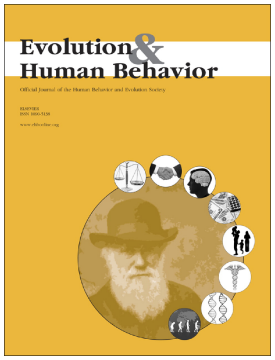


On the use of "life history theory" in evolutionary psychology

Stephen C. Stearns, António M.M. Rodrigues

PII: S1090-5138(20)30026-X

DOI: https://doi.org/10.1016/j.evolhumbehav.2020.02.001

Reference: ENS 6308

To appear in: *Evolution and Human Behavior*

Received date: 7 July 2019

Revised date: 3 February 2020

Accepted date: 12 February 2020

Please cite this article as: S.C. Stearns and A.M.M. Rodrigues, On the use of "life history theory" in evolutionary psychology, *Evolution and Human Behavior*(2020), https://doi.org/10.1016/j.evolhumbehav.2020.02.001

On the use of "life history theory" in evolutionary psychology

Stephen C. Stearns[1] and António M.M. Rodrigues[2]

[1]Department of Ecology and Evolutionary Biology

Yale University

Box 208106

New Haven, CT 06520-8106 USA

[2]Department of Zoology

University of Cambridge

Downing Street

Cambridge CB2 3EJ UK

Email of corresponding author: stephen.stearns@yale.edu

Declaration of interest: We have no financial or personal relationships with people or organizations that could bias this work.

Submission declaration: This work has not been published previously and is not under consideration elsewhere.

On the use of "life history theory" in evolutionary psychology

Declaration of interest: We have no financial or personal relationships with people or organizations that could bias this work.

Submission declaration: This work has not been published previously and is not under consideration elsewhere.

Abstract

We critically review the use of the term "life history theory" in recent publications on evolutionary psychology, focusing on how the idea of a fast-slow continuum is deployed in that literature. We raise four issues:

First, concerning plasticity, should we expect the effects of plasticity on the developmental response of a trait to mirror the effects of selection on the mean of that trait? We conclude that we should not. Do only plastic responses to harsh or unpredictable environments accelerate maturation, or are there plausible alternatives, such as nutrition? In many situations better nutrition is a plausible alternative.

Second, how should we conceive of the harshness of an environment? It has several important dimensions. It could mean an increase in the mean mortality rate, a decrease in the mean growth rate or fertility rate, or increases in the variances of any of those rates. Our judgement of harshness will also be affected by the distribution of such effects across patches in space and through generations in time. The combination and distribution of effects make important differences to predictions.

Third, where did the fast-slow idea come from, and how much does it explain? It was initially detected in comparisons across higher taxonomic levels, whose relevance to variation among individuals is unclear and where it fails to explain much of the variation.

Fourth, what sorts of processes could generate the fast-slow pattern? Here we expand on insights mentioned earlier in passing to make clear how spatial population structure and class effects generate alternative predictions.

We conclude with some thoughts on the nature of theories and research strategies and on how one might respond to empirical puzzles.

Keywords: life history theory; fast-slow; patterns; causes; research strategies

1. Introduction

Because the term "life history theory" is increasingly used by some in evolutionary psychology to support explanations of developmental responses of human behavior, cognition, and physiology along a fast-slow continuum associated with psychosocial stress (Nettle and Frankenhuis 2019), this is a good time to take a step back to ask, what exactly does life history theory predict, and is the way in which it is deployed in evolutionary psychology appropriate? Specifically, we address these questions:

1. Should one expect the effects of plasticity to mirror the effects of selection on the mean of a trait? What effects can accelerate maturation?
2. How should we represent the harshness of an environment? Several combinations of the means and variances of age-specific vital rates are candidates.
3. Where did the idea of a fast-slow pattern come from, how much variation does it explain, and should we expect it to explain variation among individuals within populations?
4. What sorts of processes could each generate the fast-slow pattern? Spatial population structure and class effects can both generate alternative predictions.

These considerations reveal a general issue that arises when one borrows ideas from other fields and has surfaced in many, including our own. While borrowing can be powerful, it has pitfalls. Its power comes from a new point of view that suggests new questions – a precious commodity in science. The usual reaction is to analogize: this pattern in my field (e.g. inter-individual correlations among

traits related to behaviour, cognition, physiology, and psychosocial stress) is like that pattern in the other field (e.g. among-species correlations among life history traits), therefore what they found there should apply here.

But analogies mislead if the underlying causes differ. Insights from analogies must be validated by showing that the predicted effects also result from tracing causes in the new context. That requires building and testing formal models that demonstrate that the borrowed predictions also arise naturally from the causal relationships of the new field. In formal models, careful attention to assumptions is important, for predictions are only meaningful if the assumptions hold. For example, in the population under study, can one demonstrate that individuals in which age at first birth is earlier are also those in which both psychosocial stress and risk of adult mortality are higher, or are they simply the ones who are better nourished?

Why do people doing research on life history evolution find formal models useful? First, by setting down explicit assumptions and analyzing their consequences, it was, very early on, possible to distinguish between two major causal claims: r&K selection on the one hand, age-specific selection on the other. The clarifying effect of that experience shaped the field for decades to follow. Second, the interactions with various intrinsic tradeoffs of age-specific changes in fertility and mortality produced a complex conceptual structure that could only be handled with confidence by building mathematical models capable of keeping track of all those interactions. Third, many of the predictions depend on quantitative effects, for the costs and benefits of tradeoffs change both with trait values and with environmental gradients. Qualitative, verbal models cannot handle such effects, which can lead to major changes in predictions.

The quantitative nature of relationships is thus central. Drawing qualitative conclusions from a body of theory in which quantitative effects are important can be dangerous when predictions change qualitatively for quantitative reasons – as when the balance of benefits and costs in a trade-off shifts from positive to negative.

While many of the tests of quantitative life history theory in biology have been done with qualitative comparisons, as have applications of life history theory in evolutionary psychology, there are some in which it has been possible to predict evolutionary trajectories quantitatively. For example, when Ackermann and colleagues (Ackermann et al. 2007) made quantitative predictions of changes in age at first birth and interbirth interval in bacteria undergoing experimental evolution, they found changes in age at first birth close to those predicted, but changes in interbirth interval much smaller than those predicted. This result suggested that interbirth interval is constrained in ways not previously expected and productively focused attention on why the mechanisms that determine reproductive rate are not easily changed by strong selection. Here the useful feature of the quantitative prediction was its capacity to fail in an interesting way.

Those using life history theory to support claims in evolutionary psychology therefore need answers to these questions: What is assumed? How were the models built? When assumptions change among models, do predictions change? Is what is thought to be a general prediction in fact context-dependent?

To be clear, we define the senses in which we use terms in a Glossary.

[Glossary near here]

We next unpack the fast-slow continuum as it is understood in life history theory through answers to four expanded sets of questions:

- How do the reaction norms for age and size at maturity as functions of environmental quality evolve? Is the evolved mean population response to a deterioration in environmental quality expected to be the same as the evolved individual plastic response? The answer is often no, but not always. Note that the predictions of reaction norms for age and size at maturity are the only part of life history theory in which explicit predictions have been made for plastic responses of individuals to environmental variation encountered during development, a type of response that is particularly relevant to evolutionary psychology.
- What are the impacts of the means and variances of mortality rates on the evolution of repeated reproduction (iteroparity) and therefore lifespan? One could define a high-mean mortality rate as reflecting a "harsh" environment, or a high-variance mortality rate as reflecting an "unpredictable" environment. All depends on whether such variation affects juveniles or adults more strongly – the issue is at its heart quantitative, not qualitative (Bulmer 1985). Increased temporal variation in juvenile mortality selects for longer, not shorter, reproductive lifespan. Increased temporal variation in adult mortality does select for a shorter reproductive lifespan. We are not aware of any formal models that predict plastic responses of reproductive lifespan to changes in the means or variances of mortality rates.
- Where did the idea of a fast-slow continuum come from, and how much of the overall variation in life histories among species is in fact explained by it? It is an important but not the only axis of variation (a second axis ranks mammals from precocial to altricial). The effects that it represents originate mostly at higher

taxonomic levels and are correlated with body size; they are not strong among or within species nor among individuals.

- How do spatial structure and class effects impact predictions? With spatial structure there can be local mate competition and local kin competition; both can produce surprising effects. With class effects the environment in which development occurs differs among individuals, who may use quite different strategies to produce the same phenotype. In such cases, inferring evolutionary causes from phenotypic patterns is not reliable, for the same patterns can be produced for quite different reasons.

We begin with a model for the evolution of maturation reaction norms, the phenotypically plastic reactions that occur as individuals respond developmentally to environmental variation.

2. Reaction norms for age and size at maturity

In this section we state the assumptions and predictions of one life-history model to make these points: (a) one can only predict how selection will change traits when one can write down equations that link changes in those traits to changes in fitness; (b) changing the assumptions about the trade-offs between traits changes the predictions; (c) the model presented has been criticized for several reasons, and the resulting modifications to it have revealed an entire range of possible predictions (something that could happen when formal models are produced in evolutionary psychology); (d) one frequent prediction remains the initial one presented here, which also appears to be the one consistent with plasticity in humans: organisms should mature young and large when growing rapidly and old and small when growing slowly. If growth is correlated with the harshness of the

environment, then this is not consistent with first-order expectations based on fast-slow continuum as it is used in some of the evolutionary psychology literature. Although some evolutionary psychologists – to their credit – have suggested that environmental gradients can affect life history traits in ways that are inconsistent with the proposed fast-slow continuum (e.g. Ellis 2004, Coall and Chisholm 2003), the idea of the continuum remains widespread.

2.1 The classical model of an age-structured population

Please bear with us while we step through the setup for the model. Such equations are essential for understanding how biologists model life histories. The devil is truly in the details.

Define terms as follows: *t* = time, *x* = age, *B(t)* = the sum of all births at time *t,* *n(x,t)* = the number of individuals of age *x* at time *t, b(x)* = the expected number of births for a female of age *x, l(x)* = survival from birth to age *x, r* = instantaneous rate of growth of a clone with the given life history traits, i.e. fitness in this context (see Glossary), α = age at first birth, and ω = age at last birth.

