REVIEW

Direct fitness of group living mammals varies with breeding strategy, climate and fitness estimates

Luis A. Ebensperger[1]*, Daniela S. Rivera[1] and Loren D. Hayes[2]

[1]*Centro de Estudios Avanzados en Ecología and Biodiversidad (CASEB), and Departamento de Ecología, Facultad de Ciencias Biológicas, Pontificia Universidad Católica de Chile, Casilla 114-D, Santiago, Chile; and* [2]*Department of Biological and Environmental Sciences, University of Tennessee at Chattanooga, Chattanooga, TN 37403, USA*

Summary

1. Understanding how variation in fitness relates to variation in group living remains critical to determine whether this major aspect of social behaviour is currently adaptive.

2. Available evidence in social mammals aimed to examine this issue remains controversial. Studies show positive (i.e. potentially adaptive), neutral or even negative fitness effects of group living.

3. Attempts to explain this variation rely on intrinsic and extrinsic factors to social groups. Thus, relatively more positive fitness effects are predicted in singularly breeding as opposed to plural breeding species. Fitness effects of sociality in turn may depend on ecological conditions (i.e. extrinsic factors) that influence associated benefits and costs.

4. We used meta-analytic tools to review how breeding strategy or ecological conditions influence the effect size associated with direct fitness-sociality relationships reported in the mammalian literature. Additionally, we determined how taxonomic affiliation of species studied, different fitness and sociality measures used, and major climatic conditions of study sites explained any variation in direct fitness effect size.

5. We found group living had modest, yet positive effects on direct fitness. This generally adaptive scenario was contingent not only upon breeding strategy and climate of study sites, but also on fitness measures examined. Thus, positive and significant effects characterized singular as opposed to plural breeding strategies.

6. We found more positive fitness effects on studies conducted in tropical as opposed to temperate or arid climates. More positive and significant effects were noted on studies that relied on group fecundity, male fecundity and offspring survival as measures of fitness.

7. To conclude, direct fitness consequences of mammalian group living are driven by interspecific differences in breeding strategy and climate conditions. Other factors not examined in this study, namely individual variation in direct and indirect fitness benefits and potential interactions between social and ecological conditions, may be important and require further studies.

Key-words: breeding structure, ecological conditions, fitness, reproductive benefits, social behaviour, sociality

Introduction

Animal social groups range from short-term associations and aggregations (e.g. foraging or roosting groups) to relatively long-term socially cohesive units (e.g. communally rearing or cooperatively breeding groups; Parrish, Hamner & Prewitt 1997; Solomon & Getz 1997; Krause & Ruxton 2002). Sociality is thought to reflect a trade-off between current fitness benefits and costs that emerge from individuals' decision to join or leave groups (Clark & Mangel 1986; Vásquez & Kacelnik 2000; Krause & Ruxton 2002). Benefits include decreased predation risk, decreased burrowing costs, defence against infanticide, increased access to resources and decreased thermoregulatory costs (Alexander 1974; Bertram 1978; Manning *et al.* 1995; Hayes 2000; Ebensperger 2001; Ebensperger & Cofré 2001; Ebensperger & Blumstein 2006). Fitness costs, on the other hand, include increased horizontal transfer of pathogens and ectoparasites, competition over resources and infanticide (Alexander 1974; Hoogland 1995). Mammals that rear young communally or live in large

colonies may experience additional costs of milk theft and pseudo-vertical transfer of pathogens during allonursing (Roulin & Heeb 1999; Hayes 2000).

Given the above potential costs and benefits, an important theme in behavioural ecology research is to determine the fitness consequences of mammalian group living. To understand whether group living is currently adaptive, researchers need to determine how much variation in fitness relates to variation in sociality (i.e. Reeve & Sherman 1993). The striking picture that emerges from the available literature aimed to this end is that fitness consequences of group living vary widely across studies. For instance, some studies have reported the number of offspring produced per year (a common proxy of direct fitness) to increase with group size (a common estimate of sociality; e.g. Boyce & Boyce 1988; Robinson 1988; Takahata *et al.* 1998; Cant 2000; Watts & Holekamp 2009). In other species, the number of offspring produced per female per year does not change (e.g. Cowan 1987; Da Silva, Macdonald & Evans 1994; Stokes, Parnell & Olejniczak 2003; Hsu, Lin & Agoramoorthy 2006) or decreases with sociality (e.g. Hoogland 1981; Watts 1996; McComb *et al.* 2001; Lacey 2004). Efforts to quantify and summarize this variation are needed to establish general patterns in group living and direct fitness.

Attempts to explain and predict differences in how fitness covaries with sociality have emphasized factors that are intrinsic and extrinsic to social groups. Among intrinsic factors, how direct reproduction and offspring care are shared within social groups is thought to influence fitness consequences across social mammals (Silk 2007). Variation in direct reproduction and how offspring are reared within groups has been used to define a continuum of breeding strategies across species (Brown 1987). On one end of this continuum are singularly breeding species in which a male–female adult pair monopolizes breeding, but in which offspring of these breeders are reared communally (e.g. African wild dogs, meerkats). On the other end, plural breeders without communal care include species in which most group members breed but care for their own offspring independently (e.g. ground squirrels, macaques). Plural breeders with communal care are intermediate and consist of species in which most group members breed and offspring are reared communally (e.g. African lions, rabbits). Based on the greater reliance of breeders on nonbreeding group members, Silk (2007) predicted increasingly positive direct fitness consequences of sociality as we move along this continuum from plural breeding without communal care to singular breeding.

Formal tests of Silk's (2007) prediction are lacking and would make a significant contribution to our understanding of the evolutionary significance of mammalian social group living. Data on the fitness consequences of sociality are available for several mammalian species that span the breeding strategy continuum. Most critical, differences between species sharing similar social and breeding strategies have been reported. For instance, while fitness benefits have been observed in some plural breeders with communal care, including carnivores (Cant 2000; Packer, Pusey & Eberly

2001) and rodents (McGuire, Getz & Oli 2002), this strategy is costly (Boyce & Boyce 1988; Da Silva, Macdonald & Evans 1994; Hoogland 1995; Solomon & Crist 2008) or has no effect on fitness (Wolff 1994; Pilastro, Missiaglia & Marin 1996; Randall *et al.* 2005; Hayes *et al.* 2009) in other species. Thus, a more quantitative measure of how fitness consequences of group living vary across studies on species with different breeding strategies is justified.

Our efforts to understand whether group living is currently adaptive in mammals have been challenged by the diversity of life history variables (i.e. other intrinsic factors) chosen to assess lifetime fitness. For a variety of reasons, all of the major components of lifetime fitness, namely adult fecundity and offspring survival to reproduction (Clutton-Brock, Albon & Guinness 1988; King & Allainé 2002), are not always examined. This relatively incomplete information may affect our interpretations because different fitness components may be affected across species or within species across different environmental conditions. On the other hand, the time span of studies in relation to life span of organisms studied may also be of critical importance. Given that short- versus long-term fitness consequences of phenotypic variation may differ (e.g. Sparkman *et al.* 2011), establishing the adaptive value of group living across mammals may be difficult if we need to rely on relatively short-term as opposed to long-term fitness measures.

