Research

A standardized approach to estimate life history tradeoffs in evolutionary ecology

Sandra Hamel, Nigel Gilles Yoccoz and Jean-Michel Gaillard

S. Hamel (sandra.hamel@uit.no) and N. G. Yoccoz, Dept. of Arctic and Marine Biology, Faculty of Biosciences, Fisheries and Economics, Univ. of Tromsø, NO–9037 Tromsø, Norway. – J.-M. Gaillard, CNRS, UMR 5558 "Biométrie et Biologie Evolutive", Univ. de Lyon, FR-69622, Villeurbanne, France.

As tradeoffs limit the maximum Darwinian fitness individuals can reach, measuring reliably the strength of tradeoffs using appropriate metrics is of prime importance to understand the evolution of traits under constraints. Tradeoffs involving phenotypic traits and fitness components, however, are difficult to quantify in free-ranging populations because of confounding effects due to environmental variation and individual heterogeneity. Furthermore, although some methods have been used previously to quantify tradeoffs, these methods cannot be applied with respect to binary traits, which are common to describe life histories (e.g. probability of reproduction, nesting success, offspring survival). Here, we demonstrate how to measure reliably the strength of tradeoffs involving binary traits using (auto)correlation estimates obtained from generalized linear (mixed) models. We first propose a standardized approach that accounts for the variation in the nature of the tradeoffs being compared (e.g. continuous/binary traits, repeated/non-repeated measures), and then apply this method to longitudinal data from two contrasting species of large herbivores. Empirical estimates of tradeoffs varied among traits, and between-species comparisons suggested that reproductive tradeoffs between successive breeding attempts might only occur in capital breeders. The empirical results we obtained clearly demonstrate that the method we provide allows measuring reliably the strength of tradeoffs under most circumstances, including tradeoffs on binary traits. Our original approach therefore offers an important first step for comparing the strength and, hence, the relative importance of different tradeoffs, and opens the door to a better understanding of the evolution of life history traits in free-ranging populations.

A major goal of life history studies is to understand how natural selection shapes individual fitness-related traits, such as growth, reproduction, and survival (Roff 1992, Stearns 1992). Without any resource limitation, a 'Darwinian demon', would evolve (Law 1979). Nevertheless, resources are usually limited so that individuals have to allocate resources to a given function at the cost of other functions (Williams 1966). According to this principle of energy allocation, the maximum fitness individuals can reach is limited by tradeoffs that occur among fitness components during a lifetime (Williams 1966). Although tradeoffs among life history traits have shaped the diversity of life history strategies we currently observe, tradeoffs involving phenotypic traits and fitness components are difficult to measure in free-ranging populations for two main reasons. First, tradeoff measures are subject to confounding effects of environmental variation in resource availability (van Noordwijk and de Jong 1986). Second, tradeoffs are usually assessed from traits measured on different individuals within a population, leading to between-individual variation, and the data generally consist of repeated measures (i.e. recurrent events, Zhao et al. 2012) of a set of individuals sampled at different ages, leading to within-individual variation. As tradeoffs are likely

to vary with age and among individuals, both a marked individual heterogeneity and age-specific changes in individual performance can lead to biased estimates of tradeoffs when not properly corrected for (van de Pol and Verhulst 2006). To understand the evolution of traits under constraints, which is a central issue in evolutionary ecology (Roff 1992, Stearns 1992), measuring reliably the strength of tradeoffs using appropriate metrics is of prime importance.

Tradeoffs can be looked for between two (or more – Roff and Fairbairn 2007) distinct life history traits. These traits can be measured at the same period (e.g. mass vs number of offspring produced at a given reproductive occasion; Smith and Fretwell 1974, Lloyd 1987), or at different periods (e.g. mass of offspring produced at one reproductive occasion vs the number of offspring produced at the next reproductive occasion; Clutton-Brock 1991) of the life cycle. They also can occur for the same trait but measured at successive periods of the life cycle (recurrent events, e.g. offspring mass at one reproductive occasion vs offspring mass at the next reproductive occasion; Clutton-Brock 1991), or between two traits that are not repeated in time (unique events, e.g. number of offspring produced at a given age vs longevity; Reiter and Le Boeuf 1991, Descamps et al. 2006).

Typically, the presence of a tradeoff is looked for by using either a hypothesis-based testing (Fisher 1956) or a model selection approach (Burnham and Anderson 2002) to assess the statistical significance of the link between two traits (or one trait at one period of the life cycle and the same trait later on) or to compare the relative support of models including and excluding a potential tradeoff, respectively. Since the influential syntheses on life history tradeoffs written by Stearns (1992) and Roff (1992, 2002), a large number of studies have used either approaches to demonstrate the occurrence of many tradeoffs (e.g. number vs size of offspring (Smith and Fretwell 1974, Smith et al. 1989, Wilson et al. 2009), current reproduction vs future body condition (Pomeroy et al. 1999, Skjæraasen et al. 2010), age at first reproduction and longevity (Charmantier et al. 2006, Descamps et al. 2006)), thereby providing substantial empirical evidence for their existence in free-ranging populations.

Nevertheless, studies so far have concentrated on detecting tradeoffs by assessing the regression slope between two traits, but these slope estimates are not often comparable among studies. Indeed, life history traits follow a large range of distributions, and so standardized measures of tradeoffs that can be compared among traits are not straightforward. For instance, some traits like body mass, body size, mass gain, or their log-transformations, are continuous and model residuals are approximately normally distributed, so they are usually modelled using linear models (LM). Other traits, like survival and several reproductive traits, often follow a Bernoulli process (e.g. probability of reproduction, nesting success, offspring survival – these traits will be referred here as binary traits, Table 2) that can be modelled through generalized linear models (GLM; Nelder and Wedderburn 1972). In GLM, the relationship between the mean response and the predictors is linearized by a transformation, which is usually the function logit in the case of a Bernoulli process (Table 2). Therefore, parameters estimated from GLM, and hence the quantification of the tradeoff between two traits, are harder to interpret and standardize than those from LM. As a result, direct comparisons of regression slopes from life history studies are usually inappropriate to measure the strength of tradeoffs.

