and 33.8% of infants born into small and large
 315 units died before weaning, while only 10.4% of infants born into mid-sized units died before reaching
 316 this age.

317

318
319
320
321
322
323
324
325
326
327

Figure 2. Unit size does not influence interbirth intervals, but does impact infant survival. (a) Interbirth intervals across unit size categories, including intervals where takeovers occurred. **(b)** Interbirth intervals across unit size categories, excluding cases where takeovers occurred. Each red, orange, and yellow point represents an interbirth interval at each unit size category. Black points and error bars indicate means and 95% confidence intervals across the three categories **(c)** The effect of unit size at birth on infant survival, as indicated through survival curves to 1.5 years of age. Unit size categories were used for visualization purposes only.

328

329 **Takeover rates**

330 Takeover frequency increased with unit size (females estimate = 7.33 +/- 2.48 SE, p -value = 0.003,
331 figure 3a). On average, small units experienced male takeovers once every 4.15 years, mid-sized units
332 once every 3.17 years, and large units once every 1.61 years. However, the number of males did not
333 influence the frequency of takeovers (males estimate = -0.08 +/- 0.113 SE, p -value = 0.458).

334

335 **Causes of infant mortality**

336 Infanticides and maternal death accounted for 38.9% and 25.6% of all dependent infant deaths
337 respectively. Despite the fact that small units experienced the lowest takeover rates, females in mid-

sized units experienced the lowest infanticide rates (females² estimate = 10.81 +/- 3.72 SE, *p*-value = 0.004, figure 3b); 9.1% and 12.8% of infants born into small and large units experienced infanticide respectively, while only 3.0% of infants in mid-sized units experienced infanticide.

Second, infants in mid-sized units were about half as likely to die as a result of maternal death as females in small and large units (females² estimate = 5.25 +/- 3.38 SE, *p*-value = 0.12, figure 3b), though this effect was not significant. Specifically, 9.1% and 8.3% of infants born into small and large units died following their mothers' deaths, while only 4.5% of infants in mid-sized groups died under these circumstances. Lastly, there was no association between unit size and mortalities of unknown cause.

Fig. 3. Takeovers and the causes of infant mortality vary with unit size. (a) The annual takeover-rate by unit size category. The dotted line represents the population mean (0.341 takeovers per unit-year). (b) The proportion of infants that died before 1.5 years of age, partitioned by likely causes across unit size categories. For the plots above, unit size categories are used solely for visualization purposes.

356 **DISCUSSION**

357 Taken together, our results show that females in mid-sized units exhibit the highest fitness. Females in
 358 smaller units have the highest mortality rate and females in both smaller and larger units have lower
 359 reproductive performance, primarily as a result of social pressures. Specifically, male takeovers are
 360 more frequent in larger units and are more likely to lead to infanticide when they do occur in smaller
 361 units. We also show that infant death rates, but not takeover-independent female fertility (i.e.,
 362 uninterrupted interbirth intervals), vary by unit size, with the lowest infant death rates occurring in mid-
 363 sized units. As a result, females in mid-sized units demonstrate the highest reproductive performance
 364 and, when factoring in variation in adult female death rates, are expected to produce more offspring over
 365 the course of their reproductive lifespan than are females in small or large units.

366
 367 Social factors – specifically, conspecific threat of infanticide – have been proposed as one potential
 368 solution to the “folivore paradox” [37], where folivorous species tend to form smaller groups than would
 369 be expected based on resource availability (which we do not expect to vary significantly across gelada
 370 units, as units within the same band share a homerange). However, larger groups have also been
 371 suggested to act as a deterrent to male takeover and infanticide [39]. Here we show that both forces may
 372 work in conjunction in gelada society to result in optimal mid-sized groups. Infanticide in large units
 373 was driven by high rates of male takeovers. By contrast, we are less certain why infanticide in small
 374 units was also high, despite infrequent takeovers. One possibility is that small units have poorer infant
 375 defense (e.g. as suggested in langurs: [39]). Another possibility is that new males may be more likely to
 376 commit infanticide in smaller units because smaller units present fewer immediate reproductive
 377 opportunities (i.e., receptive, cycling females) than larger units. A third possibility is that infanticide risk
 378 (similar to predation risk) may also follow a “dilution effect” [60], where infants in large units are
 379 buffered by an increased number of potential targets. By contrast, our proxy for ecological factors, IBI
 380 ([7]), did not appear to explain variation in reproductive performance across units and did not differ