B(t) is the sum of births in all age classes *x* (i.e. between the age at first birth, α, and the age at last birth, ω) where each age class has *n* females with average fertility *b*:

$$B(t) = \int_{\alpha}^{\omega} b(x)n(x,t)dx$$

So, the sum of all births at time *t* equals the number of individuals of age *x* at time *t* times the expected number of births per female of age *x* summed over all ages x. We also know that the number of females at time *t* is the number born at time *t* – *x* times the probability of surviving to age *x*, *l(x):*

$$n(x,t) = B(t-x)l(x)$$

Substituting in the equation above yields the renewal equation:

$$B(t) = \int_{\alpha}^{\omega} B(t-x)l(x)b(x)dx$$

While it can be demonstrated that a population with constant birth and death schedules will grow exponentially, that demonstration (Lotka 1907, Euler 1970) takes many steps, so here we simply assume that it has an exponential solution $B(t) = e^{rt}$

$$e^{rt} = \int_{\alpha}^{\omega} e^{r(t-x)} l(x)b(x)dx$$

Dividing both sides by e^{rt} yields the classic Euler-Lotka equation:

$$1 = \int_{\alpha}^{\omega} e^{-rx}\, l(x)b(x)dx$$

This equation holds whenever the schedule of birth and death rates is constant in a population growing independent of density. Such a population will converge to a stable age distribution – one in which the proportion of individuals in each age class remains constant and in which the absolute number of individuals in each age class changes at the same rate that the population grows. That growth will be exponential, with the rate of multiplication per time unit $\lambda = e^{r}$, the measure of fitness in this model. It can be positive, zero, or negative depending on the vital rates.

This equation does something both simple and powerful: it explicitly relates key life history traits – age at first birth α, survival to age *x*, births at age *x,* and age at last birth ω – to fitness, allowing us to ask, how does a change in any of those traits change fitness? However, that power and simplicity are bought at the cost of important assumptions: constant birth and death schedules in an environment with no density dependence. Building an explicit model made these assumptions clear.

2.2 Relating vital rates to age at first birth

To understand how changes in vital rates will affect fitness, we need to represent the relationships between age at first birth and survival to age *x* and births expected at age *x*. Because survival and fecundity often depend on size, we allow them to depend on both age and individual growth rate during development *k*. With those relationships defined, we can ask how a change in age at maturity will affect fitness by examining the sensitivity of fitness to changes in birth, death, and growth rates. To understand the impact of tradeoffs, we choose the functions relating those three traits to represent the following ideas (Stearns and Koella 1986):

- Survival to maturation improves as age at first birth is delayed because mothers who are older and larger when they first give birth can produce offspring whose instantaneous survival rates are higher, more than compensating – up to a point – for the longer period in which they are exposed to mortality before they mature. So let instantaneous survivorship = some such function of age at first birth α, growth rate *k,* and age *x.*
- Births at age *x* are zero until maturation and depend on size after maturation, where size is a function of growth rate and age. So let instantaneous fertility = some such function of age at first birth α, growth rate *k,* and age *x.*

Note that earlier maturation will result in higher juvenile mortality rates, and later maturation will result in higher birth rates because females who are older when they first give birth are larger. Thus, the functions chosen for survival and fertility represent the costs and benefits of earlier versus later maturation.

We insert those functions for survival and fertility into the Euler-Lotka equation and ask, what is the age at first birth α that maximizes fitness *r*? Answering that question involves calculations that can be found in (Stearns and Koella 1986).

With one further, critical assumption – that all growth environments are encountered equally frequently – we can predict the age at first birth that maximizes fitness for each growth rate given the functions specified for survivorship and fertility. The result is a graph of the maturation reaction norm in a space defined by age and size (Figure 1).

[place Figure 1 here]

The message of Fig. 1 is clear: when growing rapidly (k_1) mature young and large; when growing slowly (k_4) mature old and small. That is an evolved rule for a conditional response to environmentally induced variation in growth rates. If organisms grow more rapidly in a more secure and nurturing environment, then this prediction is the opposite of that of the fast-slow continuum as it has sometimes been used in the evolutionary psychology literature. If slower growth caused by nutritional stress represents "harshness", then this prediction is also opposite to that of a fast-slow continuum. While we note that evolutionary psychologists have sometimes distinguished nutritional from psychosocial stress and have suggested that the two types of stress would have opposite effects on age at maturation (Ellis et al. 2009), there is to our knowledge no formal life history theory concerning the effects of psychosocial stress on age at maturation, and there is a reason for that: we have information on the relationships between growth and mortality and growth and fertility, but we do not have information on those relationships for psychosocial stress.

2.3 Separating "nature" from "nurture"

Fig. 1 describes what one would expect to see in a population in which the relationships among age and growth rates, births and deaths remained constant. But what happens when extrinsic mortality changes? The age-specificity of that change is important: whether the change in extrinsic mortality affects primarily juveniles or adults is critical. Consider a decrease in juvenile mortality, which decreases the cost of early reproduction. That change has been caused in humans over the last two centuries in developed countries by better water supplies, better nutrition, antibiotics, vaccines, and medical care. It is predicted to cause a genetically based response to selection in which the entire reaction norm moves down and to the left (Fig. 2).

[place Figure 2 here]

Figure 2 contains several important messages.

- First, the empirical response is consistent with the major prediction of Figure 1: when the nutritional environment improved (large arrow pointing up and to the left under the upper reaction norm), women did in fact mature earlier and at a larger size (Frisch and Revelle 1970).
- Second, when such environmental changes endure long enough and consistently enough to produce an evolved response to lower juvenile mortality, then the model predicts that the entire reaction norm will move down and to the left (small arrow pointing down) – under all growth conditions, women are then predicted to mature younger and smaller. This is precisely the direction of selection on age at maturity (but not on size at maturity) measured for women in developed countries like the United States, Finland, and Australia (Stearns et al.

2010). Note that the prediction of earlier maturation when juvenile mortality decreases is the opposite of that predicted by the fast-slow continuum as used in some of evolutionary psychology if juvenile mortality reflects psychosocial stress.

- Third, the model also predicts what will happen when the change in the environment affects adults more strongly than juveniles. We focus here on changes in the average extrinsic adult mortality rate, where extrinsic mortality is the component of mortality due to the environment (as opposed to intrinsic mortality, which can be ascribed to tradeoffs and aging – see Glossary). When the extrinsic mortality rate of adults increases with little change in extrinsic juvenile mortality rates, then the same shift in the reaction norm seen in Fig. 2 is predicted – down and to the left (Stearns and Koella 1986). *That* response *is* consistent with the claims about the fast-slow continuum as used in some evolutionary psychology literature, and it occurs as an evolved, not a plastic, response *if* the deterioration of the environment affects adults more strongly than juveniles.

  Thus, this (relatively) simple life history model yields predictions some of which are not consistent with a fast-slow continuum as it is understood in some of the evolutionary psychology literature. Some them are, but we do not yet know whether the assumptions that yield such predictions about plastic responses hold for the cases claimed by evolutionary psychology.

  2.4 How well has this model stood up?

  Criticisms of and modifications to the model have come primarily from three directions.

- First, the assumption that all growth environments are encountered equally frequently was pointed out as being particularly onerous. When the environment

is spatially heterogeneous and organisms flow from sources to sinks (see Glossary), populations will be better adapted to the source than to the sink, and this will change the shape and position of a reaction norm that spans source-sink conditions (Kawecki and Stearns 1993). For human populations, the source-sink structure informs analysis of large-scale movements from rural to urban habitats and of migration from developing to developed countries.

- Second, in recent years much work has been done on eco-evolutionary feedbacks and the resulting adaptive dynamics. When the environment causes plastic and demographic changes in life history traits that change population dynamics, that change in population dynamics changes the selection pressures operating on life history traits. Those traits then respond with genetically based, evolved changes that are expressed in new generations of organisms that now have different evolved responses to environmental variation. The feedback cycle continues, with ecological conditions driving evolutionary changes that result in different ecological conditions. Such an eco-evolutionary dynamic produces continued change in both evolved and plastic responses. One important change in the environment is the density of the evolving population. When eco-evolutionary dynamics are modeled to include density-dependent variation in growth and its relationship with mortality rates (Marty et al. 2011), a whole range of predictions are made that depend on the nature of the growth-mortality tradeoff (their Fig. 4). Some of the predicted plastic responses are consistent with Fig. 1, whereas others make a contrasting prediction that could be interpreted as being more consistent with the picture of life history theory presented in the evolutionary psychology literature, i.e. when growing rapidly mature young and small, when growing slowly mature old and large. Such an interpretation depends on

assuming that faster growth reflects greater psychosocial stress. The question then becomes, do the assumptions made in the model when it produced that prediction hold in the human populations for which those effects are claimed? That open research question needs an answer.

- Third, at the same time that work on eco-evolutionary dynamics advanced, work on evolutionary medicine developed in parallel. In evolutionary medicine mismatch is a core concept: modern humans can be mismatched to their current environments because they spent most of their evolutionary past under different conditions. We cannot expect traits to be well adapted to circumstances that are being encountered at quite different frequencies than they were in our evolutionary past – in some cases, for the first time. When that happens, a reasonable expectation is that variation in responses will increase, some of which may be maladaptive or pathological (Stearns and Medzhitov 2015): one alternative explanation for responses to psychosocial stress is that they are not adaptive but mismatched.

  Thus, the message from the part of life history theory that tries to predict optimal plastic responses to environmental variation is complex. Some of its predictions are opposite to the claims made in evolutionary psychology; some are not; but we do not yet know whether the assumptions that one needs to make to get predictions that match those claims hold in the relevant cases. We would not be aware of any of those nuances if formal models had not been built for which assumptions were stated explicitly.

  Now we turn to the second issue, what is the predicted impact of changes in means and variances of adult and juvenile mortality rates on lifespan?

3. Means and variances of mortality rates: the evolution of reproductive lifespans.

Here we rehearse some history to make clear that important qualifications to the fast-slow paradigm were known practically from the start.

3.1 Semelparity, iteroparity, and mean mortality rates

Formal modeling of life history evolution began with Cole's 1954 paper (Cole 1954), which posed a paradox that was not resolved until 1973 by Charnov and Schaffer (Charnov and Schaffer 1973). Cole found the existence of any organisms that reproduced more than once puzzling, for under his assumptions, any organism that reproduced just once and then died (semelparity) could outcompete an organism that reproduced annually forever (extreme iteroparity) by just increasing its fertility by one offspring. In other words, "a clutch size of 101 at age 1 would serve the same purpose as having a clutch of size 100 every year forever" (Charnov and Schaffer 1973, p. 791).