Measuring the strength of a tradeoff, however, would be highly valuable for determining its relative importance (Glazier 2000). One potential measure to quantify tradeoffs is to use the covariance. In genetic for instance, the simple covariance between two traits used to be calculated to measure genetic tradeoffs. Recently, however, more sophisticated models have been developed, which allow determining the covariance among traits while accounting for covariates and individual heterogeneity (Hadfield 2010, Wilson et al. 2010), which can both importantly influence covariance. While these methods are very promising, they cannot be applied to binary traits. Indeed, the covariance is estimated from the within-individual variance (i.e. variance among residuals, σ^2_{RES} in Eq. 1a, Table 1) of the two traits. The residual variance, however, is not identifiable in binary data (because such data follow a Bernoulli distribution (Table 2) and thereby have a fixed variance and cannot be under- or overdispersed; Molenberghs et al. 2012) and this variance is therefore 'fixed' to 1. This constrains the covariance to be 0. The only variance that can be estimated

is the between-individual variance (i.e. variance among individuals, σ^2_{ID} in Eq. 1a, Table 1), which is thus a combined measure of both within- and between-individual variances that are impossible to tease apart. Analyses of binary data are thus more complex and often more limited compared with other types of data (i.e. binomial – Table 2, Poisson, normal). This issue, however, is often swept under the carpet in methodological articles, which usually simply mention that the binary case is just a simpler example of the binomial case. It is true that the Bernoulli distribution is a special case of the binomial distribution since it describes the success of a single event rather than of numerous events (Table 2), but the reduction to a single event constrains the variance and complicates analyses (see Molenberghs et al. 2012 for an explicit discussion of this issue). Nevertheless, binary traits are common in life history, and therefore these analytical issues are important when trying to measure the strength of life history tradeoffs because tradeoff comparisons have to include these central traits. Yet, if tradeoffs could be accurately measured and then standardized across a large range of life history traits and environments, we could envision performing comparative analyses of specific case studies to assess the relative role of various tradeoffs in shaping the evolution of life history strategies across species.

In this study, we use an original correlation-based approach based on generalized linear (mixed) models to measure the strength of life history tradeoffs involving binary traits. To our knowledge, no study has yet focused on estimating the strength of tradeoffs using an accurate and standardized method that can also be applied to binary traits. We focus on tradeoffs involving phenotypic traits and fitness components because they are more prone to environmental influences, but this method could similarly be applied to genetic tradeoffs. Previously, we performed a simulation study to evaluate the statistical precision and accuracy of parameters estimated from models suited for the analysis of longitudinal data (Hamel et al. 2012; see also Martin et al. 2011). Based on that study, two models could provide reliable correlation estimates (a 'time series' model and a 'state dependent' model; Table 1). Here, we first use these two models to estimate between-trait correlations from longitudinal data on two large mammalian herbivores, typical of long-term studies available on vertebrates (Clutton-Brock and Sheldon 2010). We then measure life history tradeoffs and report the first quantitative comparison of tradeoff intensity among a set of life history traits including binary traits. This quantitative analysis demonstrates how standardized correlation measures from complex models can allow estimating the strength of different types of tradeoffs.

Material and methods

Correlation-based approach to measure tradeoffs

The solution we propose here to measure tradeoffs involves measuring the correlation between the two traits. Because correlations are standardized values ranging between −1 and 1, a highly negative correlation would correspond to a stronger tradeoff than a weak negative or any positive

Table 1. Description of the time series and state dependence models.

Description	Model
The 'time series' model:	
For a continuous and normally distributed variable, like individual mass in consecutive years, a time series model would evaluate how mass, y_{ij}, changes with age, x_{Aj}, while assessing the autocorrelation ρ among residuals ε_{ij}. The sample would consist of observations from n_{IND} individuals with n_{AGE} repeated measurements over the lifetime (i.e. at each age). The linear mixed model with a first-order autoregressive error process with parameter ρ takes the following form of Eq. 1a. This model can include age effects of various functional forms, $f(x_{Aj})$, and can also allow the inclusion of covariates, $f(x_{Cj})$. In addition to the mean population intercept β_0, the model includes a random intercept α_i that corresponds to heterogeneity among individuals and allows fitting a different intercept for each individual. This accounts for bias occurring on either the parameter estimate or its uncertainty when $\sigma^2_{ID} > 0$ and the sample includes repeated measures. For a binary variable, such as offspring survival measured at successive breeding attempts for an individual female, a time series model would evaluate variation in offspring survival, y_{ij}, in relation with mother age, x_{Aj}, while assessing the autocorrelation ρ between the successive realizations y_{ij}. The generalized linear mixed model with a first-order autoregressive process modeling a binary variable takes the following form of Eq. 1b.	Equation 1a: $y_{ij} = \beta_0 + f(x_{Aj}) + f(x_{Cj}) + \alpha_i + \varepsilon_{ij}$ $\alpha_i \sim N(0, \sigma^2_{ID})$ $\varepsilon_{ij} \sim N(0, \sigma^2_{RES})$ $\text{Corr}[\varepsilon_{ij}, \varepsilon_{ij'}] \sim \rho^{|j'-j|}$ (for $j = 1, \ldots, n_{AGE}$ and $i = 1, \ldots, n_{IND}$) Equation 1b: $\text{logit}(\text{Pr}(y_{ij} = 1 \mid \alpha_i)) = \beta_0 + f(x_{Aj}) + f(x_{Cj}) + \alpha_i$ $\alpha_i \sim N(0, \sigma^2_{ID})$ $\text{Corr}[y_{ij}, y_{ij'}] \sim \rho^{|j'-j|}$ (for $j = 1, \ldots, n_{AGE}$ and $i = 1, \ldots, n_{IND}$)
The 'state dependence' model:	
The state dependence model is commonly used by econometricians to model first-order autocorrelation and individual heterogeneity (Heckman 2001, Wooldridge 2005, Berridge and Crouchley 2011). It allows assessing the influence of the state of a variable in the past at time t, on the expression of the same variable at the current time (i.e. modeling dependence on previous state). For estimating the correlation within a single variable, the response and explanatory variables take the form of y_{ij} and $y_{i,j-t'}$ respectively. When the response follows a Bernoulli process, this model takes the form of a multilevel logistic regression, where β_T measures the correlation between the states, as expressed in Eq. 2a. When estimating the correlation between two different traits, however, the response and explanatory variables become y_{ij} and $x_{i,j-t'}$ respectively, as shown in Eq. 2b.	Equation 2a: $\text{logit}(\text{Pr}(y_{ij} = 1 \mid \alpha_i)) = \beta_0 + \beta_T y_{i,j-t} + f(x_{Aj}) + f(x_{Cj}) + \alpha_i$ $\alpha_i \sim N(0, \sigma^2_{ID})$ (for $j = 1, \ldots, n_{AGE}$ and $i = 1, \ldots, n_{IND}$) Equation 2b: $\text{logit}(\text{Pr}(y_{ij} = 1 \mid \alpha_i)) = \beta_0 + \beta_T x_{i,j-t} + f(x_{Aj}) + f(x_{Cj}) + \alpha_i$ $\alpha_i \sim N(0, \sigma^2_{ID})$ (for $j = 1, \ldots, n_{AGE}$ and $i = 1, \ldots, n_{IND}$)
The state dependence model can also estimate the correlation for two different traits that represent unique events, such as between age at first reproduction and longevity, and therefore estimate the strength of lifetime tradeoffs. Lifetime tradeoffs are not measured on repeated measures on individuals, and hence individual variance doesn't need to be accounted for. When the response follows a normally distributed process, this model takes the form of a simple linear regression presented in Eq. 3.	Equation 3: $y = \beta_0 + \beta_T x + f(x_C) + \varepsilon$

correlation (assuming that the values describing each trait are both positively correlated with fitness – otherwise signs can be changed, see the example regarding age at first reproduction below). Furthermore, autocorrelation can be applied to estimate the correlation of successive trait values, and could be used to measure a tradeoff for a trait that changes with time. As life history strategies encompass a large number of traits that can be each influenced by a large number of covariates (e.g. age) and confounding factors (e.g. environmental variation), simply using the correlation between two traits or the autocorrelation among successive measures of a given trait while neglecting these effects is likely to be flawed (Nakagawa and Cuthill 2007).