Extrinsic factors such as the distribution and abundance of food resources, predation risk and thermal conditions (e.g. Ebensperger 2001) may lead to variation in sociality (Emlen & Oring 1977; Brashares & Arcese 2002). Some of these factors may also interact with sociality and influence how fitness relates to sociality. For example, social mammals may benefit the most from group living when conditions are 'mild' (e.g. food is abundant). Under these conditions, competition for resources among group members may be low, allowing breeders and nonbreeders to allocate more energy to offspring care (Harrington, Mech & Fritts 1983). Evidence for this hypothesis comes from studies on wolves in which litter size and pack size are negatively correlated in a population with low prey availability and positively correlated in a population with abundant prey (Harrington, Mech & Fritts 1983). Similarly, Solomon & Crist (2008) observed a positive relationship between per capita reproductive success and group size in food-supplemented prairie vole populations living in semi-natural conditions. If this hypothesis is supported, we predict a positive fitness-sociality relationship in species living in 'mild' (high food abundance, low predation) conditions and negative fitness-sociality relationship in species living in 'harsh' (high predation, low food abundance) conditions. Alternatively, sociality may have evolved to improve reproductive success under conditions that are challenging to reproduction (sensu Hayes & Solomon 2004). In some species, adults are constrained to groups because successful reproduction is not possible without communal care of offspring (Moehlman 1979). This hypothesis predicts positive sociality-fitness relationship when conditions are 'harsh' but not when conditions are 'mild'.

Climate and its temporal effects on productivity have been linked to sociality in mammals through limitations to offspring dispersal and independent breeding. Sociality may be favoured in temperate, seasonally cooler habitats because the time needed for sexual maturity and dispersal of offspring is longer relative to the time of abundant food in the habitat (Barash 1974; Armitage 1981). On the other hand, relatively high aridity conditions may favour sociality though limiting dispersal opportunities (Jarvis *et al.* 1994), or from selection to reduce aggression (and increase social tolerance, Ganem & Nevo 1996). Taken together, available theory suggests one or more links between mammalian group living and climatic variation. However, the same theory is not clear in terms of under what climatic conditions mammalian group living is most advantageous. Compared with arid or temperate environments, temporally stable and warm environments of the tropics may represent milder conditions for social mammals, a hypothesis that has yet to be tested in a comparative context. Thus, meta-analytic analyses of sociality-fitness variation across different climates are needed to determine how environmental harshness influences sociality-fitness covariation.

Herein, we conducted a meta-analysis to examine variability in fitness consequences of group living across published studies on mammalian sociality. In particular, we aimed to determine whether group living is currently adaptive in the mammals. We examined how differences in the breeding strategy (singular vs. plural breeding) and ecological conditions (i.e. natural vs. abundant food vs. decreased predation; across different climates) of studies influence the direct fitness consequences of group living. A meta-analysis uses a common and standardized measure (effect size) to summarize the extent to which a given biological trend is present in a sample of studies. The biological trend of interest in our study was the direct fitness-sociality correlations (e.g. number of offspring per female vs. group size).

Materials and methods

DATA SET

We first searched for relevant studies in the Biological Abstracts and Web of Science data bases for all articles relating fitness and sociality measures. We used the following terms during our search: sociality, fitness, breeding strategy, group size, survival, reproductive success, fecundity, offspring survival, mortality, reproduction. All combinations of fitness- and sociality-related search terms yielded 89 publications that met our criteria of reporting correlations or contrasts between fitness and sociality measures of field-based studies on mammalian species. We only included studies that reported direct measures of survival and fecundity (reproductive success). Studies reporting measures that are more indirect or distantly related to fitness (e.g. based on growth or energy acquisition) were excluded. On a very limited number of comparisons, studies simultaneously reported correlations with and without the effects of territory quality, density, or other covariates. Our decision was to give preference to statistics that controlled for these factors. The following information and variables were extracted from each study: (i) study species and mamma-

lian order, (ii) sociality measures examined, (iii) direct fitness measures, (iv) main breeding strategy of species studied, (v) ecological conditions of studies, (vi) overall climate of study site, (vii) number of social groups or other social entities studied (e.g. families, clans), (viii) the actual number of statistical replicates entered in the original analysis reported, and (ix) statistics and associated *P*-values reported.

Following Brown (1987) and Silk (2007), model species studied were placed into one of three breeding strategies: singular breeder, plural breeder with communal care, and plural breeder without communal care. For species in which social organization is flexible (e.g. Arctic fox, prairie vole), we assigned the species to a breeding strategy based on how adults in multi-female or multi-male groups are commonly reported to share reproduction and parental care.

In several studies, the effect of group size and other sociality measures on fitness is nonlinear, implying beneficial consequences for individuals up to a certain threshold. To account for these patterns, we examined how effect size differed between studies in which authors reported significantly 'positive', 'negative', 'nonlinear' and statistically nonsignificant (i.e. 'no effect') sociality-fitness comparisons.

We aimed to determine how effect size changed with different sociality measures reported as a way to quantify the extent to which different components of social groups (i.e. group structure) have different fitness consequences. To do so, we reduced the multiple terminology used by different studies into five measures: total group size, number of adults, number of females, number of males and number of nonbreeders. Total group size typically included adults and offspring of different ages or developmental stages. The number of nonbreeders category used in this study included a few comparisons that distinguished male ($n = 5$) and female ($n = 7$) nonbreeders. A complete list of sociality measures used as synonyms during our analysis is given in Table S1 of Supporting Information.

We followed a similar strategy to examine major components of fitness. In particular, we classified fitness measures reported into estimates of fecundity and survival. Fecundity measures were (i) adult (male and female) fecundity, (ii) female fecundity, (iii) male fecundity and (iv) group fecundity. While measures (i) through (iii) included per individual or per capita estimates of reproductive success, group fecundity referred to collective success of social groups. Survival components of fitness examined were (v) adult survival (including male and female) and (vi) offspring survival. A complete list of fitness measures used as synonyms during the analysis is given in Table S2 of Supporting Information.

To contrast short- versus long-term fitness measures, we categorized fitness variables reported into these two categories. Fitness measures calculated by authorities before or near weaning time of the studied species, expressed on a seasonal or annual basis, were considered short-term measures. Fitness measures calculated by authorities beyond weaning time and spanning multiple seasons or years were classified as long-term measures.

The most frequent approach used to estimate fitness measures relies on the combined use of demographic information and behavioural records of individuals. Alternatively, studies rely on 'molecular' tools for fitness estimates. Thus, we aimed to determine how effect size differed between 'traditional' and 'molecular' approaches.

We used two methods to evaluate the relationship between ecological variation and direct fitness in social groups. First, we tallied comparisons coming from studies that were explicitly conducted under conditions of 'low food', 'high food', 'low predation' and 'high predation'. Comparisons coming from studies in which authors were not explicit about how food or predation were high or low were included in the analysis under the 'natural' category. Two comparisons that

came from studies conducted under low- and high-density conditions, respectively, were excluded from subsequent analysis because of insufficient replication in the data set. Laboratory-based studies were excluded from the analysis. Secondly, and to examine an influence of main climate of study location, we used three main categories: tropical, temperate (including tundra) and arid (including semi-arid) conditions.

To examine how taxonomy explained variation in sociality-fitness effect size, we distinguished Carnivora, Primates, Rodentia and lumped Artiodactyla, Chiroptera, Lagomorpha, Perissodactyla, Proboscidea into an 'other' category because of the relatively small number of studies available for these mammals. Table 1 summarizes the number of species, studies, and main category of breeding strategy, sociality, fitness and fitness approaches, contributed by each mammalian order represented in the study.

In all analyses (see below), effect sizes from individual comparisons were weighted by their sample size to the proportion of the total sample size in the meta-analysis. The calculation of effect size (see below) requires a minimum number of four associated replicates (Rosenberg, Adams & Gurevitch 2000). Observations on multiple species or fitness measures reported in single publications were considered as independent. The complete data base with all comparisons used in the analysis is given in Table S3 of Supporting Information.

DATA ANALYSIS

An effect size is a standardized measure of the magnitude of the effect of an independent variable (either categorical or continuous) on a dependent variable (Rosenberg, Adams & Gurevitch 2000). We used the Pearson's product-moment correlation coefficient, r, as an appropriate measure of effect size (Hunter, Schmith & Jackson 1982; Rosenthal 1991), because it has been shown recently to be appropriate for analysing results from behavioural and psychological studies (Fiske, Rintamäki & Karvonen 1998; Møller & Saino 2004; Segerstrom & Miller 2004; Stankowich & Blumstein 2005). During our analyses, r represented the magnitude of the effect on fitness measures of varying sociality measures. The benefit of using Pearson correlation coefficients is that these can be calculated even if information on descriptive statistics is lacking (a relatively recurrent situation in the literature, Rosenberg, Adams & Gurevitch 2000).