Charnov and Schaffer resolved the paradox by noting that Cole had assumed that juvenile and adult annual survival were equal. Because annual juvenile survival (p_j) is usually much lower than annual adult survival (p_a), they suggested that for a semelparous species, the absolute gain in fitness that could be achieved by changing to iteroparity would be equivalent to adding p_a / p_j to the average annual fertility. For example, if adult annual survival were 0.90 and juvenile annual survival were 0.30, a semelparous mutant would have to produce at least 3 more offspring than an iteroparous organism to invade a resident iteroparous population. That implies that higher annual adult survival and lower annual juvenile survival select for a longer reproductive lifespan. This confirmed one prediction made by Gadgil and Bossert in 1970, who stated it the other way

round, in terms of mortality rather than survival (Gadgil and Bossert 1970): if mortality increases in all age classes after a certain age *x*, age at first birth will decrease, reproductive effort in all age classes younger than *x* will increase, and, because of the tradeoff between reproduction and survival, lifespan will decrease. In other words, if life becomes harsher on average for older adults, faster life histories will evolve. This is consistent with some versions of the fast-slow interpretation of evolutionary psychology.

However, there was, even at that time, an exception relevant here. Schaffer found that either semelparity or iteroparity could evolve and remain stable under realistic assumptions about fertility and growth, simply depending on where the species started in trait space (Schaffer 1974). Life could end up being fast or slow depending only on starting conditions, not on differences in the harshness or unpredictability of the environment.

These models all considered the impact of changes in the means, not the variances, of age-specific mortality rates. So what about variances?

3.2 Bet-hedging, variance in mortality rates, and lifespan

The idea that it pays to hedge bets by spreading risk originated with Bernoulli in 1738 (translated in Bernoulli 1954) in a paper in which he also first noted that the appropriate way to measure the outcome of a multiplicative process, such as fitness measured across generations, is the geometric mean. The geometric mean weights small values more strongly than large ones; it emphasizes the avoidance of failure more strongly than the benefits of success. Bet-hedging was first applied to life history evolution by Murphy, who noted a strong positive

relationship between variation in spawning success in fish and their reproductive lifespan (Murphy 1968; Murphy was not aware of Bernoulli's results).

Bulmer formalized and generalized Murphy's insight by analyzing the ability of annuals (semelparous) to invade perennials (iteroparous) and vice versa under two forms of density dependence, depending on the relative means and variances of juvenile and adult mortality rates (Bulmer 1985). He found that, other things being equal, variation across generations in the juvenile mortality rate favors iteroparity and variation across generations in the adult mortality rate favors semelparity. Many other studies, summarized by Wilbur and Rudolf (Wilbur and Rudolf 2006), confirm this prediction. They add to it the prediction that certain types of environmental stochasticity can lead to the evolution of both longer reproductive lifespans and delayed age at maturation (Wilbur and Rudolf 2006).

Now translate those results into the language used to describe life history theory by some evolutionary psychologists. One might describe an environment in which *adult* mortality varies stochastically from generation to generation as *unpredictable*. Because of the way that the geometric mean measures fitness, such an environment reduces the contribution of adult survival to fitness and favors the evolution of faster life histories. In contrast, one might describe an environment in which *juvenile* mortality varies from year to year and from generation to generation as *unpredictable*. Such an environment reduces the contribution of juvenile survival to fitness, increases the relative contribution of adult survival to fitness, and favors the evolution of slower life histories. Thus, whether one should expect a faster or slower life history when the unpredictability of the environment increases depends very much on whether it is primarily juveniles or primarily adults who experience that increase in

unpredictability. The actual outcome will depend on the quantitative relationship of both the means and the variances of adult and juvenile mortality. While this nuance has been recognized by some evolutionary psychologists (e.g. Ellis et al. 2009, Del Giudice et al. 2015), it often gets buried when one categorizes environments simply as unpredictable without specifying precise impacts on age classes.

3.3 Conclusions on reproductive lifespan

The impact of environmental changes on lifespan can be compactly summarized by recognizing that an increase in variation in mortality rates in a given age class has an impact on fitness similar to an increase in the mean mortality rate of that age class. Either effect reduces the amount that age class contributes to fitness. We therefore arrive at the following prediction: from a given starting point, things that reduce the relative contribution of *juveniles* to fitness will increase lifespan; things that reduce the relative contribution of *adults* to fitness will decrease lifespan (Stearns 1992, Chapter 8). Note that this is a prediction about the mean value of the evolved lifespan in a population experiencing such conditions; it is *not* a prediction of the plastic response of a developing individual to the perception of such conditions. As we saw above (Section 2.2), plastic responses can be in the opposite direction to population mean responses. Not only does formal theory for plastic responses of lifespan to environmental variation not yet exist – there is not yet any formal theory for plastic responses of the suites of correlated life history traits to which evolutionary psychologists often refer. The reason is simple: in such cases the interactions are so complex that analytic theory has proven intractable. Although computer simulations can be used, these often come at a cost, as it can be difficult to gauge with precision the underlying

causes of the results. A compromise solution is to develop the model analytically and solve the equations numerically. This is certainly an area where further developments in analytical methods could be useful.

4. Where did the fast-slow idea come from, and how much does it explain?

4.1 Where it came from.

An early and influential prediction about the evolution of life histories was made by MacArthur and Wilson in their book on island biogeography (MacArthur and Wilson 1967). They called it *r&K selection,* a pattern that arranged species along a spectrum from those that evolved under density-independent regulation where rapid population growth was advantageous to those that evolved under density-dependent regulation where ability to reproduce in spite of competition was advantageous. Species that were *r-selected* were supposed to be smaller, reproduce earlier, have more smaller offspring, and live shorter lives than those that were *K-selected* (Pianka 1970). That is the origin of the idea that life histories can be arranged along a single dimension from fast (*r-selected*) to slow (*K-selected*).

Much of the subsequent work in life history theory showed that concentrating on the age-specific impacts of the environment on vital rates, rather than on mode of population regulation, was a more powerful and precise way to make predictions that could survive both comparative and experimental tests (Stearns 1992). However, much of that subsequent work did not include the important effects of density that had originally been represented by *K-selection*. This deficiency has since been redressed with ever-increasing success by more recent work in

adaptive dynamics and eco-evolutionary feedbacks (e.g. Marty et al. 2011, one of many examples).

4.2 How much it explains.

Although the initial descriptions of *r&K* selection were based on an analysis of causes that was at best a partial truth, they did point to an important pattern that is most easily perceived in large-scale comparisons of the mean values of life history traits in species, genera, families, and orders. That pattern depends strongly on the correlations of life history traits with body size.

An early comparative study of patterns of variation of life history traits in the mammals using principal component analysis found that body weight was associated with much of the tendency for species to fall on a "single axis ranging from early maturing, highly fecund, and short-lived small animals to the opposite" (Stearns 1983, p. 173). That first principal component accounted for either 68% or 75% of the variation among mammal species, depending on which of two data sets was used. When order and family effects were removed by subtracting order and family means and repeating the analysis on the residuals, the variation accounted for by the first principle component dropped to either 29% or 36%. It still described the fast-slow continuum, but that continuum accounted for much less of the variation. Analyses done within families showed that the loadings of the life history traits on the first two principle components varied among families. The point here is that much of the variation captured by the fast-slow idea resides at the level of differences among higher taxonomic units, not at the level of individual responses. The closer one approaches the individual level, the weaker the pattern becomes, and it weakens in a heterogeneous fashion, differently in

each taxonomic family. What causes the pattern to shift as the focus shifts down the taxonomic hierarchy is not currently known and relates to the difficulty of understanding the meaning of terms such as "phylogenetic constraint." The patterns are real; the causes remain obscure.

Then, starting in the mid-1980's, methods of comparative analysis developed rapidly, driven by improvements in the logic of phylogenetic analyses and remarkable increases in the amount of reliable DNA sequence data. These methods allow more reliable estimates of the amount of variation in life history traits that should be attributed to phylogenetic effects inherited from ancestors and to body size. It is the residual variation, the variation that remains after accounting for such effects, that might "in part represent species-specific responses to unique ecological selective regimes" (Miles and Dunham 1992, p. 848). For example, in the lizard family Iguanidae, only 4-22% of variation in age at first reproduction, clutch size, and neonate size was explained by species-level effects after body size was controlled in nested analyses of covariance. The proportion of variation in body size, age at first reproduction, and neonate size that was attributable to shared common ancestry in a phylogenetic autocorrelation model was 51%, 36%, and 56%, respectively (Miles and Dunham 1992). This study confirmed that most of the variation among species in life history traits, and thus most of the fast-slow pattern, should be attributed to the effects of body size and shared ancestry rather than to local adaptation to specific environments.

In a comprehensive study of fish, mammals, and birds based on data from nearly 2,300 species, Jeschke and Kokko (2008) reinforced the earlier conclusions: the traits associated with a fast-slow continuum change if one removes body size; the

associations change differently in different clades; and the amount of variation explained by a single fast-slow axis drops significantly when size and phylogeny are taken into account.

Thus, multiple studies report that the comparative fast-slow pattern is neither as strong nor as consistent as is claimed in many recent papers in evolutionary psychology, and where it is found, it is much stronger in comparisons of higher levels of the taxonomic hierarchy than it is in comparisons among species, much less among individuals within a population. Thus, the effects of body size and phylogeny do much to explain the perception of a fast-slow pattern at higher taxonomic levels, where the mechanisms that might have produced the patterns remain obscure, as does their relationship to patterns that arise when comparing individuals within populations, an entirely different level of analysis. Now we turn to the effects of spatial structure and classes within and among populations.

5. Similar patterns for different reasons: spatial structure and class effects

Here we expand on some of our comments above to further clarify the challenges posed by alternative explanations when analyzing life history traits within populations. In contrast to Section 2, which deals with maturation, here we deal with fertility.

5.1 Models and mechanisms

Because life history patterns can be generated by multiple mechanisms, the assumptions of models must be tested, and the structural assumptions of models matter. Whereas most life history models assume spatially-unstructured populations, human populations are highly structured.