Furthermore, several studies have demonstrated the need for longitudinal rather than cross-sectional data to assess reliably the response of life history traits to biological processes, like aging (Nussey et al. 2008). Such analyses require long-term data sets consisting of repeated measures on individuals, and are thereby affected by individual heterogeneity (i.e. the among-individual variance). Therefore, to obtain reliable correlation values between traits to estimate tradeoffs, we need to extract correlation values from models that can allow the integration of complex effects while accounting for individual heterogeneity at the same time (van de Pol and Verhulst 2006, van de Pol and Wright 2009).

Table 2. Distinctions between binary data, Bernoulli distribution, and binomial distribution.

Binary data	The term binary refers to the nature of the data. The data are described by binary variables that can take only two possible values. They are used to represent the outcomes of Bernoulli events/trials.
Bernoulli distribution	The term Bernoulli refers to the distribution of the data. It is a discrete probability distribution, where a success (value of 1) takes place with a probability p, and a failure (value of 0) occurs with a probability of $q = 1 - p$.
Binomial distribution	A binomial distribution is the distribution of the sum of n independent Bernoulli events, each having the same probability p. It is therefore a discrete probability distribution of the number of successes out of n trials/events. Each event produces a success with a probability of p, and the variance is $np(1 - p)$.
Bernoulli vs binomial	The Bernoulli distribution is a special case of the binomial distribution with $n = 1$. The variance of the Bernoulli distribution, however, is necessarily $p(1 - p)$ (Molenberghs et al. 2012, p1−2), and is thereby constrained, which is not the case for proportions ($n > 1$).

Consequently, the first model we considered is a 'time series' model in which repeated measures of an individual as it ages represent a time series for which both the autocorrelation ρ, measuring the tradeoff, and the individual variance have to be modelled (Eq. 1a, b). The second one is a 'state dependence' model that allows assessing the influence of the state of a variable in the past at time t, on the expression of the same variable at the current time – i.e. modelling dependence on previous state (Eq. 2a). Whereas ρ estimates the autocorrelation in the time series model, it is the coefficient between the two states, β_T (Table 1), that allows estimating the correlation in the state dependence model. For normally distributed traits, β_T will be similar to ρ if both the response and explanatory variables are standardized (Schielzeth 2010). For binary traits, however, β_T needs to be standardized according to $\beta_T \times ([f'(pv)] \times [2.5 \times pv])^{-1}$ (where f' is the derivative and pv is the mean probability value) to be equivalent to ρ and to β_T estimated for normally distributed traits (Hamel et al. 2012).

The time series model has the advantage of modelling the autocorrelation directly and thereby provides the most accurate correlation estimates. The major disadvantage, however, is that the time series model can only measure a tradeoff within a single trait measured at successive periods of the life cycle, and so tradeoffs between two different recurrent traits cannot be analyzed. Compared with the time series model, the state dependence model provides slightly less accurate correlation estimates because it models the autocorrelation indirectly (Hamel et al. 2012), but it offers the possibility to estimate correlation between two different recurrent traits (Eq. 2b). In addition, it can estimate the correlation for two different traits that are unique events, such as between age at first reproduction and longevity, and therefore estimate the strength of lifetime tradeoffs (Eq. 3).

Correlation constraint on binary data

As mentioned earlier, binary data are often more complex to analyse than other types of data because their variance is constrained. Another difficulty with binary data is the existence of a mathematical constraint on the maximum negative correlation that can occur. The maximum correlation is constrained by the mean probability (Prentice 1988, Chaganty and Harry 2006), and by the standard deviation representing heterogeneity among individuals (Supplementary material Appendix A1). For example, strong negative autocorrelation values cannot occur for very high or very low mean probabilities. Furthermore, for a given probability, the possibility for a strong negative autocorrelation to occur decreases as the standard deviation of the between-individual variance (representing individual heterogeneity) increases. Importantly, this means that the maximum value of a tradeoff that we can estimate, or that can occur in a population, is constrained by both the variance among individuals and the mean population value of the trait. Consequently, the definition and the measure of the strength of a tradeoff must take into account that the maximum value for a tradeoff does vary with individual heterogeneity and mean population value of traits involved in that tradeoff. This mathematical constraint for binary data has been previously acknowledged in the literature. When measuring the variability of demographic rates,

Sim et al. (2011) accounted for such a constraint by using the ratio of the observed variance to the maximum variance possible for a certain probability (see also Gaillard and Yoccoz 2003, Morris and Doak 2004). Likewise, using the ratio of the observed (auto)correlation (ρ_{OBS} for time series models or standardized β_T for state dependence models) on the absolute value of the maximum (auto)correlation possible for a specific variance and mean population probability (ρ_{MAX} determined from simulations; Supplementary material Appendix A1) should provide an accurate measure of the strength of a tradeoff. We need to use $|\rho_{MAX}|$ to maintain the sign of the relationship between the traits.

Therefore, we will use two approaches and discuss their advantages and limitations. The first approach accounts for the variation in the nature of the traits (e.g. continuous versus binary traits, repeated versus non-repeated measures) involved in the tradeoffs to be compared, but it ignores the constraint on binary traits. In the second approach, we will account for this constraint by adjusting the correlation using the ratio of the observed correlation to ρ_{MAX}, i.e. adjusted correlation = ($\rho_{OBS}/|\rho_{MAX}|$) or (standardized $\beta_T/|\rho_{MAX}|$).

Populations and traits

We used life history data from three intensively monitored populations of large herbivores, two populations of roe deer *Capreolus capreolus* and one of mountain goats *Oreamnos americanus*, to measure the strength of life history tradeoffs using our proposed approach. These data sets had previously been analyzed for other purposes, thereby providing us with prior knowledge of the covariates needed in the models and of the occurrence or not of a tradeoff (Gaillard et al. 2003a, Hamel et al. 2009a, b, Plard et al. 2012). We selected various traits to estimate both tradeoffs within a single trait measured at successive periods of the life cycle, which correspond to recurrent tradeoffs that can be estimated by both time series and state dependence models, and tradeoffs between two different traits, including recurrent and lifetime tradeoffs, which can only be estimated using state dependence models. We included life history traits well approximated by normal and Bernoulli distributions, with and without missing values. The diversity of tradeoffs we chose was intended to demonstrate the potential and limitation of the approach we are proposing. To model lifetime tradeoffs, we ran the state dependence model using the R function 'lmer' from the package 'lme4' (Bates et al. 2012), which fits a generalized linear or non-linear mixed model with multivariate normal random effects. Recurrent tradeoffs, however, included an age effect and we wanted to model age non-linearly to avoid potential spurious positive correlation that can arise from an inappropriate modelling of the fixed effects (Supplementary material Appendix A2, Wakefield 2007, Hodges and Reich 2010). We therefore ran the time series and state dependence models for recurrent tradeoffs using the R function 'gamm' from the package 'mgcv' (Wood 2006), which allows fitting generalized additive mixed models estimated by penalized quasi-likelihood (PQL) maximization, using the 'glmmPQL' function. Although PQL can sometimes provide biased estimates (Bolker et al. 2009), we used 'glmmPQL' because it is one of the rare functions that allow estimating simultaneously auto-correlated errors and random effects (Zhang

et al. 2011) and we have shown that this function can provide reliable auto(correlation) estimates (Hamel et al. 2012).