across unit sizes (after removing cases where male-mediated pregnancy loss may have occurred). In sum, variation in gelada reproductive performance was largely explained by social factors.

Adult female mortality, which contributes to low lifetime reproductive success, was highest in small units. Although female deaths could be triggered by takeover-related injuries [52], smaller units experienced less frequent takeovers than mid-sized or large units; and, we did not observe an increase in female deaths after takeovers (in contrast to infant deaths, which do increase following takeovers: [52]). Females in small units may be more likely to die because females in larger units can monopolize food resources. Although folivores are thought to experience less feeding competition (due to the dispersed nature of their resources: [32]), research suggests that both within-group and between-group competition over resources can still influence female nutrition and energetics (e.g., [61,62]). Female geladas form strict dominance hierarchies within units [63], but our understanding of between-unit dominance is lacking. However, given that feeding-related aggression between units is relatively rare and low-intensity [64], we consider it more likely that females in small units may be more vulnerable to predation. Although the large foraging bands characteristic of gelada social structure likely act as a deterrence to predators [44], females in larger units may monopolize locations within the band during the day or sleeping position on the cliffs at night, thus leaving small units vulnerable at the periphery. This question will require simultaneous targeted data collection on relative group position - a focus of future research.

What (if anything) can females in suboptimal groups do? Units consist of close maternal kin and females tend to remain in the same unit they were born into [65]; in this population, individual female dispersal is rare ($n = 2$; 0.008 per unit-year). Small units, with their lower reproductive performance, require more time to grow in size via juvenile maturation. Preliminary observations from our study population suggest that small units can sometimes fuse together into one larger unit, but such occurrences are also rare ($n =$

3; 0.012 per unit-year) and involve a number of potential hurdles (e.g., resistance of leader males and/or unrelated females). In contrast, gelada social organization may offer a unique opportunity for fissions to occur, as bands may buffer the ecological costs of fissions, but fissions resulting in too-small units may be more costly than remaining in a too-large unit (and again, females might be limited by the presence of a resident leader male). In our population, large units have been observed to fission ($n = 11$; 0.045 per unit-year), however, more long-term data (particularly focused on social network metrics among females) are needed to understand the factors that lead up to both fissions (and subsequent group membership decisions) and fusions.

As previously reported in this gelada population [48], takeover frequency was positively correlated with the number of females (but not males) in a unit. This stands in contrast to our previous research showing that multimale units (consisting of a leader male and 1+ “follower” males) experienced fewer takeovers per unit-year than units with only one adult leader male across all unit sizes [42]. Indeed, in a number of primates, group composition (e.g., the ratio of males to females) has been shown to reduce the risk of takeovers and/or infanticide (e.g. [66], reviewed in [67]). For example, reproductive success of female white-faced capuchins (*Cebus capucinus*) was both negatively correlated with group size and positively correlated with the ratio of males to females in the group [68]. In contrast, multimale groups of ursine colobus monkeys (*Colobus vellerosus*) had higher rates of male immigration and infanticide, perhaps because these dominant males were ineffective at deterring immigrants [38]. Thus, the consequences of follower males in geladas may rather depend on their individual reproductive tactics and the ability of leader males to regulate group membership, which require further data and analysis; whether females in small or large *multimale* units demonstrate higher reproductive performance than females in small or large *single-male* units remains to be seen.