We illustrate these points with a specific problem. Several populations show fertility and life expectancy gradients that are inversely correlated. That is, while some individuals have higher fertility rates and die younger, other individuals have lower fertility rates and live longer. A common explanation for the negative correlation between the fertility rate and life expectancy gradients is a corresponding gradient in maternal extrinsic mortality (Fig. 3B; see Appendix for details). However, this negative correlation among gradients can also emerge for nutritional reasons (Fig. 3A). Both better nutritional state and higher maternal extrinsic mortality predict higher birth rates and shorter lifespans (Fig. 3A and 3B). Conversely, both poorer nutritional state and lower adult extrinsic mortality predict lower birth rates and longer lifespans (Fig. 3A and 3B).

At first sight, the impact of nutrition on lifespan may seem counter-intuitive. After all, we often observe a positive correlation between nutritional state and longevity. However, a closer inspection of the model explains this apparent contradiction and yields some insights. Nutrition affects organisms in many ways, and while some of these effects influence fertility and growth patterns, others influence survival. Poor nutrition, for instance, can lead to amenorrhea, which reduces fertility, but it can also depress the immune system, which impairs survival rates. The model assumes that nutrition has an impact on fertility-related traits only (i.e. we hold the impact of nutrition on survival constant). Under such conditions, natural selection favours mothers who increase their reproductive effort in proportion to their nutritional state, a pattern that is often seen across natural populations.

But why does the model predict a negative association between nutritional state and lifespan? This happens due to the costs of reproduction. Because mothers in

better nutritional state invest more in reproduction, their intrinsic mortality rate (see Glossary) increases, which, all else being equal, leads to shorter lifespans. In natural populations, it is often the case that nutritional state is correlated with phenotypes and social conditions that contribute to both fertility (or growth) and survival. For instance, nutritional state may be correlated with food availability but also with an extended and supportive network of kin members. These factors improve fertility and also allow individuals to live longer, with the result that at the population level nutritional state, fertility, and lifespan are positively correlated. Here the direct positive effect of nutrition on longevity buffers the indirect negative effects of nutrition (via increased reproductive effort) on longevity, which may be difficult to detect in natural populations (e.g. Hurt et al. 2006).

Our relatively simple model illustrates an important point about the use of mathematical models: a match between the predictions of a model and observed patterns does not mean that the model explains the observations. Unless the assumptions and mechanisms of the model are themselves tested, one risks mistaking correlation for causation. To avoid such errors, it is useful to ask: do the parameter values required to generate the observed patterns correspond to those measured in the study? For studies involving cognition, is there a conceivable mechanism that allows individuals to infer these parameter values?

Often mathematical predictions are also sensitive to the underlying assumptions of the mathematical framework, a problem that is frequently neglected (Rodrigues and Kokko 2016). For instance, the predictions outlined above assume a population without group structure. If populations are structured into groups, then the predictions may change. In group-structured populations, we still find

that higher maternal extrinsic mortality is associated with higher reproductive rates and shorter lifespans (Fig. 3D). However, nutritional state no longer has an effect on optimal reproductive effort strategies, and therefore lifespan remains invariant in relation to nutritional state. This is because in a group-structured setting, better nutritional state increases the number of offspring born in the local habitat, which, in turn, increases local competition for resources. These two quantities affect fitness in opposite directions, with the result that nutritional state has little or no effect on reproductive effort and lifespan (Fig. 3C).

[place Figure 3 here]

5.2 Class-effects

Such results illustrate how difficult it can be to infer behaviour from life history data in complex environments. Natural selection operates through the differential reproductive success of genes. All else being equal, genetic variants that have higher fitness will spread throughout the population. Often, however, all else is not equal. Within a population, some individuals will live in supportive areas, others in deprived areas; some will be strong and vigorous, others will be weak and feeble; some will be females, others will be males; and some will be in a dominant position, others in a subordinate position.

Such categories are what evolutionary geneticists call classes. They are the focus of much interest in part because they impact the way natural selection operates. Individuals in privileged classes (e.g. areas with abundant resources) may have higher reproductive success, not (necessarily) because they have superior genes but because they live in an affluent area. This will lead to evolutionary change (i.e. change in gene frequency; see Glossary) that is due to what we can call "class-effects." "Classes" can change the costs and benefits of behaviours and

strategies. For instance, in a harsh environment natural selection may favour riskier behaviours depending on whether they increase or decrease fitness. The details of the cost and benefit functions will matter, and these details will be class-dependent.

Consider Fig. 3 again. In Fig. 3C, variation in fertility rates emerges solely because of class-effects. Empirically, this can be quite difficult to detect, especially when we only have a limited amount of life history data, as is often the case. Without complete information, one may mistakenly infer that behaviour differs *because* life history traits vary, when in fact it differs for other reasons.

In this scenario, natural selection favours a simple optimal resource-allocation rule: "invest a proportion *x* of the resources in reproduction, irrespective of the environment". Here, individuals that have access to more resources will also produce more offspring at no additional cost to their survival. To infer the rule that governs the expression of these life history traits, one would need to measure at least the correlations between the nutritional gradient and the fertility and survival rates.

Cases in which different behaviours or strategies cause the expression of similar phenotypes are also common in life history evolution. Under such circumstances, inferring the underlying strategy directly from the observed phenotype can lead to error. Organisms often evolve plastic responses that can be represented by reaction norms, which give the conditional rules for producing phenotypes in response to environments that vary. Such phenotypes depend both on the condition of organisms and on their evolved strategies. For instance, an organism born in privilege may have a natural tendency for higher birth rates that are fully explained by its privileged condition (i.e. its class), rather than some genetic

effect. A female born in deprived conditions may be able to equal the birth rates of her privileged counterparts, but only if she works harder (i.e. has a different strategy).

Sometimes, natural selection favours adjustments in behaviour by mothers in different condition that can lead to interesting predictions of phenotypic patterns. For instance, under local mate competition, when mothers adjust their sex-ratio according to their condition, they are predicted to produce exactly the same number of sons whatever their condition (Yamaguchi 1985). Under kin competition, when mothers adjust the dispersal behaviour of their offspring according to their condition, they are predicted to produce exactly the same number of philopatric offspring (Rodrigues and Gardner 2016). Thus, one has to take class-effects into account when inferring behaviour from life history data. For instance, simply observing that the number of sons in a patch remains constant tells us little about the strategic and behavioral "decisions" that have shaped such patterns.

In sum, inferring behaviour or strategic allocation of resources requires a careful analysis of the problem at hand to identify the portion of the phenotype that is attributable to "class-effects" and the portion of the phenotype that is attributable to inherited behaviour or strategy. Observations of human responses to psychosocial stress are often made in the context of highly variable environments. Estimating how much of such variation can be accounted for by "class-effects" is challenging but essential for understanding the inter-individual differences in behavioural and cognitive traits studied in evolutionary psychology.

5.3 Social mobility, ageing and lifespan

Our model in section 5.1 highlights that lifespan is the product of biological ageing, as measured by intrinsic mortality and reproductive effort, and extrinsic mortality (see Glossary). Therefore, biological ageing and lifespan need not be negatively correlated within populations. Good nutritional state can accelerate biological ageing through the costs of reproduction, but good nutritional state can also be correlated with low extrinsic mortality. This can lead to longer lifespans but higher rates of biological ageing, and conversely to shorter lifespans but lower rates of biological ageing as measured when nutrition is controlled.

Although intrinsic mortality and biological ageing can be difficult to detect by observational studies alone (Stearns et al. 2000), it is possible to find molecular markers of senescence, which provide a proxy for intrinsic mortality rates and biological ageing. For instance, a molecular analysis of telomere attrition in meerkats (*Suricata suricatta*), a measure of senescence rates, has shown that dominant female meerkats enjoy longer lifespans but suffer faster telomere loss and therefore higher rates of biological ageing (Cram et al. 2018). This contrasts with subordinate female meerkats, which show slower rates of telomere loss but shorter lifespans.

There are at least two potential hypotheses to explain this unusual positive association between biological ageing and lifespan. The first focuses on the fact that dominant females experience better breeding conditions and live in safer conditions than their subordinates, which is in part maintained by the aggressive behaviour of dominant females towards subordinate females. Dominant females experience better breeding conditions because they enjoy better nutritional state and higher offspring viability, and they experience safer conditions because they occupy inner areas of the colony that are less exposed to extrinsic mortality

factors. This hypothesis hinges on how dominance affects the relationship between breeding conditions (i.e. fecundity effects) and safety (i.e. survival effects). More specifically, it assumes that dominant mothers enjoy both better breeding conditions and lower extrinsic mortality than subordinate females, but dominance confers proportionally better breeding conditions than safety (i.e. extrinsic mortality). If the hypothesis is correct, we expect that dominant females have higher breeding rates – and therefore higher ageing rates – as well as longer lifespans, while subordinate meerkats should have fewer breeding attempts – therefore lower ageing rates – and shorter lifespans.

The second hypothesis, which we can call the *social mobility hypothesis*, proposes that subordinate females exercise reproductive restraint as a dominance-seeking strategy. Under this hypothesis subordinate females show slower rates of biological ageing because of reproductive restraint rather than because of relatively worse breeding conditions - they are cautiously waiting for an opportunity to move up the dominance hierarchy when the dominant female dies.

From the verbal description of these two hypotheses it is difficult to evaluate which of them best explains the observed patterns in meerkats. In such cases, mathematical models provide a valuable research tool, partially because the differences between the competing hypotheses are quantitative, and because it is difficult to determine how the different life history variables are intertwined.

To formalise the two hypotheses, we elaborate on a model (Rodrigues 2018b) that studies the evolution of rank-dependent life history (see Appendix B for details). From the analysis of the model, we find support for the *social mobility*

hypothesis. In particular, the model shows that higher levels of inequality in fertility rates and lifespan between mothers of different ranks are more likely to be associated with faster biological ageing in dominant (high-ranked) females when the social mobility scenario is considered (Fig. 4). The competing hypothesis, i.e. the *social immobility hypothesis*, can also lead to faster rates of biological ageing among dominant females, but this is associated with less inequality in biological ageing and lifespan among mothers of different ranks. Overall, our model suggests that the faster rates of biological ageing of dominant meerkats can arise because of reproductive restraint of subordinate meerkats as a status-acquisition strategy.