Recurrent tradeoffs in roe deer: mass and probability of pregnancy

Roe deer pregnancy data come from the Chizé population in France, whereas the mass data come from the Trois Fontaines population also in France (see Gaillard et al. 2003a, b for more details on both populations). In the winter of each year between 1988 and 2011, marked females of known age were captured, weighed (Gaillard et al. 1998), and pregnancy diagnostics were performed (Gaillard et al. 2003a), thereby providing individual mass and pregnancy time series. We did not need to adjust mass measurements to a certain date because females were all weighed at the same time of the year and roe deer do not show seasonal variation in body mass (Andersen et al. 2000). Not all females were captured every year (the capture probability was around 0.5; Gaillard et al. 2003a). As the state dependence models required at least two values in consecutive years to include a female in the analysis, we only kept time series for females with at least two consecutive values recorded. The resulting data set consisted of about 40% missing values for body mass and 32% for pregnancy. Because a large proportion of missing data ($\geq 25\%$) can introduce bias when estimating (auto)correlation and individual heterogeneity (Hamel et al. 2012), we built reduced data sets that included only females with a time series including 20% or less missing values to minimize bias (see also Nakagawa and Freckleton 2008 for other solutions for accounting for missing values). Therefore, the length of the time series for the reduced mass data set at Trois Fontaines ranged between 2 and 12 (mean $n_{AGE} = 5.0$; $n_{IND} = 118$; $n_{TOTAL} = 394$), whereas that at Chizé for pregnancy ranged between 2 and 16 (mean $n_{AGE} = 5.8$; $n_{IND} = 75$; $n_{TOTAL} = 252$), and both data sets had less than 10% missing values overall. For both mass and pregnancy, we included cohort, population density, and a non-linear age effect as covariates (Gaillard et al. 2003b, Hamel et al. 2009b, 2010a, Nilsen et al. 2009). Overall, we measured two potential tradeoffs based on recurrent traits in roe deer. The first one assessed the influence of mass measured at a given year on mass measured the following year, traits approximated by a normal distribution. A negative correlation would indicate a tradeoff: individuals allocating more to growth one year are unable to allocate as much the following year, leading to compensatory growth. The second tradeoff we looked for was between pregnancy at time t and time $t - 1$, binary traits following a Bernoulli distribution. We thus tested for the occurrence of reproductive pauses (sensu Cameron 1994) interspersed among pregnancies during the female lifetime. The mean pregnancy rate was 0.89, leading to a value to standardize β_T (i.e. $[f'(pv)] \times [2.5 \times pv]$) of 22.72.

Recurrent tradeoffs in mountain goats: parturition and offspring survival probabilities

Mountain goat data come from the Caw Ridge population in Canada (more details in Hamel et al. 2010a). Each year between 1988 and 2011, reproduction of individually marked females of known age and survival of their offspring

were recorded (Hamel et al. 2010a). Therefore, we had access to time series on parturition – if a female gave birth or not each year – and on offspring survival – if the single offspring produced by a female each year survived or not from birth to one year of age. The data set on parturition contained no missing value because the reproductive status of each female was recorded every year (Hamel et al. 2010a), and so $n_{IND} = 89$ and $n_{TOTAL} = 683$ for this data set. Because female mountain goats do not give birth every year, however, the data set on offspring survival contained about 25% missing values ($n_{IND} = 89$; $n_{TOTAL} = 510$), representing years when females skipped reproduction. The length of time series for both data sets ranged between 2 and 15 (mean $n_{AGE} = 7.8$). We used a non-linear age effect in the models and included yearly environmental covariates (population density, North Pacific index, and normalized difference vegetation index, NDVI) that were previously shown to influence female reproduction in this mountain goat population (Hamel et al. 2010a). We therefore measured two potential tradeoffs based on recurrent binary measures in goats. The first one was between the probabilities of parturition at time t and time $t - 1$, and the second was between the probability of offspring survival at time t and parturition at $t - 1$. As the mean probability for parturition was 0.75 and that for offspring survival was 0.63, we used a value of 9.99 and 6.76, respectively, for standardizing β_T.

Lifetime tradeoffs in mountain goats: growth, first reproduction, and survival

For mountain goats, body mass, annual reproduction, and survival of each female were followed intensively from birth to death (details in Hamel et al. 2009b, 2010a). We therefore had access to lifetime data to measure tradeoffs on unique-event traits, i.e. lifetime tradeoffs. Because age at primiparity varied from 3 to 7 in goats, we could estimate whether there was a lifetime tradeoff between allocation to growth, measured as mass at one year old, and the timing of first reproduction (n = 36). Because the mass of mountain goats varies a lot within the year, we adjusted masses to 15 July (as described in Hamel et al. 2010a). As the existence of a tradeoff between mass at one year old and age at primiparity would result in a positive correlation, we inverted age at primiparity and used '– age at primiparity'. Age at primiparity therefore represented an index of 'early primiparity', with individuals starting reproduction at a younger age having highest values. If a tradeoff between growth and age at primiparity occurs, the model would thus estimate the negative correlation commonly expected for tradeoffs. On the other hand, a positive correlation would show that individuals allocating more to growth also start reproducing earlier. Therefore, a positive correlation among lifetime traits would represent a quality effect (Weladji et al. 2008). This would not be the case for recurrent tradeoffs, because the variation among individuals can be estimated from the repeated measurements and is accounted for in the random effect included in the model (a positive correlation in recurrent traits would illustrate a 'long-term memory', Supplementary material Appendix A2). Moreover, we evaluated whether a lifetime tradeoff occurred between reproduction and survival by looking for a negative relationship between

early primiparity and longevity (n = 81), as expected from the disposable soma theory of aging (Kirkwood 1977). Finally, as mountain goat females start reproducing before they reach asymptotic mass (at seven years of age; Côté and Festa-Bianchet 2001) and reproduction lowers summer mass gain (Hamel and Côté 2009), we also evaluated if there was a lifetime tradeoff between age at primiparity and asymptotic adult mass (n = 70) (as shown for example in bison *Bison bison*; Green and Rothstein 1991). We included the cohort year as a random effect in all models evaluating lifetime tradeoffs.