Following Stankowich & Blumstein (2005), we obtained correlation coefficients from each study, when possible, from direct reporting of r, r^2, or partial correlation coefficients. When these statistics were not reported, we converted other test statistics (e.g. F, U, t, χ^2) to r using methods in Rosenthal (1991), Rosenberg, Adams & Gurevitch (2000) and Rosenthal & Rubin (2003) with the use of MetaWin Statistical Calculator (MetaCalc), a module available in METAWIN. Whenever studies reported no exact *P*-values, these were estimated to the nearest tenth or hundredth decimal place of a given value (e.g. if $P < 0.25$ is reported, a $P = 0.2$ was used; if $P > 0.05$ we used $P = 0.06$; if $P > 0.1$ we used $P = 0.2$). We used METAWIN (version 2) statistical software (Rosenberg, Adams & Gurevitch 2000) to perform all analyses.

All analyses used the random-effects model for categorical data, an algorithm referred to as the mixed effects model (Gurevitch & Hedges 1999). This model is analogous to ANOVA and is based on the

Table 1. Number of species, studies and main category of breeding strategy, sociality measures, fitness measures and fitness approaches contributed by each mammalian order represented in the study. For further explanation or definition of main factor categories, see Materials and Methods

Order	Number of species	Number of studies	Breeding strategy	Sociality	Fitness measure	Temporal category of fitness measure	Approach used to estimate fitness measures
Artiodactyla	1	2	Plural without	# Females	Female fecundity, Offspring survival	Short term; Long term	Traditional
Carnivora	14	36	Plural with; Plural without; Singular	# Adults; # Females; # Male; # Non breeders; Total group size	Adult fecundity; Adult survival; Female fecundity; Group fecundity; Male fecundity; Offspring survival	Short term; Long term	Molecular; Traditional
Chiroptera	3	3	Plural without	# Adults; # Females; Total group size	Adult fecundity; Female fecundity; Male fecundity; Offspring survival	Short term; Long term	Molecular; Traditional
Lagomorpha	1	1	Plural with	# Females	Female fecundity; Group fecundity; Offspring survival	Short term; Long term	Traditional
Perissodactyla	1	4	Plural without	# Females; # Male	Female fecundity; Group fecundity; Male fecundity Offspring survival	Short term; Long term	Molecular; Traditional
Primates	16	25	Plural without; Singular	# Adults; # Females; # Male; # Non breeders; Total group size	Adult fecundity; Female fecundity; Adult survival; Group fecundity; Male fecundity; Offspring survival	Short term; Long term	Traditional
Proboscidea	1	2	Plural with	Total group size; # Females	Female fecundity; Offspring survival	Short term	Traditional
Rodentia	14	22	Plural with; Plural without; Singular	# Adults; # Females; # Male; # Non breeders; Total group size	Adult survival; Female fecundity; Group fecundity; Male fecundity; Offspring survival	Short term; Long term	Molecular; Traditional

more realistic assumption that a given class of studies shares a common effect and that random variation among studies exists (Gurevitch & Hedges 1993; Rosenberg, Adams & Gurevitch 2000). We used bootstrapping approaches to determine whether effect sizes were statistically significant whenever the calculated 95% confidence intervals excluded the zero (Cohen 1969; Verdolin 2006). This resampling method does not assume normality and provide valid sampling distributions for test statistics (Gurevitch & Hedges 1999). When discussing the magnitude of effects, we used the convention whereby effects sizes of 0·2 were 'small', 0·5 were 'medium', and 0·8 were 'large' (Cohen 1969; Verdolin 2006).

The number of statistical replicates used to report fitness-sociality correlations was different from the actual number of social groups studied (i.e. 'social replicates') in 63% of the 308 comparisons included in our analyses. For instance, some studies considered the number of adult females or litters as replicates across social groups (or other social entities such as clans) studied. Therefore, we first verified whether consideration of original replicates used vs. actual social replicates influenced effect size in our data set. We ran two separate analyses, one in which effect size was weighted by statistical replicates and an alternative in which effect size was weighted with the use of social replicates. Given that these analyses rendered identical results, we opted to report results based on statistical replicates.

The influence of intrinsic (e.g. breeding strategy) and extrinsic (e.g. climate) factors on effect size was assessed through analyses of heterogeneity (Huedo-Medina, Sanchez-Meca & Martin-Martinez 2006). Under this approach, heterogeneity of results across comparisons or studies (i.e. the amount of variation in *r*-scores) was estimated by the Q statistic, a measure that partitions total heterogeneity into variance explained by the model (QM) and residual error not explained by the model (QE; i.e. QT = QM + QE; Rosenberg, Adams & Gurevitch 2000). This partitioning is analogous to F in ANOVA tests (Rosenberg, Adams & Gurevitch 2000). Both QM and QE were tested against a X^2-distribution (significance level $P < 0·05$). A statistically significant QM implies that here are differences among cumulative effect sizes for the groups; statistically significant values of QE imply that there is heterogeneity among effect sizes not explained by the model (Rosenberg, Adams & Gurevitch 2000). Upon detecting statistically significant heterogeneity we considered the bootstrapped 95% confidence interval linked to each effect size to determine which categories were different.

Given some statistical shortcomings linked to the Q statistic (Alexander, Scozzaro & Borodkin 1989; Sanchez-Meca & Martin-Martinez 1997; Hardy & Thompson 1998; Higgins *et al.* 2003; Huedo-Medina, Sanchez-Meca & Martin-Martinez 2006), we followed Higgins & Thompson (2002) and Higgins *et al.* (2003) and complemented our Q estimates with reports of the I^2 index. The I^2 index can be interpreted as the percentage of total variability in a set of effect sizes because of true heterogeneity, that is, between-study (or between-comparison) variability. For instance, $I^2 = 50$ means that half of the total variability among effect sizes is caused not by sampling error but by true heterogeneity between studies or comparisons. For each weighted mean *r*, we calculated the fail-safe number of studies, the number of unpublished studies required to bring the meta-analytic mean effect size down to a statistically insignificant level (Rosenthal 1991). The larger the fail-safe number of studies, the greater our confidence in that the observed results are a reliable estimate of the true effect is high (Rosenberg, Adams & Gurevitch 2000).

We assessed the relevance of publication bias by means of the Spearman rank correlation. In particular, we correlated effect size against sample size across studies and comparisons (Rosenberg, Adams & Gurevitch 2000). A significant correlation would indicate bias towards the publication of larger effects in which larger effect sizes are more likely to be published than smaller effect sizes.

Results

GENERAL PATTERN OF SOCIALITY-FITNESS COVARIATION

Our data base of 91 studies rendered 308 sociality-fitness correlations (hereafter referred to as comparison) that were included in the analysis (Table S3 of Supporting Information). The mean effect size linked to sociality-fitness covariation across comparisons reached 0·074 (95% CI: 0·039–0·111), a small yet statistically significant effect size. Splitting the data set among statistical comparisons reported as significantly 'positive', 'negative', 'nonlinear' and statistically nonsignificant, that is, 'no effect' revealed differences in effect size ($Q_{B} = 312·29$; d.f. = 3; $P < 0·001$; Fig. 1). Thus, the effect size was statistically significant and similarly positive in comparisons reported as positive and nonlinear (Fig. 1). The effect size calculated from comparisons reported as negative was statistically significant and different from positive, nonlinear and no effect categories. As expected, effect size associated with comparisons reporting no effects was statistically nonsignificant (Fig. 1).