Other evidence also supports the social mobility hypothesis. If it is correct, then the benefits of waiting must be substantial. In meerkats, this is determined by tenure length and fertility rate of dominant females. Both are likely to be significant: the tenure of dominant females is greater than that of males (Thorley et al. 2020), and the fertility of dominant meerkats is relatively large. Moreover, if the hypothesis is correct, then waiting must improve the chances of dominance-acquisition. This is exactly what happens: older subordinate females are more likely to inherit the dominant position than younger subordinate females (Sharp and Clutton-Brock 2011).

[place Figure 4 here]

Although relatively simple, the model identifies the causal effects of empirically measurable variables – such as biological ageing, extrinsic mortality, and nutrition – on lifespan and other life history traits. The meerkat study illustrates how these

variables can be measured in natural populations by combining molecular and statistical analysis (Thorley et al. 2020).

The model also suggests a different explanation for suites of life history traits. While most models in the fast-slow continuum literature propose a positive association between fertility rates and aggression and a negative association between these traits and lifespan and senescence, our model suggests a contrasting scenario in which fertility rate, aggression, lifespan and senescence are positively correlated. Under this scenario, aggression by dominant females is part of a maintenance-of-status strategy that operates when societies are organised in hierarchies. This association among status, reproductive success, and aggression is shared by several animal species, including our closest relatives (Pusey and Schroepfer-Walker 2013, Fedigan et al. 2014), and has been observed in some human societies (Chagnon 1988, Glowacki and Wrangham 2015).

Overall, the model supports the idea that social gradients have significant impact on life histories. They probably emerge through plastic responses to multiple social and ecological factors. Social gradients are widespread among human populations, shape much of human sociality and life history, and have been associated, for example, with a predisposition to develop non-communicable diseases. Despite this, models that explore the impact of social gradients on life history evolution are relatively scarce. Further work on this area is required.

5.4 Diversified bet-hedging

Above, we discussed a particular form of bet-hedging called conservative bet-hedging (see Glossary), and we showed how this can affect traits that are

associated with either faster of slower life histories. Diversified bet-hedging is a different form of hedging one's bets that may evolve when individuals face fluctuating environments (Seger and Brockmann 1987, Starrfelt and Kokko 2012). Seger and Brockmann (1987) showed that when environments fluctuate from generation to generation, phenotypes with lower mean reproductive success can nevertheless evolve if their fitness is uncorrelated with the fitness of the remainder of the population. That is, it pays to be different in environments where others perform poorly.

Because diversified bet hedging can maintain variation in the population, it has received some attention from evolutionary psychologists (e.g. Belsky 1997, Belsky and Pluess 2009, Frankenhuis et al. 2016). Because diversified bet-hedging can lead to contrasting strategies, it has the potential to influence life history traits in ways that are inconsistent with the fast-slow continuum as it is usually understood within evolutionary psychology. This is the case, for instance, when fecundity within any given environment is uncorrelated among individuals (Rodrigues and Stearns, *in preparation*). Current models of bet-hedging do not capture these effects because of their simplifying assumptions, which often include non-overlapping generations, lack of spatial or class structure, or density-independent regulation. More realistic models that address these shortcomings yield insights into how different forms of bet-hedging affect traits associated with the fast-slow continuum and test the robustness of hypotheses in evolutionary psychology that have been informed by older life-history models.

6. Discussion

6.1 The nature of life history theory

Life history theory is *not* a single statement about a fast-slow continuum. The fast-slow continuum is an observation of a pattern, not a prediction from a model; it is a pattern observed in comparisons among orders, families, genera, and species, not among individuals responding with plasticity to environmental variation encountered in the course of development; and where it has been observed in comparisons of higher taxonomic units, it accounts for some but by no means all of the variation in life history patterns (precise amounts vary among studies, which use different samples) – other dimensions of variation are significant, and patterns vary among clades.

Life history theory *is* an overarching set of ideas, an organizational paradigm, about what questions to ask, what assumptions to make, and what simplifications to accept. Within it, specific models can make different predictions about the responses to selection of population mean values of traits and of the plastic responses elicited as individuals develop from birth to adult to old age in interaction with their environments, a point also made elsewhere in this issue (Zietsch and Sidari 2019). The predictions for population mean values and for plastic developmental responses can be in opposite directions for the traits, age and size at maturation, that have been analysed in detail.

Note that the theoretical predictions for both population means and plastic responses of individuals assume that they are genetically based and have evolved in response to selection, with the plastic responses having evolved as developmental rules that are contingent on the environments encountered by individuals. Plasticity can also be caused for non-adaptive reasons that are not reflected in life history models.

One might claim that it does not matter what stimulates a hypothesis if it leads to testable predictions and a research program that makes progress. The danger, however, is that when a research program is getting the right answers for the wrong reasons, it will be a house built on sand.

Is the point of doing science to identify patterns or to understand causes? That deeper issue frames this conversation. Those who want to find patterns sometimes trim variable and inconsistent results, cleaning them up to fit a preconceived framework, as Procrustes trimmed his captives to fit his bed. In contrast, those who want to understand causes cherish exceptions for their power to reveal what is not yet understood, remaining agnostic until predictions have withstood strong tests.

6.2 The importance of formal models

For several reasons mathematical models have strengthened the foundations of evolutionary biology. First, our capacity to develop complex verbal arguments is limited. For example, Darwin's several attempts to explain verbally the evolution of 50:50 sex ratios were unsuccessful: "… I now see that the whole problem is so intricate that it is safer to leave its solution for the future" (Darwin 1874, pp. 259, 260). This problem was later solved in a few lines of algebra (Edwards 1998, 2000). Second, our intuition is often wrong. For example, population viscosity was seen as promoting the evolution of altruism (Hamilton 1964), but a simple mathematical model uncovered a flaw in the argument (Taylor 1992): while population viscosity does promote interactions among relatives, as Hamilton suggested, it also increases competition for resources among relatives. These two opposing forces can cancel each other out, yielding no net effect of population

viscosity on the evolution of altruism. This insight has led to significant theoretical and empirical work that seeks to identify biological factors that modify this cancelation result (Lehmann and Rousset 2010, Rodrigues and Kokko 2016): it has had heuristic value. Third, the sensitivity of the predictions of mathematical models to their underlying assumptions is not easy to see in verbal arguments. For example, whereas single-population models predict lower reproductive effort in poor environments, meta-population models predict higher reproductive effort under the same conditions (Ronce and Olivieri 1997).

There are many approaches to model building, ranging from proof-of-concept models, that focus on a minimal number of key factors, to simulation models, that can include a large number of variables (Servedio et al. 2014). Each approach has its advantages and disadvantages: while the simplicity of proof-of-concept models gives great clarity and tractability but sacrifices realism, simulation models can include complicated details that more closely mirror reality but sacrifice clarity and tractability to do so (Levins 1968). One might start with simple models to decide which biological factors are important, then build more realistic assumptions into the model.

The mathematical models appropriate for evolutionary psychology include those that study variation among individuals within populations by focusing on life history patterns as a function of environmental gradients (i.e. reaction norms).

That said, there is room in science for a diversity of approaches, some of them data driven. The success of an approach can best be judged by the progress made by those using it over periods of decades.

6.3 The role of multiple working hypotheses

In addition to being more careful and nuanced in leveraging ideas from life history theory, evolutionary psychologists might more frequently consider multiple working hypotheses (Chamberlin 1995). The alternatives to a fast-slow life history dichotomy include at least: (a) current incentives shaping personal behaviour beyond what evolutionary psychologists call an "evoked" response to environmental conditions, (b) cultural evolution of norms that shape behaviour in ways that may or may not converge on those predicted by biological evolution for culturally inherited reasons (Sear et al. 2019); (c) mismatch arguments, which are non-adaptive – they predict maladaptation or pathology but not necessarily any consistent response to an environment that has not previously been frequently encountered (some evolutionary psychologists acknowledge mismatch but too often seem to expect consistent responses). Those taking this approach will remain agnostic about a hypothesis until plausible alternatives have been excluded with confidence, i.e. until the hypothesis has survived strong tests. For this reason, among others, we are surprised by the degree of confidence in life history theory expressed in the evolutionary psychology literature.

6.4 The logical power of interesting failures

In some of evolutionary psychology, the response to criticism of the use of the fast-slow continuum to explain behavioural, cognitive, and physiological responses to psychosocial stress has not been to welcome the failures as informative but instead to try to preserve the paradigm by explaining away the failures. In contrast, we believe that failures are informative, for science is the art of efficiently making interesting mistakes. We find this a useful working definition for those who are more interested in knowing the truth than they are in being right.

6.5 Concluding thoughts

Evolutionary psychologists might find it productive to debate these larger issues:

- What is a theory? We consider a theory to be the tracing of the consequences of explicit assumptions for specific predictions. Such analysis can only be done mathematically when relationships are quantitative.
- What constitutes an explanation? We consider an explanation to be a body of work that connects quantitative theory to quantitative tests in which alternative hypotheses are explicitly evaluated against reliable evidence in tests that are strong enough to change ideas. Explanations are the ideas that remain standing when repeatedly confronted with strong tests.
- What kinds of evidence are admissible? We give greater credence to experiments than to observational studies, and within observational studies, greater credence to prospective, randomized, case-control studies than to those in which such methods are not used.

A self-critical conversation about such issues could strengthen the field.

7. Acknowledgements

We thank Willem Frankenhuis, Daniel Nettle, Marco del Giudice, and two anonymous reviewers for their constructive comments.

8. References

Ackermann, M., A. Schauerte, S. C. Stearns, and U. Jenal. 2007. Experimental evolution of aging in a bacterium. Bmc Evolutionary Biology **7**.

Belsky, J. 1997. Variation in susceptibility to environmental influence: An evolutionary argument. Psychological Inquiry **8**:182-186.

Belsky, J., and M. Pluess. 2009. Beyond diathesis stress: Differential susceptibility to environmental influences. Psychological bulletin **135**:885-908.

Bernoulli, D. 1954. Exposition of a new theory on the measurement of risk. Econometrica **22**:23-36.

Bulmer, M. G. 1985. Selection for iteroparity in a variable environment. The American naturalist **126**:63-71.

Bulmer, M. G. 1994. Theoretical evolutionary ecology. Sinauer Associates, Sunderland, Massachusetts.

Chagnon, N. A. 1988. Life Histories, blood revenge, and warfare in a tribal population. Science **239**:985-992.