430 The link between group size and individual fitness has long been the focus of socioecological theory, yet
 431 it has been difficult to demonstrate this relationship for long-lived species with empirical data. Here, we
 432 were able to leverage long-term data from a population of wild geladas comprising multiple groups of
 433 variable size to test the extent to which social and ecological factors shape optimal group size. Our
 434 findings provide insight into how and why group size shapes female fitness – specifically, adult female
 435 mortality and conspecific infanticide.

436 ACKNOWLEDGEMENTS

437 We would like to thank the Ethiopian Wildlife Conservation Authority (EWCA) and the wardens and
 438 staff of the Simien Mountain National Park for their permission and ongoing support for our long-term
 439 research project. Additionally, we are grateful to our excellent field team across the years, most
 440 especially E. Jejaw, A. Fanta, S. Girmay, J. Jarvey, and M. Gomery for their assistance with field data
 441 collection. We also owe thanks to A. Marshall, B. Dantzer, and J. Mitani for providing valuable
 442 feedback on data analyses. **Funding:** This work was supported by: National Science Foundation (grant
 443 numbers BCS-0715179, IOS-1255974, BCS-1340911, IOS-1854359; BCS-1723237; BCS-17123228;
 444 BCS-2010309); Leakey Foundation (multiple grants); National Institutes of Health (grant number R00-
 445 AG051764); National Geographic Society (grant numbers 8100-06, 8989-11, 50409R-18; the Fulbright
 446 Scholar Program; Sigma Xi; American Society of Primatologists; Wildlife Conservation Society; and
 447 the University of Michigan. **Author contributions:** ETJ, TJB, NSM, and JCB designed the study. ETJ,
 448 AL, TJB, NSM, and JCB collected the data. ETJ, JAF, AL, TJB, NSM, and JCB analyzed the data. ETJ,
 449 JAF, and NSM wrote the paper, with contributions and edits from all authors. ETJ, JAF, AL, TJB, NSM,
 450 and JCB provided funding support. **Competing interests:** Authors declare no competing interests. **Data**
 451 **and materials availability:** All data and code used in this analysis are available at:
 452 https://github.com/GeladaResearchProject/TinsleyJohnson_Feder_et_al_Goldilocks_2021.

453 REFERENCES:

- 454 1. Alexander RD. 1974 The evolution of social behavior. *Annu. Rev. Ecol. Syst.* **5**, 325–383.
- 455 2. Terborgh J, Janson CH. 1986 The Socioecology of Primate Groups. *Annu. Rev. Ecol. Syst.* **17**, 111–
456 136.
- 457 3. Janson CH, Goldsmith ML. 1995 Predicting group size in primates: foraging costs and predation
458 risks. *Behav. Ecol.* **6**, 326–336.
- 459 4. Grove M. 2012 Space, time, and group size: a model of constraints on primate social foraging.
460 *Anim. Behav.* **83**, 411–419.
- 461 5. Markham AC, Gesquiere LR. 2017 Costs and benefits of group living in primates: an energetic
462 perspective. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**, 20160239.
- 463 6. Pulliam HR. 1973 On the advantages of flocking. *J. Theor. Biol.* **38**, 419–422.
- 464 7. van Schaik CP. 1983 Why are diurnal primates living in groups? *Behaviour* **87**, 120–144.
- 465 8. Whiteside MA, Langley EJG, Madden JR. 2016 Males and females differentially adjust vigilance
466 levels as group size increases: effect on optimal group size. *Anim. Behav.* **118**, 11–18.
- 467 9. Rasa OAE. 1989 The costs and effectiveness of vigilance behaviour in the Dwarf Mongoose:
468 implications for fitness and optimal group size. *Ethol. Ecol. Evol.* **1**, 265–282.
- 469 10. Clutton-Brock TH. 1988 *Reproductive success: studies of individual variation in contrasting*
470 *breeding systems*. Chicago: University of Chicago Press.
- 471 11. Giraldeau L-A. 1988 The stable group and the determinants of foraging group size. In *The ecology*
472 *of social behavior* (ed CN Slobodchikoff), pp. 33–53. New York: Academic Press.
- 473 12. Carbone C, Cowlshaw G, Isaac NJB, Rowcliffe JM. 2005 How far do animals go? Determinants of
474 day range in mammals. *Am. Nat.* **165**, 290–297.
- 475 13. Majolo B, de Bortoli Vizioli A, Schino G. 2008 Costs and benefits of group living in primates:
476 group size effects on behaviour and demography. *Anim. Behav.* **76**, 1235–1247.
- 477 14. Markham AC, Gesquiere LR, Alberts SC, Altmann J. 2015 Optimal group size in a highly social
478 mammal. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 14882–14887.
- 479 15. Gesquiere LR, Altmann J, Archie EA, Alberts SC. 2018 Interbirth intervals in wild baboons:
480 environmental predictors and hormonal correlates. *Am. J. Phys. Anthropol.* **166**, 107–126.
- 481 16. Clutton-Brock TH, Albon SD, Guinness FE. 1982 Competition between female relatives in a
482 matrilineal mammal. *Nature* **300**, 178–180.
- 483 17. Borries C, Larney E, Lu A, Ossi K, Koenig A. 2008 Costs of group size: lower developmental and
484 reproductive rates in larger groups of leaf monkeys. *Behav. Ecol.* **19**, 1186–1191.
- 485 18. McComb K, Moss C, Durant SM, Baker L, Sayialel S. 2001 Matriarchs as repositories of social