BREEDING STRATEGY

The effect size of sociality on fitness differed across breeding strategies ($Q_{B} = 12·78$; d.f. = 2; $P = 0·002$; Fig. 2). The effect size for singularly breeding species was positive and statistically different from the effect size calculated for plural breeders with and without communal care (Fig. 2). The effect
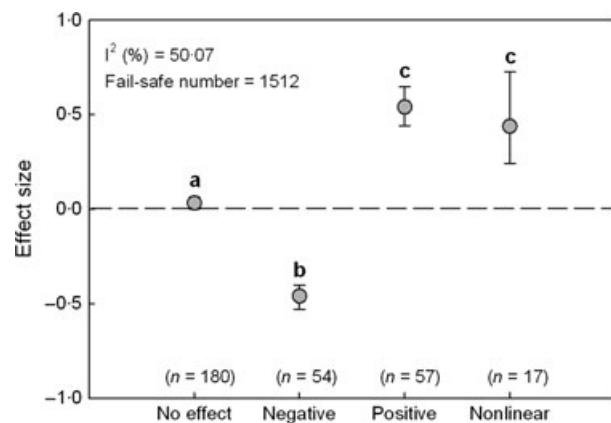

Fig. 1. Mean effect size ($\pm$95% confidence interval) of sociality on direct fitness after reported statistical comparisons is split into comparisons reported as statistically nonsignificant (i.e. 'no effect') and statistically significant 'positive', 'negative', or 'nonlinear'. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 308). In this and all other figures, an effect size is statistically significant whenever the 95% confidence interval excludes the zero. Small capital letters on means are used to indicate which categories are different only after statistically significant heterogeneity has been detected.
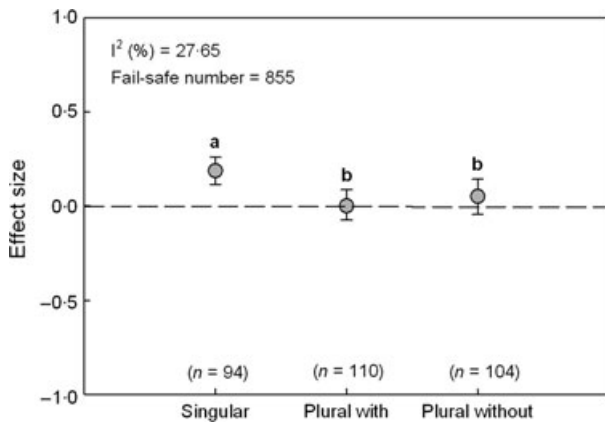
Fig. 2. Mean effect size (±95% confidence interval) of sociality on direct fitness after reported statistical comparisons is split into the prevailing breeding strategy of studied species. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 308). Small capital letters on means only are used to indicate which categories are different after statistically significant heterogeneity has been detected.

size for plural breeders with and without communal care did not differ from zero.

ECOLOGICAL CONDITIONS: NATURAL, LOW FOOD, HIGH FOOD, LOW PREDATION, HIGH PREDATION

Effect size linked to comparisons from studies under natural, high food and high predation conditions was all similarly positive and statistically significant. Thus, the splitting of comparisons into those coming from studies conducted under natural, low food, high food, low predation and high predation revealed no differences in effect size for sociality on fitness ($Q_B = 5.05$; d.f. = 4; $P = 0.282$; Fig. 3).

ECOLOGICAL CONDITIONS: MAIN CLIMATE

The effect size of sociality on fitness differed across major climate of study locations ($Q_B = 7.17$; d.f. = 2; $P = 0.028$; Fig. 4). The effect size related to comparisons from studies on tropical areas was positive and significant (Fig. 4). In contrast, the effect size calculated from studies in temperate or arid regions locations did not differ from zero (Fig. 4).

SOCIALITY (GROUP STRUCTURE) MEASURES

The effect size of sociality on fitness did not differ across sociality measures ($Q_B = 5.083$; d.f. = 4; $P = 0.279$; Fig. 5). Significantly positive effects were recorded for the number of adult members and for the number of nonbreeders in social groups (Fig. 5).

FITNESS MEASURES

The effect size of sociality on fitness was heterogeneous across different fitness measures ($Q_B = 44.82$; d.f. = 5; $P < 0.001$), where the effect size associated with group fecundity
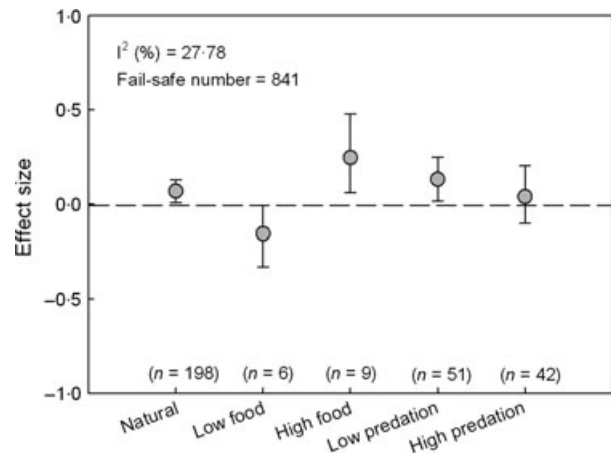

Fig. 3. Mean effect size (±95% confidence interval) of sociality on direct fitness after statistical comparisons reported is split into those coming from studies conducted under 'natural', low food, high food, low predation and high predation conditions. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 306). Two comparisons that came from studies conducted under low- and high-density conditions, respectively, were excluded from analysis because of insufficient replication in the data set. The absence of small letters on means indicate that no statistically significant heterogeneity across categories was detected.
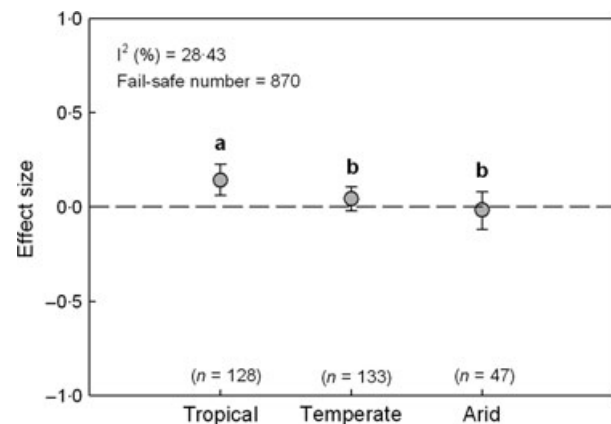

Fig. 4. Mean effect size (±95% confidence interval) of sociality on direct fitness after statistical comparisons reported is split into those coming from studies conducted in tropical, temperate (including tundra) and arid (including semi-arid) habitats. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 308). Small capital letters on means only are used to indicate which categories are different after statistically significant heterogeneity has been detected.

was higher than that of adult fecundity, female fecundity and adult survival, all three of which were also not different from zero (Fig. 6). The effect size for male fecundity and offspring survival was positive and statistically significant (Fig. 6).