Chamberlin, T. 1995. Historical essay - The method of multiple working hypotheses (reprinted from Science, 1890). Journal Of Geology **103**:349-354.

Charlesworth, B. 1994. Evolution in age-structured populations. 2nd ed edition. Cambridge University Press, Cambridge, UK.

Charnov, E. L., and W. M. Schaffer. 1973. Life-history consequences of natural-selection - Cole's result revisited. The American naturalist **107**:791-793.

Cole, L. C. 1954. The population consequences of life history phenomena. Quarterly Review Of Biology **29**:103-137.

Cram, D. L., P. Monaghan, R. Gillespie, B. Dantzer, C. Duncan, H. Spence-Jones, and T. Clutton-Brock. 2018. Rank-related contrasts in longevity arise from extra-group excursions not delayed senescence in a cooperative mammal. Current Biology **28**:2934-2939.e2934.

Darwin, C. R. 1874. The descent of man, and selection in relation to sex. 2 edition. John Murray, London, UK.

Del Giudice, M., S. W. Gangestad, and H. S. Kaplan. 2015. Life History Theory and Evolutionary Psychology. Pages 1-27 *in* D. M. Buss, editor. The Handbook of Evolutionary Psychology.

Edwards, A. W. F. 1998. Natural selection and the sex ratio: Fisher's sources. The American naturalist **151**:564-569.

Edwards, A. W. F. 2000. Carl Düsing (1884) on the regulation of the sex-ratio. Theoretical Population Biology **58**:255-257.

Ellis, B. J., A. J. Figueredo, B. H. Brumbach, and G. L. Schlomer. 2009. Fundamental dimensions of environmental risk. Human Nature **20**:204-268.

Euler, L. 1970. A general investigation into the mortality and multiplication of the human species. Theoretical Population Biology **1**:307-314.

Feldblum, Joseph T., Emily E. Wroblewski, Rebecca S. Rudicell, Beatrice H. Hahn, T. Paiva, M. Cetinkaya-Rundel, Anne E. Pusey, and Ian C. Gilby. 2014. Sexually coercive male chimpanzees sire more offspring. Current Biology **24**:2855-2860.

Fisher, R. A. 1930. The genetical theory of natural selection. Oxford University Press, Oxford, UK.

Frank, S. A. 1998. Foundations of Social Evolution. Princeton University Press, Princeton, NJ.

Frank, S. A. 2012. Natural selection. III. Selection versus transmission and the levels of selection. Journal of Evolutionary Biology **25**:227-243.

Frankenhuis, W. E., K. Panchanathan, and J. Belsky. 2016. A mathematical model of the evolution of individual differences in developmental plasticity arising through parental bet-hedging. Developmental Science **19**:251-274.

Frisch, R., and R. Revelle. 1970. Height and weight at menarche and a hypothesis of critical body weights and adolescent events. Science **169**:397-399.

Gadgil, M., and W. H. Bossert. 1970. Life historical consequences of natural selection. The American naturalist **104**:1-24.

Glowacki, L., and R. Wrangham. 2015. Warfare and reproductive success in a tribal population. Proceedings of the National Academy of Sciences **112**:348-353.

Grafen, A. 2007. The formal Darwinism project: a mid-term report. Journal of Evolutionary Biology **20**:1243-1254.

Hamilton, W. D. 1964. The genetical evolution of social behaviour. I. Journal of theoretical biology **7**:1-16.

Hurt, L. S., C. Ronsmans, and S. L. Thomas. 2006. The effect of number of births on women's mortality: Systematic review of the evidence for women who have completed their childbearing. Population Studies-a Journal of Demography **60**:55-71.

Jeschke, J. M., and H. Kokko. 2008. The roles of body size and phylogeny in fast and slow life histories. Evolutionary Ecology **23**:867-878.

Kawecki, T. J., and S. C. Stearns. 1993. The evolution of life histories in spatially heterogeneous environments - Optimal reaction norms revisited. Evolutionary Ecology **7**:155-174.

Lehmann, L., and F. Rousset. 2010. How life history and demography promote or inhibit the evolution of helping behaviours. Philosophical Transactions of the Royal Society B-Biological Sciences **365**:2599-2617.

Levins, R. 1968. Evolution in Changing Environments. Princeton University Press, Princeton, NJ.

Lieberman, D. E. 2013. The story of the human body: Evolution, health and disease. Allen Lane, London, UK.

Lotka, A. J. 1907. Studies on the mode of growth of material aggregates. American Journal of Science **24**:141-159.

MacArthur, R. H., and E. O. Wilson. 1967. *The Theory of Island Biogeography*. Princeton University Press, Princeton NJ.

Marty, L., U. Dieckmann, M.-J. Rochet, and B. Ernande. 2011. Impact of environmental covariation in growth and mortality on evolving maturation reaction norms. The American naturalist **177**:E93-118.

Maynard Smith, J. 1982. Evolution and the theory of games. Cambridge University Press, Cambridge, UK.

Miles, D. B., and A. E. Dunham. 1992. Comparative analyses of phylogenetic effects in the life-history patterns of Iguanid reptiles. The American naturalist **139**:848-869.

Murphy, G. I. 1968. Pattern in life history and the environment. The American naturalist **102**:391-403.

Nettle, D., and M. Bateson. 2015. Adaptive developmental plasticity: what is it, how can we recognize it and when can it evolve? Proceedings of the Royal Society B **282**:20151005.

Nettle, D., and W. E. Frankenhuis. 2019. The evolution of life history theory: Bibliometric analysis of an interdisciplinary research area. Proceedings of the Royal Society B **286**:20190040.

Otto, S. P., and T. Day. 2007. A Biologist's Guide to Mathematical Modeling in Ecology and Evolution. Princeton University Press, Princeton, NJ.

Parker, G. A. 1989. Hamilton's rule and conditionality. Ethology Ecology & Evolution **1**:195-211.

Pianka, E. R. 1970. On r- and K-selection. The American naturalist **104**:592-597.

Price, G. R. 1970. Selection and covariance. Nature **227**:520-521.

Pulliam, H. R. 1988. Sources, sinks, and population regulation. The American naturalist **132**:652-661.

Pusey, A. E., and K. Schroepfer-Walker. 2013. Female competition in chimpanzees. Philosophical Transactions of the Royal Society B **368**:20130077.

Rodrigues, A. M. M. 2018a. Demography, life history and the evolution of age-dependent social behaviour. Journal of Evolutionary Biology **31**:1340-1353.

Rodrigues, A. M. M. 2018b. The evolution of class-dependent reproductive effort in humans and other animals. bioRxiv:449868.

Rodrigues, A. M. M., and A. Gardner. 2016. The constant philopater hypothesis: a new life history invariant for dispersal evolution. Journal of Evolutionary Biology **29**:153-166.

Rodrigues, A. M. M., and H. Kokko. 2016. Models of social evolution: can we do better to predict "who helps whom to achieve what"? Philosophical Transactions of the Royal Society B-Biological Sciences **371**:20150088.

Ronce, O., and I. Olivieri. 1997. Evolution of reproductive effort in a metapopulation with local extinctions and ecological succession. The American Naturalist **150**:220-249.

Schaffer, W. M. 1974. Selection for optimal life histories: The effects of age structure. Ecology **55**:291-303.

Sear, R., P. Sheppard, and D. A. Coall. 2019. Cross-cultural evidence does not support universal acceleration of puberty in father-absent households. Philosophical Transactions Of The Royal Society Of London Series B-Biological Sciences **374**:20180124-20180114.

Seger, J., and H. J. Brockmann. 1987. What is bet-hedging? Pages 182-211 *in* P. H. Harvey and L. Partridge, editors. Oxford Surveys in Evolutionary Biology. Oxford University Press, Oxford, UK.

Servedio, M. R., Y. Brandvain, S. Dhole, C. L. Fitzpatrick, E. E. Goldberg, C. A. Stern, J. Van Cleve, and D. J. Yeh. 2014. Not Just a Theory-The Utility of Mathematical Models in Evolutionary Biology. Plos Biology **12**:5.

Sharp, S. P., and T. H. Clutton-Brock. 2011. Reluctant challengers: why do subordinate female meerkats rarely displace their dominant mothers? Behavioral Ecology **22**:1337-1343.

Starrfelt, J., and H. Kokko. 2012. Bet-hedging--a triple trade-off between means, variances and correlations. Biological Reviews **87**:742-755.

Stearns, S. C. 1983. The influence of size and phylogeny on patterns of covariation among life-history traits in the mammals. Oikos **41**:173-187.

Stearns, S. C. 1989. The Evolutionary Significance of Phenotypic Plasticity: Phenotypic sources of variation among organisms can be described by developmental switches and reaction norms. BioScience **39**:436-445.

Stearns, S. C. 1992. The Evolution of Life Histories. Oxford University Press, Oxford, UK.

Stearns, S. C., M. Ackermann, M. Doebeli, and M. Kaiser. 2000. Experimental evolution of aging, growth, and reproduction in fruitflies. Proceedings of the National Academy of Sciences USA **97**:3309.

Stearns, S. C., S. G. Byars, D. R. Govindaraju, and D. Ewbank. 2010. Measuring selection in contemporary human populations. Nature Reviews Genetics **11**:611-622.

Stearns, S. C., and J. C. Koella. 1986. The Evolution of Phenotypic Plasticity in Life-History Traits - Predictions of Reaction Norms for Age and Size at Maturity. Evolution **40**:893-913.

Stearns, S. C., and R. Medzhitov. 2015. Evolutionary Medicine. Sinauer, Sunderland, MA.

Taylor, P. D. 1992. Altruism in viscous populations — an inclusive fitness model. Evolutionary Ecology **6**:352-356.

Thorley, J., C. Duncan, S. P. Sharp, D. Gaynor, M. B. Manser, and T. Clutton-Brock. 2020. Sex-independent senescence in a cooperatively breeding mammal. Journal of Animal Ecology **n/a**.

Tinbergen, N. 1963. On aims and methods of Ethology. Zeitschrift für Tierpsychologie **20**:410-433.

Wilbur, H. M. H., and V. H. W. V. Rudolf. 2006. Life-history evolution in uncertain environments: bet hedging in time. The American naturalist **168**:398-411.

Wright, S. 1931. Evolution in Mendelian populations. Genetics **16**:97-159.

Wright's, S. 1931. Evolution in Mendelian populations. Genetics **16**:97-159.