- 486 knowledge in African elephants. *Science* **292**, 491–494.
- 487 19. Waterman JM. 2006 Delayed maturity, group fission and the limits of group size in female Cape
488 ground squirrels (Sciuridae: *Xerus inauris*). *J. Zool.* **256**, 113–120.
- 489 20. Creel S, Creel NM. 2015 Opposing effects of group size on reproduction and survival in African
490 wild dogs. *Behav. Ecol.* **26**, 1414–1422.
- 491 21. Sibly RM. 1983 Optimal group size is unstable. *Anim. Behav.* **31**, 947–948.
- 492 22. Krause J, Ruxton GD. 2002 *Living in Groups*. Oxford: Oxford University Press.
- 493 23. Clutton-Brock TH, Gaynor D, McIlrath GM, Maccoll ADC, Kansky R, Chadwick P, Manser M,
494 Skinner JD, Brotherton PNM. 1999 Predation, group size and mortality in a cooperative mongoose,
495 *Suricata suricatta*. *J. Anim. Ecol.* **68**, 672–683.
- 496 24. Guindre-Parker S, Rubenstein DR. 2020 Survival benefits of group living in a fluctuating
497 environment. *Am. Nat.* **195**, 1027–1036.
- 498 25. Mosser A, Packer C. 2009 Group territoriality and the benefits of sociality in the African lion,
499 *Panthera leo*. *Anim. Behav.* **78**, 359–370.
- 500 26. Cheney DL, Seyfarth RM. 1987 The influence of intergroup competition on the survival and
501 reproduction of female vervet monkeys. *Behav. Ecol. Sociobiol.* **21**, 375–386.
- 502 27. Harris TR. 2006 Between-group contest competition for food in a highly folivorous population of
503 black and white colobus monkeys (*Colobus guereza*). *Behav. Ecol. Sociobiol.* **61**, 317–329.
- 504 28. Pride RE. 2005 Optimal group size and seasonal stress in ring-tailed lemurs (*Lemur catta*). *Behav.*
505 *Ecol.* **16**, 550–560.
- 506 29. Papageorgiou D, Farine DR. 2020 Group size and composition influence collective movement in a
507 highly social terrestrial bird. *Elife* **9**, e59902.
- 508 30. Williams CK, Lutz RS, Applegate RD. 2003 Optimal group size and northern bobwhite coveys.
509 *Anim. Behav.* **66**, 377–387.
- 510 31. Rudolph K, Fichtel C, Schneider D, Heistermann M, Koch F, Daniel R, Kappeler PM. 2019 One
511 size fits all? Relationships among group size, health, and ecology indicate a lack of an optimal
512 group size in a wild lemur population. *Behav. Ecol. Sociobiol.* **73**, 132.
- 513 32. Isbell LA. 1991 Contest and scramble competition: patterns of female aggression and ranging
514 behavior among primates. *Behav. Ecol.* **2**, 143–155.
- 515 33. Koenig A, Borries C. 2002 Feeding competition and infanticide constrain group size in wild
516 hanuman langurs. *Am. J. Primatol.* **57**, 33–34.
- 517 34. Snaith TV, Chapman CA. 2007 Primate group size and interpreting socioecological models: Do
518 folivores really play by different rules? *Evol. Anthropol.* **16**, 94–106.
- 519 35. Archie EA, Morrison TA, Foley CAH, Moss CJ, Alberts SC. 2006 Dominance rank relationships