TAXONOMIC EFFECTS

Fifty-one species from 41 genera and 8 mammalian orders (Artiodactyla, Carnivora, Chiroptera, Lagomorpha, Perissodactyla, Primates, Proboscidea, Rodentia) were represented
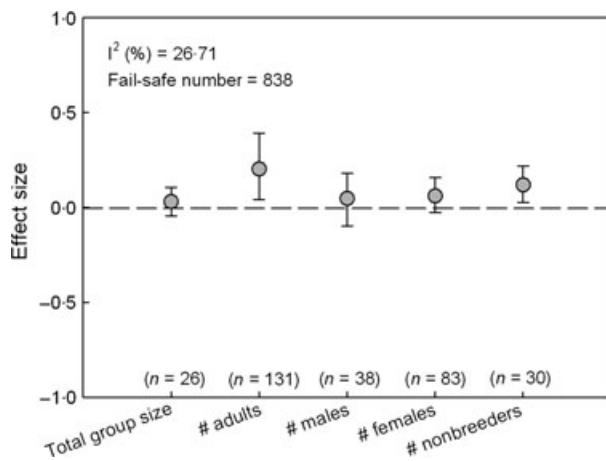
Fig. 5. Mean effect size (±95% confidence interval) of sociality on direct fitness after statistical comparisons reported is split according to sociality measure used. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 308). The absence of small letters on means indicates that no statistically significant heterogeneity across categories has been detected.
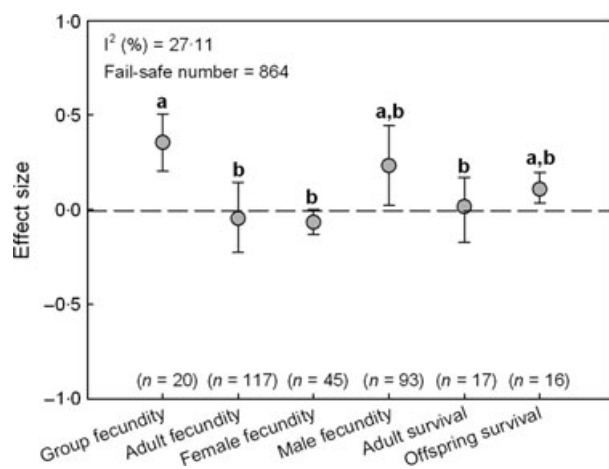

Fig. 7. Mean effect size (±95% confidence interval) of sociality on direct fitness after statistical comparisons reported is split into major taxonomic affiliation (order) of studied species. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 308). The absence of small letters on means indicates that no statistically significant heterogeneity across categories was detected.

Fig. 6. Mean effect size (±95% confidence interval) of sociality on direct fitness after statistical comparisons reported is split according to fitness measure used. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 308). Small capital letters on means only are used to indicate which categories are different after statistically significant heterogeneity has been detected.

Fig. 8. Mean effect size (±95% confidence interval) of sociality on direct fitness after statistical comparisons reported is split according to short- and long-term based fitness measures. Replicates (comparisons) for each category (on the *x*-axis) are given in parentheses (total = 308). The absence of small letters on means indicates that no statistically significant heterogeneity across categories was detected.

in our data base. The splitting of comparisons according to main taxonomic affiliation of model species examined showed no differences in effect size for sociality on fitness (Q_B = 4·13; d.f. = 3; P = 0·248; Fig. 7).

ANALYSIS OF BIAS

Overall, values of I^2 ranged from 24% to 50%, indicating that variation in effect size across factor-level categories exists and needs to be explained. The fail-safe number of studies across comparisons ranged from 838 to 1512 correlations, indicating that a relatively large number of unpublished relationships would be required to change statistically significant effects
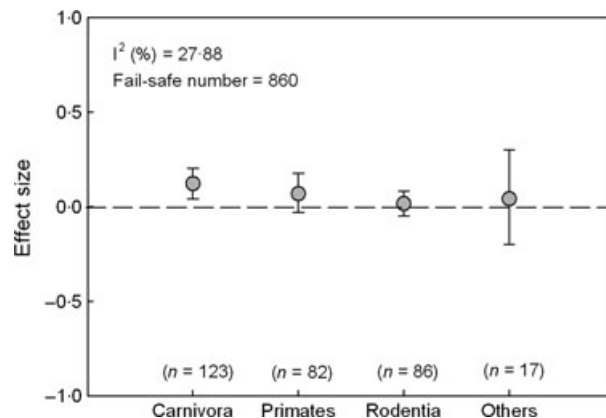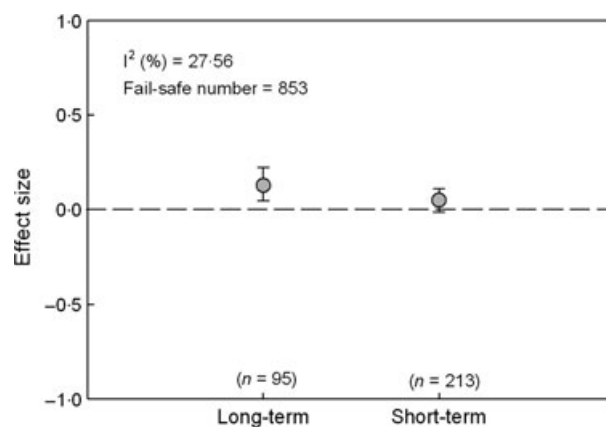
into nonsignificant effects. The Spearman rank correlation test revealed no evidence of publication bias (r_s = −0·020, P = 0·722), meaning that larger effect sizes in one direction (positive or negative) are equally likely to be published compared with smaller effect sizes. Taken together, estimates based on the fail-safe number and on correlation revealed that our analyses are robust against publication bias.

The splitting of comparisons according to short- versus long-term fitness measures involved showed no heterogeneity (Q_B = 2·67; d.f. = 1; P = 0·102), implying that sociality-fitness comparisons were positive irrespective of whether they came from long- or short-term fitness measures (Fig. 8). Similarly, there was not heterogeneity in effect size when comparisons were split according to the approach used to estimate fitness (Q_B = 2·58; d.f. = 1; P = 0·108). Thus, effect size of sociality-fitness comparisons reported in papers using

traditional ($n = 280$) or molecular ($n = 28$) approaches was statistically nonsignificant.

Discussion

How variation in reproductive fitness relates to variation in sociality across species remains critical to establish whether group living is currently adaptive (sensu Reeve & Sherman 1993). The available literature aimed to answer this question is controversial as it includes studies revealing positive (i.e. potentially adaptive; Boyce & Boyce 1988; Robinson 1988; Takahata *et al.* 1998; Cant 2000; Watts & Holekamp 2009), neutral (Cowan 1987; Da Silva, Macdonald & Evans 1994; Stokes, Parnell & Olejniczak 2003; Hsu, Lin & Agoramoorthy 2006), or negative (Hoogland 1981; Watts 1996; McComb *et al.* 2001; Lacey 2004) fitness effects of group living. Moreover, a previous analysis on primates revealed neutral fitness effects of group living across species (Majolo, De Bartolini Vizioli & Schino 2008). Our current analysis of eight mammalian orders, indicated overly positive, yet modest effect of sociality on fitness, implying that sociality is generally adaptive in mammals. When the sign (positive, negative, no effect) and linearity (linear, nonlinear) of comparisons are taken into account, most comparisons (58%, $n = 308$) reveal no effect, a pattern that probably drives the overall small effect of sociality on fitness. These different relationships suggest that several intrinsic and/or extrinsic factors drive species-specific patterns as suggested by Silk (2007). Not surprisingly, our analysis of heterogeneity also revealed that the effect size of sociality on fitness is influenced by breeding strategies of species, climate conditions and fitness estimates reported. In contrast, the analysis did not support that ecological conditions (i.e. food or predation differences), major taxonomic affiliation of species studied, sociality measures reported or the time span associated with fitness estimates influenced effect size of sociality on fitness.

BREEDING STRATEGY

Our study revealed a positive effect size for sociality and fitness in singular breeders but not in the plural breeders with and without communal care. This finding is partially in agreement with Silk's (2007) expectation that positive fitness consequences of sociality should prevail as we move towards singularly breeding species. In contrast to Silk (2007), our analysis failed to confirm an intermediate effect size among plural breeding species in which offspring are reared communally. We caution, however, that a more precise test of Silk's (2007) expectation would benefit from using more quantitative measures of how breeding is shared in social groups in studies of the singular to plural breeding continuum. As few studies provide quantitative measures across the singular–plural breeding continuum, we were precluded from using a more robust analysis. Like most previous studies, our meta-analysis reported results based on mutually exclusive breeding strategy categories. Regardless, our observation that singular, but not plural, breeders experience direct fitness benefits raises an important evolutionary question, why do some mammals form plural breeding groups? One potential answer to this question is that plural breeding confers long-term direct fitness benefits to breeders (Solomon & Hayes 2009; Dugdale, Ellwood & Macdonald 2010). Currently, our observation that the effect size linked to long-term direct fitness benefits is not significantly greater than the effect size linked to short-term direct fitness benefits suggests that both breeding strategies do not enhance long-term direct fitness. Alternatively, communally nursing females may do a better job of enhancing the survival and future reproduction of offspring produced by closely related kin than do nonbreeders in singularly breeding groups. A test of this hypothesis in flexibly social species such as voles (Lucia *et al.* 2008) or striped mice (Schradin & Pillay 2005), albeit logistically challenging, is necessary to determine how important kin selection and inclusive fitness are in the evolution of mammalian plural breeding.