Yamaguchi, Y. 1985. Sex ratios of an aphid subject to local mate competition with variable maternal condition. Nature **318**:460-462.

Zietsch, B., and M. Sidari. 2019. A critique of life history approaches to human trait variation. Evolution And Human Behavior:1-51.

Glossary

Adaptive evolution -- Changes in gene frequency that result from the action of Natural Selection, i.e. from correlations with reproductive success (Price 1970).

Age-structured populations -- In such a population the individuals present at any one time were born over a range of different times, and their fertility and survival depend on age (Charlesworth 1994).

Conservative bet-hedging -- Such strategies are recognized by a reduction in individual-level *variance* in fitness, often measured by the geometric mean (Starrfelt and Kokko 2012).

Constitutive traits -- Traits whose development is insensitive to environmental conditions; such traits are buffered against both genetic and environmental variation.

Developmental plasticity -- Environmental effects during development that produce long lasting effect in phenotypes (Nettle and Bateson 2015).

Diversified bet-hedging -- Such strategies are recognized by a reduction in between-individual *correlations* in fitness. Conservative and diversified bet-hedging are not mutually exclusive. Both can often be found in the same population (Starrfelt and Kokko 2012).

Evolutionary change -- Changes in gene frequency (Frank 2012).

Extrinsic and intrinsic mortality -- Extrinsic mortality is the portion of total mortality that can be directly ascribed to the effects of the environment, such as predation, infection, starvation. Intrinsic mortality is the portion of total mortality that is internal to the individual, such as cancer, cardiovascular disease and behaviour (Stearns et al. 2000, Stearns and Medzhitov 2015).

Fitness -- There are many definitions: in general, the maximand of Natural Selection (Grafen 2007); in particular models, r, K, w, or G; a convenient approximation for empiricists is lifetime reproductive success, whose ease of

measurement comes at the cost of missing multi-generation effects (see bet-hedging).

Mismatch -- Recognized when evolutionary traits that were once adaptive are no longer beneficial in the current environment (Lieberman 2013).

Non-overlapping generations -- Populations in which the current generation dies before the new generation start breeding, as opposed to overlapping generations, which refer to populations in which individuals of many ages are present and may be breeding.

Phenotypic plasticity -- The production of alternative phenotypes by single genotypes in response to variation in the environment (Stearns 1989).

Reaction norm -- The range of potential phenotypes a single genotype can develop when exposed to a range of environmental conditions (Stearns and Koella 1986).

Sources and sinks -- A population inhabiting patches in a spatially heterogeneous environment can consist of "sink" sub-populations maintained by immigration from "source" sub-populations: sinks absorb and the individuals in them are maladapted to local circumstances, sources export excess individuals (Pulliam 1988).

Tinbergen's four questions -- Four levels of analysis in evolutionary biology: function, phylogeny, development, and mechanism. Our understanding is not complete until all have been explored (Tinbergen 1963).

Unconditional phenotypes -- Phenotypes (or strategies) that arise from genes whose expression is independent of role or environment; genetic polymorphisms can be maintained in the population through frequency-dependent selection (Maynard Smith 1982, Parker 1989).

Vital rates -- The age-specific survival and fertility rates of a population (Charlesworth 1994).

Appendices

Content

Appendix A - Environment-dependent life history (section 5.1)

Appendix A.1. Models and methodology

We consider two types of models. While the first model assumes a population subdivided into a very large number of patches, the second model does not make such assumption. We first describe the model with patch structure, and we then describe the model without patch structure.

We assume a population subdivided into a very large number of patches (i.e. the infinite island model of Wright (1931)), and for simplicity we assume asexually-reproducing and haploid individuals. There are n breeders per patch. Patches experience different environments, such that a proportion p_i of the patches are in environment i, with $\sum_{i=1}^{Q} p_i = 1$, where Q is the total number of environments. Each breeder in environment i produce f_i offspring. Offspring may remain in the local patch with probability $1 - d_i$, and disperse with probability d_i. Dispersed offspring arrive at a random patch in the population with probability $1 - k$, where k is the cost of dispersal. Mothers in environment i survive with probability S_i, and die with probability $1 - S_i$. The death of mothers frees up breeding spots that are then contested at random by both native and immigrant offspring. After competition, offspring that obtain a breeding spot become adults and initiate their reproductive lives. At this stage, the life-cycle of our model species resumes.

The model without spatial structure follows a similar life-cycle, but some of the variables have a different interpretation. In the unstructured model, $1 - d_i$ is interpreted as the probability of remaining in the same environment, while d_i is the probability of moving to a random environment in the population. In addition, all offspring become adult breeders.

We are interested in understanding the evolution of reproductive effort. We assume a trade-off between maternal fecundity (i.e. f_i) and maternal survival (i.e. S_i; Stearns (1992)). In particular, we consider that fecundity and adult survival are mediated by reproductive effort, denoted by x_i, where $f_i = s_i x_i^{1/2}$, and

$S_i = \sigma_i(1 - x_i)^{1/2}$, where: and s_i is the nutritional state of a mother in environment *i* ; and σ_i is the baseline survival of a mother in environment *i*. We employ standard optimisation methods to determine the optimal reproductive effort strategies (Frank 1998, Otto and Day 2007). Optimal strategies are those strategies that cannot be invaded by slightly higher or slightly lower values of the optimal strategy (Otto and Day 2007). Below, we describe the different steps required to find these optimal strategies.

Appendix A.2. Reproductive success and reproductive value

Here, we start by defining the reproductive success of a focal mother. The fitness of a focal mother in any environment depends on her fecundity and survival. A focal mother produces two types of offspring, philopatric and dispersed offspring. Philopatric offspring remain in the local habitat, while dispersed offspring leave the local habitat. Let us first consider a mother's reproductive success through philopatric offspring, which is given by

$$w_i^{\phi} = \frac{f_i(1-d_i)}{nf_{i,o}(1-d_i)+\sum_{l=1}^{Q} r_l nf_{l,z}d_l(1-k)}. \quad \text{(A.2.1)}$$

where: f_i is the fecundity of the focal mother in environment *i*; $f_{j,o}$ is the average fecundity in the local patch; $f_{l,z}$ is the fecundity of a random mother in environment *l*.

The reproductive success of a focal mother in environment *i* through dispersed offspring is given by

$$w_{i\to j}^{\delta} = \frac{f_i d_i (1-k)}{n f_{j,z}(1-d_j) + \sum_{l=1}^{Q} p_l n f_{l,z} d_l (1-k)}. \quad \text{(A.2.2)}$$

The average reproductive success in the population is given by

$$\overline{w} = \sum_{i=1}^{Q} p_i \left(S_i + w_i^{\phi} o_i + \sum_{j=1}^{Q} w_{i\to j}^{\delta} p_j o_j \right) \quad \text{(A.2.3)}$$

where $o_i = (1 - S_i)n$ is the expected number of empty breeding sties in a patch in environment *i*.

The reproductive value of a mother in environment *i* via the production of offspring, both philopatric and dispersed, is given by

$$v_i^{f} = \frac{1}{\overline{w}} \left(w_i^{\phi} o_i v_i + \sum_{j=1}^{Q} w_{i\to j}^{\delta} p_j o_j v_j \right). \quad \text{(A.2.4)}$$

where v_i is the reproductive value of a focal mother in environment *i*, and where reproductive value gives the contribution of an individual to the future gene pool of the population (Fisher 1930). The reproductive value of a mother in environment *i* through her own survival is given by

$$v_i^{s} = \frac{1}{\overline{w}} S_i v_i, \quad \text{(A.2.5)}$$

The total reproductive value of a focal mother in environment *i* is then given by the reproductive value obtained via fecundity and the reproductive value obtained via survival, which is given by

$$v_{\mathrm{i}} = v_{\mathrm{i}}^{\mathrm{f}} + v_{\mathrm{i}}^{\mathrm{s}}. \tag{A.2.6}$$

Because reproductive value is a relative measure, we can set the reproductive value of mothers in environment 1 to one, and solve the remainder system of equations to find the reproductive values of other mothers. Thus, the full system of equations is given by

$$\begin{cases} v_1 = 1 \\ v_{\mathrm{i}}' = v_{\mathrm{i}} \end{cases} \quad , \mathrm{i} \in \{2, .., Q\}. \tag{A.2.7}$$

After replacing equation (A.2.6) in the system of equations (A.2.7), we can then solve that system.

Appendix A.3. Reproductive value of offspring

Alongside the reproductive value of adults, it is also useful to consider the reproductive value of offspring. The reproductive success of an offspring when she remains in the local patch is given by

$$W_{\mathrm{i}}^{\Phi} = \frac{(1-d_{\mathrm{i}})}{nf_{\mathrm{i,o}}(1-d_{\mathrm{i}}) + \sum_{\mathrm{l}=1}^{\mathrm{Q}} p_{\mathrm{l}} nf_{\mathrm{l,z}} d_{\mathrm{l}} (1-k)}. \tag{A.3.1}$$

The reproductive success of an offspring when she leaves the local patch is given by

$$W_{\mathrm{i}\to\mathrm{j}}^{\delta}=\frac{d_{\mathrm{i}}(1-k)}{nf_{\mathrm{j,o}}(1-d_{\mathrm{j}})+\sum_{\mathrm{l}=1}^{\mathrm{Q}}p_{\mathrm{l}}nf_{\mathrm{l,z}}d_{\mathrm{l}}(1-k)}. \tag{A.3.2}$$

The reproductive value of an offspring in environment *i* is then given by

$$V_{\mathrm{i}}=W_{\mathrm{i}}^{\phi}o_{\mathrm{i}}v_{\mathrm{i}}+\sum_{\mathrm{j}=1}^{\mathrm{Q}}W_{\mathrm{i}\to\mathrm{j}}^{\delta}p_{\mathrm{j}}o_{\mathrm{j}}v_{\mathrm{j}}. \tag{A.3.3}$$