Results

First approach: no correction for the constraint on binary traits

For the three recurrent tradeoffs estimated within a single trait (i.e. mass, pregnancy and parturition) measured at different periods of the life cycle, the age-specific variation was modelled with both the time series (TS) and the state dependence (SD) models. Both models provided similar measures, except for roe deer mass for which there was a difference (correlation estimates ± standard errors: mass, TS = 0 [−0.08; 0.08], SD = 0.26 [0.22; 0.30]; pregnancy, TS = 0.14 [0.06; 0.21], SD = 0.06 [0.03; 0.09]; parturition, TS = −0.11 [−0.07; −0.15], SD = −0.08 [−0.06; −0.10]; Fig. 1A). This result, however, is not too surprising considering that the data set on roe deer mass included missing values. Missing values can lead to severe underestimation of the correlation in the time series model and, to a lower extent, a slight overestimation in the state dependence model (Hamel et al. 2012). This suggests that the accurate correlation value for mass laid between the two models' estimates, likely closer to the state dependence than the time series estimate. In the present case, such slight biases do not influence the interpretation. Overall, results using both methods clearly indicated that no tradeoff occurred in successive measurements of a given female for both mass and pregnancy in roe deer (Fig. 1A). In mountain goats, parturition showed statistical evidence of a tradeoff with future parturition (Fig. 1A), as well as with offspring survival (SD = −0.04 [−0.005; −0.08]; Fig. 1A). In both cases, however, the biological significance of these tradeoffs was weak, especially for offspring survival with a correlation close to 0.

Among the three potential lifetime tradeoffs evaluated in mountain goats, a positive correlation occurred between mass as a yearling and early primiparity (SD = 0.20 [0.01; 0.39]; Fig. 1A), indicating the absence of any tradeoff. This supports a quality effect because females allocating less to body growth did not allocate more to reproduction, but they rather had to delay their first reproduction compared with heavier females that were able to start reproducing earlier. It is important to note, however, that although the mean correlation was quite high, the confidence interval was large compared with the other correlation estimates (Fig. 1A). The low precision for this correlation estimate is explained by the small sample size. Opposite to the tradeoff hypothesis, the mean correlation between early primiparity and longevity was slightly positive, although low and non-statistically

significant (SD = 0.03 [−0.06; 0.10]; Fig. 1A) On the other hand, a tradeoff tended to occur between early primiparity and adult mass (SD = −0.06 [−0.19; 0.06]; Fig. 1A), but again this effect was weak and non-statistically significant.

Second approach: correcting for the constraint on binary traits

Out of the three binary traits we analyzed, only two traits had to be corrected for the constraint because positive correlations (as found for pregnancy in roe deer) are not constrained (Supplementary material Appendix A1). We therefore only applied a correction to parturition and offspring survival in mountain goats. For parturition, the mean probability was 0.75 and the standard deviation measuring individual heterogeneity was 0.84, leading to a simulated ρ_{MAX} of −0.16. In Fig. 1B, the correlation estimates corrected using ρ_{MAX} are presented in red. Based on the time series model, the parturition tradeoff measured using the first approach represented 69% of the maximum tradeoff possible to be observed for the level of individual heterogeneity and the mean probability recorded for this trait ($\rho_{OBS}/|\rho_{MAX}|$ = −0.11 / 0.16 = −0.69, Fig. 1B). Thus, the correlation reported for this tradeoff with the second approach was −0.69 [−0.94; −0.44] (and −0.50 [−0.62; −0.38] for the state dependence model), much stronger than that obtained with the first approach. For offspring survival, the mean probability was 0.63 and the standard deviation was 0.01, leading to a ρ_{MAX} of −0.57. This ρ_{MAX} was much higher than for the tradeoff on parturition because the mean probability was close to 0.5 and the standard deviation was close to 0 (Supplementary material Appendix A1). Therefore, the tradeoff measured for offspring survival with the first approach represented 7% of the maximum tradeoff possible to be observed for this trait (Fig. 1B). The corrected correlation for the second approach was −0.07 [−0.14; 0.00], similar to that obtained from the first approach. Overall, both approaches did not differ in identifying the tradeoffs, but the strength of the tradeoff varied quite a lot (Fig. 1B), being placed in perspective to the maximum tradeoff possible to be observed when using the second approach.

Discussion

We presented an original correlation-based statistical approach to estimate life history tradeoffs in two contrasted species of large herbivores. Although the method we proposed has its own limitations, we demonstrated the large potential in using first-order autocorrelation and standardized correlation coefficients from generalized and linear mixed models to measure the strength of life history tradeoffs. Furthermore, for traits that are not binary, our method provides similar estimates (Supplementary material Appendix A3 Fig. A3) of the strength of tradeoffs compared with the MCMCglmm covariance method (Hadfield 2010), which is an effective method that has been used to estimate tradeoffs but that cannot be applied to binary traits. Most importantly, our method allows estimating the strength of tradeoffs on binary traits, for which many analytical issues exist and thereby limit the possibilities to include these traits in comparative
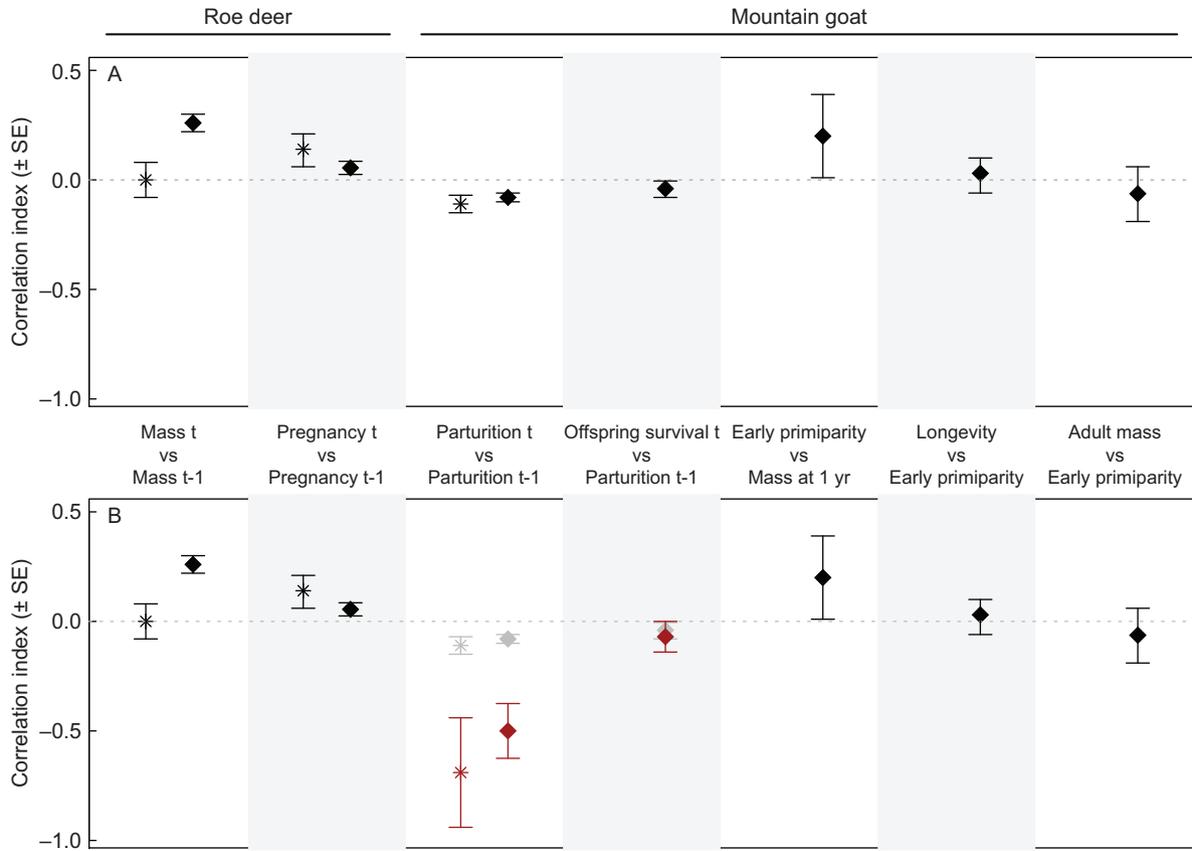
Figure 1. Comparison of the strength of tradeoffs among different life history traits for roe deer and mountain goats, using two standardized correlation approaches. The first approach (A) accounts for variation in the nature of the tradeoffs being compared, i.e. linear (mass, early primiparity index and longevity) versus binary (pregnancy, parturition and offspring survival) distributed traits, and repeated (the first four columns) versus non-repeated (the last three columns) traits. Diamonds are the means β_T values estimated from the state dependence model and standardized according to Hamel et al. (2012). Stars are the means ρ values estimated from the time series model. This model only provides tradeoff estimates for the first three traits because this model can only estimate ρ for a tradeoff occurring within a single trait measured at successive periods of the life cycle. Whiskers represent standard errors. The second approach (B) also takes into account that the negative correlation within binary data is constrained by the mean probability and by the level of among-individual variation for traits that are repeated in time. Therefore, correlations estimated for binary traits have been corrected (in red) using the ratio of the observed correlation (in grey; first approach) on $|\rho_{MAX}|$ estimated by simulations, and standard errors have been adjusted using the ratio of the observed lower or upper bound on $|\rho_{MAX}|$. No correction was applied for pregnancy because the correlation was positive and positive correlations do not share this constraint.