ECOLOGICAL CONDITIONS

Contrary to previous studies on canids (Harrington, Mech & Fritts 1983) and rodents (Solomon & Crist 2008), our analysis did not support the hypothesis that the benefits of group living are greatest when conditions are 'mild', that is, food availability is high and predation risk low. Based on this observation, we argue that Harrington, Mech & Fritts (1983) hypothesis according to which group members allocate resources to offspring based on the availability of food is not generally supported. Our analysis also did not support the alternate hypothesis that group living enhances reproductive success when conditions are challenging or 'harsh' (Moehlman 1979). On the surface, this observation counters previous arguments that sociality is constrained by the environment in which species are evolving (Moehlman 1979; Jarvis *et al.* 1994). However, the lack of habitat-specific sociality-fitness trends suggests that other factors that could vary between environments, such as thermal conditions or parasite loads, need to be examined. It is possible that in singular breeders group living could also ensure fitness in harsh environments if nonbreeders provide care to nondescendent kin.

In contrast, we found more positive fitness effects from comparisons and studies conducted in tropical as opposed to temperate or arid climates. It is possible that the tropics represent warm, temporarily more stable environments compared with arid or temperate regimes. Group living under these conditions would be relatively advantageous and not the consequence of dispersal limitations. Alternatively, some of these positive trends may have been driven in part by the fact that some mammalian orders such as primates contribute with generally positive fitness effects and also are commonly associated with tropical regions.

TAXONOMIC AFFILIATION

Major taxonomic affiliation of species did not explain variation in effect size of sociality on fitness. However, some

caution is needed given that most comparisons coming from singularly breeding mammals have been conducted on carnivores, the order in which fitness effects are relatively more positive. Similarly, comparisons involving nonbreeders as the sociality measure come from studies on carnivores. Unfortunately, our ability to examine factor interactions or to isolate main factor effects is limited by the nature of the data available. For instance, data reported would need to come from a same study in a well-replicated factorial design (Gurevitch, Morrison & Hedges 2000).

Subsequent studies on under-represented mammals such as large-sized marine (i.e. pinnipeds, cetaceans) and small-sized volant (i.e. Chiroptera) species are needed to increase taxonomic representation. Most pinnipeds, cetaceans and chiropterans are social animals that would fit to plural breeders without communal care (Altringham 1996; Whitehead & Mann 2000). These species are found under widely different habitat conditions (Whitehead & Mann 2000; Simmons & Conway 2003), a feature that may further clarify a potential social by environment interaction on fitness.

CAVEATS

Our analysis is the first major step towards a comprehensive understanding of the evolutionary significance of mammalian sociality. However, our conclusions are limited by the nature of data in the literature and methodological constraints of meta-analyses (e.g. no factor interactions). First, our analysis relied on studies of plural breeders reporting the 'mean' or per capita direct fitness of all group members, estimates that do not capture individual variation in direct fitness in the same social group. Although egalitarian reproduction is observed in some carnivores (Packer, Pusey & Eberly 2001), within-group variation in the direct fitness occurs in other plural breeders (Pilastro, Missiaglia & Marin 1996; Dugdale *et al.* 2008). Future analyses of variation in direct fitness across species will allow researchers to examine how differences between group members, such as dominance status (Altmann & Alberts 2003), affects sociality-fitness covariation. Second, we did not include metrics of sociality that account for the extent and nature of social interactions within groups such as social cohesion (including social bonds) of group members (Krause & Ruxton 2002; Silk 2007) in our analysis. Recent advances in social network theory (Wey *et al.* 2008; Whitehead 2008) will improve our ability to quantify these variables, improving the quality of future meta-analyses. Third, our analysis of climatic conditions dealt with spatial, but not temporal, variation in environmental conditions. Thus, the scale of the available data might be inappropriate to determine if sociality is currently adaptive for some species, or a nonadaptive response to temporal environmental variation. Future analyses need to test the hypothesis that flexible social strategies observed in some social mammals may have evolved to maximize reproductive success in response to seasonal or interannual variation in thermal conditions (Madison 1984; Johnson, Jetz & Macdonald 2002) and availability of food resources (Woodroffe

& Macdonald 1993). Finally, our method of analysis cannot rule out the impact of as historical predation pressures (Byers 1997; Blumstein 2006) and historical differences in food resources (Furuichi 2009) on sociality and fitness. Thus, our interpretations are limited to the current utility (or current adaptive significance) of group living.

Conclusions

Despite its limitations, our study revealed that the underlying sources of variation in direct fitness of social mammals are complex, with some variation explained by breeding strategy of species, climatic conditions and fitness measures examined. Together, these observations imply that the fitness consequences of group living are influenced by intrinsic (life history) and extrinsic (climatic) factors of species studied, as suggested by Silk (2007). Our study did not support that studies relying on molecular tools used to assign paternity or maternity during direct fitness estimates, or on long-term direct fitness estimates determine greater effect sizes compared with traditional studies that rely more heavily on demographical and behavioural records, or on short-term direct fitness measures. These findings imply that researchers need to determine the direct fitness consequences of group living using multiple measures of direct fitness and environments to provide general conclusions. To advance theory, we need to determine the relationship between inclusive fitness and sociality measures. However, these advances will require greater emphasis on the development of methods to quantify communal care and models to estimate indirect fitness in natural populations.

References

Alexander, R.D. (1974) The evolution of social behavior. *Annual Review of Ecology, Evolution and Systematics*, **5**, 325–383.

Alexander, R.A., Scozzaro, M.J. & Borodkin, L.J. (1989) Statistical and empirical examination of the chi-square test for homogeneity of correlations in meta-analysis. *Psychological Bulletin*, **106**, 329–331.

Altmann, J. & Alberts, S.C. (2003) Intraspecific variability in fertility and offspring survival in a nonhuman primate: behavioural control of ecological and social sources. *Offspring: The Biodemography of Fertility and Family Behavior* (eds K.W. Wachter & R.A. Bulatao), pp. 140–169. National Academy Press, Washington, District of Columbia, USA.

Altringham, J.D. (1996) *Bats: Biology and Behaviour*. Oxford University Press, Oxford, UK.

Armitage, K.B. (1981) Sociality as a life history tactic of ground squirrels. *Oecologia*, **48**, 36–49.

Barash, D.P. (1974) The evolution of marmot societies: a general theory. *Science*, **185**, 415–420.

Bertram, B.C.R. (1978) Living in groups: predators and prey. *Behavioural Ecology: And Evolutionary Approach* (eds J.R. Krebs & N.B. Davies), pp. 64–96. Blackwell Scientific Publications, Oxford, UK.

Blumstein, D.T. (2006) The multipredator hypothesis and the evolutionary persistence of antipredator behavior. *Ethology*, **112**, 209–217.

Boyce, C.C.K. & Boyce, J.L. (1988) Population biology of *Microtus arvalis*. I. Lifetime reproductive success of solitary and grouped breeding females. *Journal of Animal Ecology*, **57**, 711–722.

Brashares, J.S. & Arcese, P. (2002) Role of forage, habitat and predation in the behavioral plasticity of a small African antelope. *Journal of Animal Ecology*, **71**, 626–638.