- among wild female African elephants, *Loxodonta africana*. *Anim. Behav.* **71**, 117–127.
36. Focardi S, Pecchioli E. 2005 Social cohesion and foraging decrease with group size in fallow deer (*Dama dama*). *Behav. Ecol. Sociobiol.* **59**, 84–91.
37. Steenbeek R, van Schaik CP. 2001 Competition and group size in Thomass langurs (*Presbytis thomasi*): the folivore paradox revisited. *Behav. Ecol. Sociobiol.* **49**, 100–110.
38. Teichroeb JA, Wikberg EC, Bădescu I, Macdonald LJ, Sicotte P. 2012 Infanticide risk and male quality influence optimal group composition for *Colobus vellerosus*. *Behav. Ecol.* **23**, 1348–1359.
39. Treves A, Chapman CA. 1996 Conspecific threat, predation avoidance, and resource defense: implications for grouping in langurs. *Behav. Ecol. Sociobiol.* **39**, 43–53.
40. Jarvey JC, Low BS, Pappano DJ, Bergman TJ, Beehner JC. 2018 Graminivory and fallback foods: annual diet profile of geladas (*Theropithecus gelada*) living in the Simien Mountains National Park, Ethiopia. *Int. J. Primatol.* **39**, 105–126.
41. Fashing PJ, Nguyen N, Venkataraman VV, Kerby JT. 2014 Gelada feeding ecology in an intact ecosystem at Guassa, Ethiopia: variability over time and implications for theropith and hominin dietary evolution. *Am. J. Phys. Anthropol.* **155**, 1–16.
42. Snyder-Mackler N, Alberts SC, Bergman TJ. 2012 Concessions of an alpha male? Cooperative defence and shared reproduction in multi-male primate groups. *Proc. Biol. Sci.* **279**, 3788–3795.
43. Snyder-Mackler N, Beehner JC, Bergman TJ. 2012 Defining higher levels in the multilevel societies of geladas (*Theropithecus gelada*). *Int. J. Primatol.* **30**, 1054–1068.
44. Dunbar RIM. 1986 The social ecology of gelada baboons. In *Ecological aspects of social evolution* (eds DI Rubenstein, RW Wrangham), pp. 332–351. Princeton: Princeton University Press.
45. Roberts EK, Lu A, Bergman TJ, Beehner JC. 2012 A Bruce effect in wild geladas. *Science* **335**, 1222–1225.
46. Tinsley Johnson E, Snyder-Mackler N, Lu A, Bergman TJ, Beehner JC. 2018 Social and ecological drivers of reproductive seasonality in geladas. *Behav. Ecol.* **29**, 574–588.
47. Beehner JC, Bergman TJ. 2008 Infant mortality following male takeovers in wild geladas. *Am. J. Primatol.* **70**, 1152–1159.
48. Bergman TJ, Ho L, Beehner JC. 2009 Chest color and social status in male geladas (*Theropithecus gelada*). *Int. J. Primatol.* **30**, 791–806.
49. Ezenwa VO. 2004 Host social behavior and parasitic infection: a multifactorial approach. *Behavioral Ecology*. **15**, 446–454. (doi:10.1093/beheco/arh028)
50. Koenig A, Borries C, Chalise MK, Winkler P. 1997 Ecology, nutrition, and timing of reproductive events in an Asian primate, the Hanuman langur (*Presbytis entellus*). *J. Zool.* **243**, 215–235.
51. Emery Thompson M, Muller MN, Wrangham RW. 2012 The energetics of lactation and the return to fecundity in wild chimpanzees. *Behav. Ecol.* **23**, 1234–1241.