analyses. Because binary traits are common in life history and tradeoff comparisons have to include these central traits, the method we presented is very promising. Furthermore, although our focus here is on binary traits, the method we propose can also be applied to traits following other distributions as long as the appropriate standardisation for β_T is used. For instance, we demonstrate in the Supplementary material Appendix A4 how our method can be applied for tradeoffs on traits following a Poisson distribution (e.g. litter size, lifetime reproductive success). In such case, the response and explanatory variables cannot be standardized before the analyses, as is required for normal traits, so the standardization needs to be done after analyses by multiplying the β_T estimated with the variance of the trait (Supplementary material Appendix A4).

The first approach we proposed builds on the demonstration that the first-order autocorrelations estimated from the time series models are similar to the β_T coefficients estimated from the state dependence models, for both normal and binary traits, if 1) both the response and explanatory

variables are standardized and 2) the estimated β_T is divided by $[f'(pv)] \times [2.5 \times pv]$ when the trait is binary (see details in Hamel et al. 2012 – a simulation study that evaluated the reliability of different parameters estimated by many statistical models). In the current study, we found that the strengths of the tradeoffs were similar for the time series and the state dependence models. Nevertheless, the strength of the correlation differed between the two models when data included missing values, and we recommend that the choice of the model to be used to estimate a correlation should be made in the light of the suggestions provided by the simulation study (see Hamel et al. 2012 for a detailed discussion on the potential biases resulting from missing values, as well as limitations for specific data and population structures with respect to correlation estimates obtained from the statistical models used in the current study). Since tradeoffs are estimated from correlations obtained from these models, it is imperative that their interpretation is made according to the models' limitations. One fundamental aspect to always keep in mind is that temporal autocorrelation can result from

an inappropriate modelling of the fixed effects, for instance by neglecting to account for some kind of environmental variation or modelling age inadequately (as illustrated in the Supplementary material Appendix A2). A good biological understanding is therefore required to model properly both the form of dependence and the fixed effects (Wakefield 2007, Hodges and Reich 2010).

Constraints on binary traits are well acknowledged in the literature and are still a source of debate as to which is the best strategy to standardize estimates (Link and Doherty 2002). Although everyone agrees that some adjustments must be applied, some proposed a correction using the ratio of observed variance on the maximum achievable variance (Gaillard and Yoccoz 2003, Sim et al. 2011), while others proposed using the variance on the logit scale (Link and Doherty 2002). Obviously, each strategy has its pros and cons, and the best strategy will depend on the question we are trying to answer. In our case, the problem with the constraint on binary traits is that the correlation estimating a tradeoff will be dependent on the mean probability and the among-individual variation, so that a tradeoff estimated for a binary trait close to 0 or 1 will necessarily be low. Thus, as the probability value approaches 0 or 1, the maximum (auto)correlation that can be reached gets smaller. This can be interpreted in two ways. The tradeoff could be truly small, so the value observed is not really constrained, a case covered by our first approach (Fig. 1A). A small (auto)correlation value for a high or a low probability could also correspond to a more important tradeoff than the same value for a probability of 0.5. This case was represented by our second approach, which accounts for this constraint on binary traits by representing the correlation estimates in relation with the maximum correlation that is possible to simulate for the level of individual heterogeneity and the mean probability recorded for the trait (i.e. $\rho_{OBS}/|\rho_{MAX}|$ or standardized $\beta_T/|\rho_{MAX}|$). Nevertheless, irrespective of which of the two interpretations cause small measures of tradeoffs based on traits with high and low probabilities, such tradeoffs will be more difficult to detect statistically. This agrees with our recent suggestion based on an empirical review in mammals that tradeoffs in terms of costs of reproduction are more difficult to observe in traits that are less variable, as is the case for large and small probabilities because their variance is statistically constrained (Hamel et al. 2010b).

Clearly, each of the two approaches we proposed for estimating the strength of tradeoffs is not perfect, and it will be up to the researcher to evaluate which approach is better suited to a specific study. Nevertheless, the mathematical constraint on binary traits must be accounted for when estimating tradeoffs, particularly for comparing (auto)correlation estimates across binary traits with contrasting probability values, as well as across binary and continuous traits. Therefore, we recommend using the two approaches in concert, as illustrated in Fig. 1B. Moreover, although the use and understanding of generalized linear mixed models has greatly increased in the recent decade, these complex models are in active development and provide estimates that are still considered to be 'approximate' (Gelman and Hill 2007, Bolker et al. 2009). Consequently, we also strongly recommend presenting and interpreting

estimates along with their variance to emphasize the uncertainty on the tradeoffs estimated. In addition, although we have shown here that (auto)correlation estimates can reliably measure the strength of tradeoffs, we recommend using these estimates relative to one another, and not as an absolute value of the strength, because an estimate on its own would not be so meaningful.