Brown, J.L. (1987) *Helping and Communal Breeding in Birds*. Princeton University Press, Princeton, New Jersey, USA.

Byers, J.A. (1997) *American Pronghorn. Social Adaptations and the Ghosts of Predators Past*. University of Chicago Press, Chicago, Illinois, USA.

Cant, M.A. (2000) Social control of reproduction in banded mongooses. *Animal Behaviour*, **59**, 147–158.

Clark, C.W. & Mangel, M. (1986) The evolutionary advantages of group foraging. *Theoretical Population Biology*, **30**, 45–75.

Clutton-Brock, T.H., Albon, S.D. & Guinness, F.E. (1988) Reproductive success in male and female red deer. *Reproductive Success: Studies of Individual Variation in Contrasting Breeding Systems* (ed. T.H. Clutton-Brock), pp. 325–343. University of Chicago Press, Chicago, Illinois, USA.

Cohen, J. (1969) *Statistical Power Analysis for the Behavioural Sciences*. Academic Press, New York, USA.

Cowan, D.P. (1987) Group living in the European rabbit (*Oryctolagus cuniculus*): mutual benefit or resource localization? *Journal of Animal Ecology*, **56**, 779–795.

Da Silva, J., Macdonald, D.W. & Evans, P.G.H. (1994) Net costs of group living in a solitary forager, the Eurasian badger (*Meles meles*). *Behavioral Ecology*, **5**, 151–158.

Dugdale, H.L., Ellwood, S.A. & Macdonald, D.W. (2010) Alloparental behavior and long-term cost of mother tolerating other members of the group in a plurally breeding mammal. *Animal Behaviour*, **80**, 721–735.

Dugdale, H.L., Macdonald, D.W., Pope, L.C., Johnson, P.J. & Burke, T. (2008) Reproductive skew and relatedness in social groups of European badgers, *Meles meles*. *Molecular Ecology*, **17**, 1815–1827.

Ebensperger, L.A. (2001) A review of the evolutionary causes of rodent group-living. *Acta Theriologica*, **46**, 115–144.

Ebensperger, L.A. & Blumstein, D.T. (2006) Sociality in New World hystricognath rodents is linked to predators and burrow digging. *Behavioral Ecology*, **17**, 410–418.

Ebensperger, L.A. & Cofré, H. (2001) On the evolution of group-living in the New World cursorial hystricognath rodents. *Behavioral Ecology*, **12**, 227–236.

Emlen, S.T. & Oring, L.W. (1977) Ecology, sexual selection, and the evolution of mating systems. *Science*, **197**, 215–223.

Fiske, P., Rintamäki, P.T. & Karvonen, E. (1998) Mating success in lekking males: a meta-analysis. *Behavioral Ecology*, **9**, 328–338.

Furuichi, T. (2009) Factors underlying party size differences between chimpanzees and bonobos: a review and hypotheses for future study. *Primates*, **50**, 197–209.

Ganem, G. & Nevo, E. (1996) Ecophysiological constraints associated with aggression, and evolution toward pacifism in *Spalax ehrenbergi*. *Behavioral Ecology and Sociobiology*, **38**, 245–252.

Gurevitch, J. & Hedges, L.V. (1993) Meta-analysis: combining the results of independent experiments. *Design and Analysis of Ecological Experiments*, 1st edn (eds S.M. Scheiner & J. Gurevitch), pp. 378–398. Champan and Hall, New York, USA.

Gurevitch, J. & Hedges, L.V. (1999) Statistical issues in ecological meta-analyses. *Ecology*, **80**, 1142–1149.

Gurevitch, J., Morrison, J.A. & Hedges, L.V. (2000) The interaction between competition and predation: a meta-analysis of field experiments. *The American Naturalist*, **155**, 435–453.

Hardy, L.V. & Thompson, S.G. (1998) Detecting and describing heterogeneity in meta-analysis. *Statistics in Medicine*, **17**, 841–856.

Harrington, F.H., Mech, L.D. & Fritts, S.H. (1983) Pack size and wolf pup survival: their relationship under varying ecological conditions. *Behavioral Ecology and Sociobiology*, **13**, 19–26.

Hayes, L.D. (2000) To nest communally or not to nest communally: a review of rodent communal nesting and nursing. *Animal Behaviour*, **59**, 677–688.

Hayes, L.D., Chesh, A.S., Castro, R.A., Tolhuysen, L.O., Burger, J.R., Bhattacharjee, J. & Ebensperger, L.A. (2009) Fitness consequences of group living in the degu *Octodon degus*, a plural breeder rodent with communal care. *Animal Behaviour*, **78**, 131–139.

Hayes, L.D. & Solomon, N.G. (2004) Costs and benefits of communal rearing to female prairie voles (*Microtus ochrogaster*). *Behavioral Ecology and Sociobiology*, **56**, 585–593.

Higgins, J.P.T. & Thompson, S.G. (2002) Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, **21**, 1539–1558.

Higgins, J.P.T., Thompson, S.G., Deeks, J.J. & Altman, D.G. (2003) Measuring inconsistency in meta- analyses. *British Medical Journal*, **327**, 557–560.

Hoogland, J.L. (1981) Nepotism and cooperative breeding in the black-tailed prairie dog (Sciuridae: *Cynomys ludovicianus*). *Natural Selection and Social Behavior: Recent Research and New Theory* (eds R.D. Alexander & D.W. Tinkle), pp. 283–310. Chiron Press, New York, USA.

Hoogland, J.L. (1995) *The Black-Tailed Prairie Dog: Social Life of a Burrowing Mammal*. University of Chicago Press, Chicago, Illinois, USA.

Hsu, M.J., Lin, J.F. & Agoramoorthy, G. (2006) Effects of group size on birth rate, infant mortality and social interactions in Formosan macaques at Mt. Longevity, Taiwan. *Ethology Ecology and Evolution*, **18**, 3–17.

Huedo-Medina, T.B., Sanchez-Meca, J. & Martin-Martinez, F. (2006) Assessing Heterogeneity in Meta-Analysis: q statistic or I2 index? *Psychological Methods*, **11**, 193–206.

Hunter, J.E., Schmith, F.L. & Jackson, G.B. (1982) *Meta-Analysis: Cumulating Research Findings across Studies*. Sage Publications, Beverlys Hills, California, USA.

Jarvis, J.U.M., O'Riain, M.J., Bennett, N.C. & Sherman, P.W. (1994) Mammalian eusociality: a family affair. *Trends in Ecology and Evolution*, **9**, 47–51.

Johnson, D.D.P., Jetz, W. & Macdonald, D.W. (2002) Environmental correlates of badger social spacing across Europe. *Journal of Biogeography*, **29**, 411–425.

King, W.J. & Allainé, D. (2002) Social, maternal, and environmental influences in reproductive success in female Alpine marmots (*Marmota marmota*). *Canadian Journal of Zoology*, **80**, 2137–2143.

Krause, J. & Ruxton, G.D. (2002) *Living in Groups*. Oxford University Press, Oxford, UK.

Lacey, E.A. (2004) Sociality reduces individual fitness in a communally breeding rodent, the colonial tuco-tuco (*Ctenomys sociabilis*). *Behavioral Ecology and Sociobiology*, **56**, 449–457.

Lucia, K.E., Keane, B., Hayes, L.D., Lin, Y.K., Schaefer, R.L. & Solomon, N.G. (2008) Philopatry in prairie voles: an evaluation of the habitat saturation hypothesis. *Behavioral Ecology*, **19**, 774–783.

Madison, D.M. (1984) Group nesting and its ecological and evolutionary significance in overwintering microtine rodents. *Winter Ecology of Small Mammals* (ed. J.F. Merritt), pp. 267–274. Carnegie Museum of Natural History, Special Publication 10, Pennsylvania, USA.