- 555 52. Schneider-Crease I, Chiou KL, Snyder-Mackler N, Bergman TJ, Beehner JC, Lu A. 2020 Beyond
556 infant death: the hidden costs of male immigration in geladas. *Animal Behaviour*. **159**, 89–95.
557 (doi:10.1016/j.anbehav.2019.11.010)
- 558 53. Roberts EK, Lu A, Bergman TJ, Beehner JC. 2017 Female reproductive parameters in wild geladas
559 (*Theropithecus gelada*). *Int. J. Primatol.* **38**, 1–20.
- 560 54. Bates D, Maechler M, Bolker B, Walker S, Christensen R, Singmann H, Others. 2019 Lme4:
561 Linear mixed-effects models using Eigen and S4 (Version 1.1-20).
- 562 55. R Core Team. 2020 R: A language and environment for statistical computing. R Core Team,
563 Vienna, Austria.
- 564 56. Pusey A. 2012 Magnitude and sources of variation in female reproductive performance. In *The*
565 *evolution of primate societies* (eds JC Mitani, J Call, PM Kappeler, RM Palombit, JB Silk), pp.
566 343–366. Chicago: University of Chicago Press.
- 567 57. Therneau T. 2015 coxme: mixed effects Cox models. R package version 2.2-3. URL: [http://CRAN.](http://CRAN.R-project.org/package=coxme)
568 [R-project.org/package= coxme](http://CRAN.R-project.org/package=coxme)
- 569 58. Wickham H. 2011 ggplot2. *WIREs Comp Stat* **3**, 180–185.
- 570 59. Fox J, Others. 2003 Effect displays in R for generalised linear models. *J. Stat. Softw.* **8**, 1–27.
- 571 60. Foster WA, Treherne JE. 1981 Evidence for the dilution effect in the selfish herd from fish
572 predation on a marine insect. *Nature* **293**, 466–467.
- 573 61. Koenig A. 2000 Competitive regimes in forest-dwelling Hanuman langur females (*Semnopithecus*
574 *entellus*). *Behav. Ecol. Sociobiol.* **48**, 93–109.
- 575 62. Harris TR, Chapman CA, Monfort SL. 2009 Small folivorous primate groups exhibit behavioral
576 and physiological effects of food scarcity. *Behav. Ecol.* **21**, 46–56.
- 577 63. le Roux A, Beehner JC, Bergman TJ. 2011 Female philopatry and dominance patterns in wild
578 geladas. *Am. J. Primatol.* **73**, 422–430.
- 579 64. Jarvey JC. 2016 The importance of underground foods in female gelada (*Theropithecus gelada*)
580 socioecology. University of Michigan. See <https://deepblue.lib.umich.edu/handle/2027.42/134702>.
- 581 65. Snyder-Mackler N, Alberts SC, Bergman TJ. 2014 The socio-genetics of a complex society: female
582 gelada relatedness patterns mirror association patterns in a multilevel society. *Mol. Ecol.* **23**, 6179–
583 6191.
- 584 66. Ostner J, Kappeler PM. 2004 Male life history and the unusual adult sex ratios of redfronted lemur,
585 *Eulemur fulvus rufus*, groups. *Anim. Behav.* **67**, 249–259.
- 586 67. Teichroeb JA, Jack KM. 2017 Alpha male replacements in nonhuman primates: Variability in
587 processes, outcomes, and terminology. *Am. J. Primatol.* **79**, e22674.
- 588 68. Fedigan LM, Jack KM. 2011 Two girls for every boy: The effects of group size and composition on
589 the reproductive success of male and female white-faced capuchins. *Am. J. Phys. Anthropol.* **144**,

590 317–326.