The possibility to measure accurately and in a standardized way life history tradeoffs is of paramount importance in ecology and evolution because the availability of such measures will allow contrasting the strength of different tradeoffs detected in populations (Roff 1992, Stearns 1992). In this study, we provide a comparative analysis of the strength of seven potential life history tradeoffs in two contrasted species of large herbivores. In mountain goats, the strongest, and hence the most evolutionary important tradeoff, occurred on the successive probabilities of parturition. Two other tradeoffs were detected: the negative influence of previous parturition on offspring survival and that of early primiparity on adult mass. Compared with the parturition tradeoff, however, these two tradeoffs were weak and the one on adult mass was not statistically supported. The two other traits we analyzed in goats presented a positive correlation, illustrating that these traits did not present a tradeoff (van Noordwijk and de Jong 1986). Although the positive estimate between juvenile growth and age at primiparity showed a large variation, this positive covariance is still likely much stronger than that between early primiparity and longevity, as the latter was overall weak and not statistically significant. Overall, the strongest selection pressures seem to occur on the short-term rather than at the scale of the lifetime in goats. In roe deer, only two potential tradeoffs were analyzed: the occurrence of compensatory growth and that of costs of reproduction on pregnancy. In both cases, however, a positive effect was found, and the positive influence of previous state on growth was stronger than that on pregnancy. When comparing similar traits between the two species, goats presented a strong tradeoff on future reproduction, whereas roe deer did not and rather showed a weak but positive effect. Such between-species differences might result from different tactics of energy allocation to reproduction. In capital breeders (sensu Jönsson 1997) like mountain goats, giving birth and raising an offspring in a given spring/summer prevent mothers from fully replenishing their body reserves and lead them to more likely skip reproduction the year after so that a reproductive tradeoff occurs between two successive years (Hamel et al. 2010a). On the contrary, for income breeders (sensu Jönsson 1997) like roe deer, the ability of giving birth in a given year mostly depends on food resources in summer/fall and is not affected by the reproductive status of the females the previous year. Nevertheless, our analysis is limited to two species, and hence further investigation will be required to determine whether such a functional link between tactics of energy allocation to reproduction and strength of tradeoffs between successive reproductive attempts really corresponds to a life history pattern. Overall, the results based on the seven tradeoffs we selected clearly illustrate that the method we proposed makes it possible to perform fundamental comparative analyses to assess the relative importance of different tradeoffs, e.g. for different traits within a species,

for the same trait but measured for different cohorts or for individuals living in contrasting environments, or for the same trait but across different populations or species.

To sum up, we proposed using a correlation-based approach to measure the strength of life history tradeoffs, and adjusting tradeoffs estimated for binary traits by using the ratio of the observed correlation on the maximum achievable correlation to account for the constraint on binary data. Results clearly showed that using (auto)correlation estimates will generally provide a reliable measure of tradeoffs that can be used for comparative analyses of traits with different distributions. In addition, although our study focused on life history tradeoffs, the approach we proposed can also be used to measure tradeoffs between any kind of traits, such as morphological, physiological, or behavioural traits. Furthermore, we quantified tradeoffs over the entire lifetime of all individuals of a population, but the strength of tradeoffs could vary with age or among cohorts. Our method allows evaluating these possibilities, for example by performing separate analyses for specific group of individuals or for different age stages, or by including an interaction with β_T in the state dependence model. Our work therefore opens the door to a deeper understanding of the evolution of life history traits under constraints, as well as to further developments for expanding the method and overcoming its limitations. For instance, Roff and Fairbairn (2007) suggested that more than two traits could be involved in a tradeoff, leading to the fact that analyses based on only two traits might not allow detecting these tradeoffs. The development of statistical methods like the 'aster' models (Shaw and Geyer 2010) and multivariate cross-classified models (Browne et al. 2007), which allow the joint analysis of multiple life history traits measured over several years, could provide a genuine approach to start answering this question. Furthermore, positive associations between traits resulting from greater variation in resource acquisition than resource allocation are also common in life histories (van Noordwijk and de Jong 1986). The method we proposed could also be extended to quantify positive associations between traits (with caution when strong correlations are expected, i.e. > 0.7, Hamel et al. 2012), thereby providing insights into the strength of the influence of individual variation on associations between traits. Finally, a fundamental aspect that needs to be explored is the quantitative measurement of the relative contribution of different life history tradeoffs on fitness.

References

Andersen, R. et al. 2000. Factors affecting maternal care in an income breeder, the European roe deer. – J. Anim. Ecol. 69: 672–682.

Bates, D. et al. 2012. lme4: Linear mixed-effects models using S4 classes. – R package ver. 0.999999-0.

Berridge, D. M. and Crouchley, R. 2011. Multivariate generalized linear mixed models using R. – CRC Press.

Bolker, B. M. et al. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. – Trends Ecol. Evol. 24: 127–135.

Browne, W. J. et al. 2007. Using cross-classified multivariate mixed response models with application to life history traits in great tits (*Parus major*). – Stat. Modell. 7: 217–238.

Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. – Springer.

Cameron, R. D. 1994. Reproductive pauses by female caribou. – J. Mammal. 75: 10–13.

Chaganty, N. R. and Harry, J. 2006. Range of correlation matrices for dependent Bernoulli random variables. – Biometrika 93:197–206.

Charmantier, A. et al. 2006. Quantitative genetics of age at reproduction in wild swans: support for antagonistic pleiotropy models of senescence. – Proc. Natl Acad. Sci. USA 103: 6587–6592.

Clutton-Brock, T. H. 1991. The evolution of parental care. – Princeton Univ. Press.

Clutton-Brock, T. and Sheldon, B. C. 2010. Individuals and populations: the role of long-term, individual-based studies of animals in ecology and evolutionary biology. – Trends Ecol. Evol. 25: 562–573.

Côté, S. D. and Festa-Bianchet, M. 2001. Reproductive success in female mountain goats: the influence of maternal age and social rank. – Anim. Behav. 62: 173–181.

Descamps, S. et al. 2006. Best squirrels trade a long life for an early reproduction. – Proc. R. Soc. B 273: 2369–2374.

Fisher, R. A. 1956. Statistical methods and scientific inference. – Oliver and Boyd.

Gaillard, J.-M. and Yoccoz, N. G. 2003. Temporal variation in survival of mammals: a case of environmental canalization? – Ecology 84: 3294–3306.

Gaillard, J.-M. et al. 1998. Population dynamics of roe deer. – In: Andersen R. et al. (eds), The European roe deer: the biology of success. Scandinavian Univ. Press, pp. 309–335.

Gaillard, J.-M. et al. 2003a. Effects of hurricane Lothar on the population dynamics of European roe deer. – J. Wildlife Manage. 67: 767–773.

Gaillard, J.-M. et al. 2003b. Cohort effects and deer population dynamics. – Ecoscience 10: 412–420.

Gelman, A. and Hill, J. 2007. Data analysis using regression and multilevel/hierarchical models. – Cambridge Univ. Press.

Glazier, D. S. 2000. Smaller amphipod mothers show stronger tradeoffs between offspring size and number. – Ecol. Lett. 3: 142–149.

Green, W. C. H. and Rothstein, A. 1991. Tradeoffs between growth and reproduction in female bison. – Oecologia 86: 521–527.