Majolo, B., De Bartolini Vizioli, A. & Schino, G. (2008) Cost and benefits of group living in primates: group size effects on behaviour and demography. *Animal Behaviour*, **76**, 1235–1247.

Manning, C.J., Dewsbury, D.A., Wakeland, E.K. & Potts, W.K. (1995) Communal nesting and communal nursing in house mice, *Mus musculus domesticus*. *Animal Behaviour*, **50**, 741–751.

McComb, K., Moss, C., Durant, S.M., Baker, L. & Sayialel, S. (2001) Matriarchs as repositories of social knowledge in African elephants. *Science*, **292**, 491–494.

McGuire, B., Getz, L.L. & Oli, M.K. (2002) Fitness consequences of sociality in prairie voles, *Microtus ochrogaster*: influence of group size and composition. *Animal Behaviour*, **64**, 645–654.

Moehlman, P.D. (1979) Jackal helpers and pup survival. *Nature*, **277**, 382–383.

Møller, A.P. & Saino, N. (2004) Immune response and survival. *Oikos*, **104**, 299–304.

Packer, C., Pusey, A.E. & Eberly, L.E. (2001) Egalitarism in female African lions. *Science*, **293**, 690–693.

Parrish, J.K., Hamner, W.M. & Prewitt, C.T. (1997) Introduction – From individuals to aggregations: unifying properties, global framework, and the holy grails of congregation. *Animal Groups in Three Dimensions* (eds J.K. Parrish & W. Hamner), pp. 1–13. Cambridge University Press, Cambridge, UK.

Pilastro, A., Missiaglia, E. & Marin, G. (1996) Age-related reproductive success in solitarily and communally nesting female dormice (*Glis glis*). *Journal of Zoology, London*, **239**, 601–608.

Randall, J.A., Rogovin, K., Parker, P.G. & Eimes, J.A. (2005) Flexible social structure of a desert rodent, *Rhombomys opimus*: philopatry, kinship, and ecological constraints. *Behavioral Ecology*, **16**, 961–973.

Reeve, H.K. & Sherman, P.W. (1993) Adaptation and the goals of evolutionary research. *Quarterly Review of Biology*, **68**, 1–32.

Robinson, J.G. (1988) Group size in wedge-capped capuchin monkeys *Cebus olivaceus* and the reproductive success of males and females. *Behavioral Ecology and Sociobiology*, **23**, 187–197.

Rosenberg, M.S., Adams, D.C. & Gurevitch, J. (2000) *MetaWin: Statistical Software for Meta-Analysis*, Version 2. Sinauer Associates, Sunderland, Massachusetts, USA.

Rosenthal, R. (1991) *Meta-Analytic Procedures for Social Research*. Sage Publications, Newbury Park, California, USA.

Rosenthal, R. & Rubin, D.B. (2003) r equivalent: a simple effect size Indicator. *Psychological Methods*, **8**, 492–496.

Roulin, A. & Heeb, P. (1999) The immunological function of allosuckling. *Ecology Letters*, **2**, 319–324.

Sanchez-Meca, J. & Martin-Martinez, F. (1997) Homogeneity test in meta-analysis: a Monte Carlo comparison of statistical power and Type I error. *Quality and Quantity*, **31**, 385–399.

Schradin, C. & Pillay, N. (2005) Intraspecific variation in the spatial and social organization of the African striped mouse. *Journal of Mammalogy*, **86**, 99–107.

Segerstrom, S.C. & Miller, G.E. (2004) Psychological stress and the human immune system: a meta-analytic study of 30 years of inquiry. *Psychological Bulletin*, **130**, 601–630.

Silk, J.B. (2007) The adaptive value of sociality in mammalian groups. *Philosophical Transactions of the Royal Society B, Biological Sciences*, **362**, 539–559.

Simmons, N.B. & Conway, T.M. (2003) Evolution of ecological diversity of bats. *Bat Ecology* (eds T.H. Kunz & M. Brock Fenton), pp. 493–535. University of Chicago Press, Chicago, Illinois, USA.

Solomon, N.G. & Crist, T.O. (2008) Estimates of reproductive success for group-living prairie voles *Microtus ochrogaster*, in high-density populations. *Animal behavior*, **76**, 881–892.

Solomon, N.G. & Getz, L.L. (1997) Examination of alternative hypotheses for cooperative breeding in rodents. *Cooperative Breeding in Mammals* (eds N.G. Solomon & J.A. French), pp. 199–230. Cambridge University Press, Cambridge, UK.

Solomon, N.G. & Hayes, L.D. (2009) The biological basis of alloparental behaviour in mammals. *Substitute Parents: Biological and Social Perspectives on Alloparenting in Human Societies* (eds G. Bentley & R. Mace), pp. 13–49. Berghahn Books, Herndon, Virginia, USA.

Sparkman, A.M., Adams, J., Beyer, A., Steury, T.D., Waits, L. & Murray, D.L. (2011) Helper effects on pup lifetime fitness in the cooperatively breeding red wolf (*Canis rufus*). *Proceedings of the Royal Society of London, Biological Series*, **278**, 1381–1389.

Stankowich, T. & Blumstein, D.T. (2005) Fear in animals: a meta-analysis and review of risk assessment. *Proceeding of the Royal Society of London, Biological Sciences*, **272**, 2627–2634.

Stokes, E.J., Parnell, R.J. & Olejniczak, C. (2003) Female dispersal and reproductive success in wild western lowland gorillas (*Gorilla gorilla gorilla*). *Behavioral Ecology and Sociobiology*, **54**, 329–339.

Takahata, Y., Suzuki, S., Okayasu, N., Sugiura, H., Takahashi, H., Yamagiwa, J., Izawa, K., Agetsuma, N., Hill, D., Saito, C., Sato, S., Tanaka, T. & Sprague, D. (1998) Does troop size of wild Japanese macaques influence birth rate and infant mortality in the absence of predators? *Primates*, **39**, 245–251.

Vásquez, R.A. & Kacelnik, A. (2000) Foraging rate versus sociality in the starling *Sturnus vulgaris*. *Proceedings of the Royal Society of London, Biological Series*, **267**, 157–164.

Verdolin, J.L. (2006) Meta-analysis of foraging and predation risk trade-offs in terrestrial systems. *Behavioral Ecology and Sociobiology*, **60**, 457–464.

Watts, D.P. (1996) Comparative socioecology of gorillas. *Great Ape Societies* (eds W.C. McGrew, L.F. Marchant & T. Nishida), pp. 16–28. Cambridge University Press, Cambridge, UK.

Watts, H.E. & Holekamp, K.E. (2009) Ecological determinants of survival and reproduction in the spotted hyena. *Journal of Mammalogy*, **90**, 461–471.

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

Whitehead, H. (2008) *Analyzing Animal Societies: Quantitative Methods for Vertebrate Social Analysis*. Chicago University Press, Chicago, Illinois, USA.

Whitehead, H. & Mann, J. (2000) Female reproductive strategies of cetaceans. *Cetacean Societies: Field Studies of Dolphins and Whales* (eds J. Mann, R.C. Connor, P.L. Tyack & H. Whitehead), pp. 219–246. University of Chicago Press, Chicago, Illinois, USA.

Wolff, J.O. (1994) Reproductive success of solitarily and communally nesting white-footed and deer mice. *Behavioral Ecology*, **5**, 206–209.

Woodroffe, R. & Macdonald, D.W. (1993) Badger sociality–models of spatial grouping. *Symposia of the Zoological Society of London*, **65**, 145–169.

Supporting information

Additional Supporting Information may be found in the online version of this article.

Table S1. Sociality measures reported in the literature that were considered synonyms of sociality main categories used in the analyses.

Table S2. Fitness measures reported in the literature that were considered synonyms of fitness main categories used in the analyses.

Table S3. Sample size, coefficient of correlation, effect size, sample size, and factor categories for each comparison reported by different authorities.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.