Hadfield, J. D. 2010. MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. – J. Stat. Softw. 33: 1–22.

Hamel, S. and Côté, S. D. 2009. Foraging decisions in a capital breeder: tradeoffs between mass gain and lactation. – Oecologia 161: 421–432.

Hamel, S. et al. 2009a. Individual variation in reproductive costs of reproduction: high-quality females always do better. – J. Anim. Ecol. 78: 143–151.

Hamel, S. et al. 2009b. Individual quality, early-life conditions, and reproductive success in contrasted populations of large herbivores. – Ecology 90: 1981–1995.

Hamel, S. et al. 2010a. Maternal characteristics and environment affect the costs of reproduction in female mountain goats. – Ecology 91: 2034–2043.

Hamel, S. et al. 2010b. Fitness costs of reproduction depend on life speed: empirical evidence from mammalian populations. – Ecol. Lett. 13: 915–935.

Hamel, S. et al. 2012. Statistical evaluation of parameters estimating autocorrelation and individual heterogeneity in longitudinal studies. – Meth. Ecol. Evol. 3: 731–742.

Heckman, J. J. 2001. Micro data, heterogeneity, and the evaluation of public policy: Nobel lecture. – J. Polit. Econ. 109: 673–748.

Hodges, J. S. and Reich, B. J. 2010. Adding spatially-correlated errors can mess up the fixed effect you love. – Am. Stat. 64: 325–334.

Jönsson, K. I. 1997. Capital and income breeding as alternative tactics of resource use in reproduction. – Oikos 78: 57–66.

Kirkwood, T. B. L. 1977. Evolution of ageing. – Nature 270: 301–304.

Law, R. 1979. The cost of reproduction in annual meadow grass. – Am. Nat. 113: 3–16.

Link, W. A. and Doherty Jr., P. F. 2002. Scaling in sensitivity analysis. – Ecology 83: 3299–3305.

Lloyd, D. G. 1987. Selection of offspring size at independence and other size-versus-number strategies. – Am. Nat. 129: 800–817.

Martin, J. G. A. et al. 2011. Measuring individual differences in reaction norms in field and experimental studies: a power analysis of random regression coefficient models. – Meth. Ecol. Evol. 2: 362–374.

Molenberghs, G. et al. 2012. A combined beta and normal random-effects model for repeated, overdispersed binary and binomial data. – J. Multivariate Anal. 111: 94–109.

Morris, W. F. and Doak, D. F. 2004. Buffering of life histories against environmental stochasticity: accounting for a spurious correlation between the variabilities of vital rates and their contributions to fitness. – Am. Nat. 163: 579–590.

Nakagawa, S. and Cuthill, I. C. 2007. Effect size, confidence interval and statistical significance: a practical guide for biologists. – Biol. Rev. 82: 591–605.

Nakagawa, S. and Freckleton, R. P. 2008. Missing inaction: the dangers of ignoring missing data. – Trends Ecol. Evol. 23: 592–596.

Nelder, J. A. and Wedderburn, R. W. M. 1972. Generalized linear models. – J. R. Stat. Soc. A 135: 370–384.

Nilsen, E. B. et al. 2009. A slow life in hell or a fast life in heaven: demographic analyses of contrasting roe deer populations. – J. Anim. Ecol. 78: 585–594.

Nussey, D. H. et al. 2008. Measuring senescence in wild animal populations: towards a longitudinal approach. – Funct. Ecol. 22: 393–406.

Plard, F. et al. 2012. Modeling reproductive trajectories of roe deer females: fixed or dynamic heterogeneity? – Theor. Popul. Biol. 82: 317–328.

Pomeroy, P. P. et al. 1999. Consequences of maternal size for reproductive expenditure and pupping success of grey seals at North Rona, Scotland. – J. Anim. Ecol. 68: 235–253.

Prentice, R. L. 1988. Correlated binary regression with covariates specific to each binary observation. – Biometrics 44: 1033–1048.

Reiter, J. and Le Boeuf, B. J. 1991. Life history consequences of variation in age at primiparity in northern elephant seals. – Behav. Ecol. Sociobiol. 28: 153–160.

Roff, D. 1992. The evolution of life histories: theory and analysis. – Chapman and Hall.

Roff, D. A. 2002. Life history evolution. – Sinauer.

Roff, D. A. and Fairbairn, D. J. 2007. The evolution of tradeoffs: where are we? – J. Evol. Biol. 20: 433–447.

Schielzeth, H. 2010. Simple means to improve the interpretability of regression coefficients. – Meth. Ecol. Evol. 1: 103–113.

Shaw, R. G. and Geyer, C. J. 2010. Inferring fitness landscapes. – Evolution 64: 2510–2520.

Sim, I. M. W. et al. 2011. Characterizing demographic variation and contributions to population growth rate in a declining population. – J. Anim. Ecol. 80: 159–170.

Skjæraasen, J. E. et al. 2010. A cost of reproduction in male Atlantic cod (*Gadus morhua*). – Can. J. Zool. 88: 595–600.

Smith, C. C. and Fretwell, S. D. 1974. The optimal balance between size and number of offspring. – Am. Nat. 108: 499–506.

Smith, H. G. et al. 1989. The tradeoff between offspring number and quality in the great tit *Parus major*. – J. Anim. Ecol. 58: 383–401.

Stearns, S. C. 1992. The evolution of life histories. – Oxford Univ. Press.

van de Pol, M. and Verhulst, S. 2006. Age-dependent traits: a new statistical model to separate within- and between-individual effects. – Am. Nat. 167: 766–773.

van de Pol, M. and Wright, J. 2009. A simple method for distinguishing within- versus between-subject effects using mixed models. – Anim. Behav. 77: 753–758.

van Noordwijk, A. J. and de Jong, G. 1986. Acquisition and allocation of resources: their influence on variation in life history tactics. – Am. Nat. 128: 137–142.

Wakefield, J. 2007. Disease mapping and spatial regression with count data. – Biostatistics 8: 158–183.

Weladji, R. B. et al. 2008. Heterogeneity in individual quality overrides costs of reproduction in female reindeer. – Oecologia 156: 237–247.

Williams, G. C. 1966. Natural selection, the costs of reproduction, and a refinement of Lack's principle. – Am. Nat. 100: 687–690.

Wilson, A. J. et al. 2009. Trading offspring size for number in a variable environment: selection on reproductive investment in female Soay sheep. – J. Anim. Ecol. 78: 354–364.

Wilson, A. J. et al. 2010. An ecologist's guide to the animal model. – J. Anim. Ecol. 79: 13–26.

Wood, S. N. 2006. Generalized additive models: an introduction with R. – Chapman and Hall.

Wooldridge, J. M. 2005. Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. – J. Appl. Econ. 20: 39–54.

Zhang, H. et al. 2011. On fitting generalized linear mixed-effects models for binary responses using different statistical packages. – Statist. Med. 36: 2562–2572.

Zhao, X. et al. 2012. Semiparametric model for recurrent event data with excess zeros and informative censoring. – J. Stat. Plan. Infer. 142: 289–300.