Note that the reproductive value of a mother through the production of offspring is given by her fecundity times the reproductive value of each offspring, that is

$$v_{\mathrm{i}}=f_{\mathrm{i}}V_{\mathrm{i}}. \tag{A.3.4}$$

Appendix A.4. Selection gradients

We now want to obtain the effect of a slight increase in the breeding value of the focal individual, denoted by g_{α}, on the fitness of the focal individual (Frank 1998, Rodrigues 2018a). The selection gradient is given by

$$\frac{\mathrm{d}v_{\mathrm{i}}}{\mathrm{d}g_{\alpha}}=\frac{\mathrm{d}v_{\mathrm{i}}^{\mathrm{f}}}{\mathrm{d}g_{\alpha}}+\frac{\mathrm{d}v_{\mathrm{i}}^{\mathrm{s}}}{\mathrm{d}g_{\alpha}}. \tag{A.4.1}$$

If we expand the right-hand side of equation (A.4.1), we obtain

$$\frac{\mathrm{d}v_\mathrm{i}}{\mathrm{d}g_\alpha} = \frac{1}{\bar{w}}\left(\frac{\mathrm{d}S_\mathrm{i}}{\mathrm{d}g_\alpha}v_\mathrm{i} + \frac{\mathrm{d}w_\mathrm{i}^{\phi}}{\mathrm{d}g_\alpha}o_\mathrm{i}v_\mathrm{i} + \sum_{\mathrm{j}=1}^{\mathrm{Q}}\frac{\mathrm{d}w_{\mathrm{i}\to\mathrm{j}}^{\delta}}{\mathrm{d}g_\alpha}p_\mathrm{j}o_\mathrm{j}v_\mathrm{j}\right). \quad \text{(A.4.2)}$$

We can now expand the derivatives with respect to the breeding value of the focal individual. This is given by

$$\frac{dv_\mathrm{i}}{dg_\alpha} = \frac{1}{\bar{w}}\left(\frac{\partial S_\mathrm{i}}{\partial x_\alpha}\frac{\mathrm{d}x_\alpha}{\mathrm{d}g_\alpha}\frac{\mathrm{d}g_\alpha}{\mathrm{d}g_\rho}v_\mathrm{i} + \frac{\partial w_\mathrm{i}^{\phi}}{\partial x_\alpha}\frac{\mathrm{d}x_\alpha}{\mathrm{d}g_\alpha}\frac{\mathrm{d}g_\alpha}{\mathrm{d}g_\rho}o_\mathrm{i}v_\mathrm{i} + \sum_{\mathrm{j}=1}^{\mathrm{Q}}\frac{\partial w_{\mathrm{i}\to\mathrm{j}}^{\delta}}{\partial x_\alpha}\frac{\mathrm{d}x_\alpha}{\mathrm{d}g_\alpha}\frac{\mathrm{d}g_\alpha}{\mathrm{d}g_\rho}p_\mathrm{j}o_\mathrm{j}v_\mathrm{j}\right). \quad \text{(A.4.3)}$$

where g_ρ is the breeding value of the recipient. The partial derivative means that a slight increase in the breeding value of the focal individual influences the fitness of the recipient-*i* by the impact of the individual-α phenotype on the fitness of the recipient. The derivative of the breeding value on the phenotype, i.e. the genotype-phenotype mapping, can be arbitrarily set to one (i.e. $\mathrm{d}x_\alpha/\mathrm{d}g_\alpha = 1$). The slope of the focal individual's breeding value on the recipient's breeding value gives the coefficients of consanguinity between the two individuals (Bulmer 1994). When the recipient of the behaviour is the focal individual, then the slope is one. Because we assume a very large number of individuals in each patch (i.e. $n \to \infty$), this slope is zero when the recipients of the behaviour are all other individuals in the patch. Thus, if we expand equation (A.4.3), we obtain

$$\frac{dv_\mathrm{i}}{dg_\alpha} = \frac{1}{\bar{w}}\left(\frac{\partial S_\mathrm{i}}{\partial g_\alpha}v_\mathrm{i} + \frac{\partial f_\mathrm{i}}{\partial g_\alpha}\left(\frac{(1-d_\mathrm{i})}{nf_{\mathrm{i,o}}(1-d_\mathrm{i})+\sum_{\mathrm{l}=1}^{\mathrm{Q}}p_\mathrm{l}nf_{\mathrm{l,z}}d_\mathrm{l}(1-k)}o_\mathrm{i}v_\mathrm{i} + \sum_{\mathrm{j}=1}^{\mathrm{Q}}\frac{d_\mathrm{i}(1-k_\mathrm{i})}{nf_{\mathrm{j,o}}(1-d_\mathrm{j})+\sum_{\mathrm{l}=1}^{\mathrm{Q}}p_\mathrm{l}nf_{\mathrm{l,z}}d_\mathrm{l}(1-k)}p_\mathrm{j}o_\mathrm{j}v_\mathrm{j}\right)\right). \quad \text{(A.4.4)}$$

If we plug in equation (A.3.3) into equation (A.4.4), we obtain the selection gradient

$$\frac{dv_{\mathrm{i}}}{dg_{\alpha}} = \frac{1}{\bar{w}}\left(\frac{\partial S_{\mathrm{i}}}{\partial g_{\alpha}} v_{\mathrm{i}} + \frac{\partial f_{\mathrm{i}}}{\partial g_{\alpha}} V_{\mathrm{i}}\right). \quad \text{(A.4.5)}$$

If we define the cost of an increase in reproductive effort as

$$c_{\mathrm{i}} = -\frac{\partial S_{\mathrm{i}}}{\partial g_{\alpha}}, \quad \text{(A.4.6)}$$

and if we define the benefit of an increase in reproductive effort as

$$b_{\mathrm{i}} = \frac{\partial f_{\mathrm{i}}}{\partial g_{\alpha}}, \quad \text{(A.4.7)}$$

equation A.4.5 becomes

$$\frac{dv_{\mathrm{i}}}{dg_{\alpha}} = \frac{1}{\bar{w}}(-c_{\mathrm{i}} v_{\mathrm{i}} + b_{\mathrm{i}} V_{\mathrm{i}}). \quad \text{(A.4.8)}$$

Appendix A.5. Optimal strategies

To find the optimal reproductive effort strategies, i.e. x_{i}^{*}, we use an iterative algorithm (see Rodrigues (2018a) for details). We first find the reproductive value of each individual assuming a neutral population (equations A.2.7). We then determine the selection gradients (equation A.4.8). If the selection gradient is positive, we update the resident reproductive effort strategy by a slightly higher

value. If the selection gradient is negative, we update the resident reproductive effort strategy by a slightly lower value. We repeat this iterative process until we find the equilibrium values of the reproductive effort strategies.

Appendix A.6. Reproductive success and reproductive value in unstructured populations

In unstructured populations, the reproductive success of a focal mother in environment *i* through offspring that remain in environment *i* is given by

$$w_{\mathrm{i}}^{\Phi} = f_{\mathrm{i}}(1 - d_{\mathrm{i}}). \tag{A.6.1}$$

The reproductive success of a focal mother in environment *i* through offspring that become breeders in environment *j* is given by

$$w_{\mathrm{i}\to\mathrm{j}}^{\delta} = f_{\mathrm{i}} d_{\mathrm{i}}(1 - k). \tag{A.6.2}$$

The average reproductive success in the population is given by

$$\bar{w} = \sum_{\mathrm{i}=1}^{\mathrm{Q}} p_{\mathrm{i}}\left(S_{\mathrm{i}} + w_{\mathrm{i}}^{\Phi} + \sum_{\mathrm{j}=1}^{\mathrm{Q}} w_{\mathrm{i}\to\mathrm{j}}^{\delta} p_{\mathrm{j}}\right). \tag{A.6.3}$$

The reproductive value of a mother in environment *i* via the production of offspring is given by

$$v_{\mathrm{i}}^{\mathrm{f}} = \frac{1}{\bar{w}}\left(w_{\mathrm{i}}^{\phi} v_{\mathrm{i}} + \sum_{\mathrm{j}=1}^{\mathrm{Q}} w_{\mathrm{i}\to\mathrm{j}}^{\delta} p_{\mathrm{j}} v_{\mathrm{j}}\right). \quad \text{(A.6.4)}$$

The reproductive success of an offspring when she remains in the local environment is given by

$$W_{\mathrm{i}}^{\phi} = (1 - d_{\mathrm{i}}). \quad \text{(A.6.5)}$$

The reproductive success of an offspring when she leaves the local environment to a random environment is given by

$$W_{\mathrm{i}\to\mathrm{j}}^{\delta} = d_{\mathrm{i}}(1 - k). \quad \text{(A.6.6)}$$

The reproductive value of an offspring in environment *i* is then given by

$$V_{\mathrm{i}} = W_{\mathrm{i}}^{\phi} v_{\mathrm{i}} + \sum_{\mathrm{j}=1}^{\mathrm{Q}} W_{\mathrm{i}\to\mathrm{j}}^{\delta} p_{\mathrm{j}} v_{\mathrm{j}}. \quad \text{(A.6.7)}$$

As in the patch-structured model, we can now find the optimal reproductive effort strategies using equations (A.2.7) and (A.4.8) and the iterative process described in Appendix E.

Appendix A.7. Figure and results

For each model, we consider two scenarios. In the first scenario, we consider a nutritional gradient, where the fertility of mothers increases with nutritional

state. In particular, we assume that $s_1 = 0.7$, $s_2 = 0.8$, $s_3 = 0.9$, and $s_4 = 1.0$. The results for the nutritional gradient are depicted in Fig. 3, panels A and C, in the main text. In a second scenario, we consider a gradient in maternal extrinsic mortality. In particular, we assume that $\sigma_1 = 0.90$, $\sigma_2 = 0.85$, $\sigma_3 = 0.80$, and $\sigma_4 = 0.75$. The results for the maternal extrinsic mortality scenario are depicted in Fig. 3, panels B and D, in the main text. The parameter values used in Fig. 3 are as follows: $n \to \infty$, $d_i = 0.1$, $k = 0.5$, $Q = 4$, $p_i = 0.25$.

Appendix B – Social mobility, ageing and lifespan (section 5.3)

We follow the model of Rodrigues (2018b) that studies the evolution of reproductive effort in stratified societies. Here, we provide a brief description of the model and we refer to Rodrigues (2018b) for an in-depth outline of the methods and analysis of the model.

Appendix B.1. Life-cycle

We consider Wright's (1931) infinite island model and a population composed of asexually-reproducing haploid breeding mothers. Each patch is inhabited by a group of four females that form a society stratified into four different ranks. Rank-i females give birth to f_i offspring who become juveniles with probability s_i, which means that of the f_i offspring only $F_i = f_i s_i$ each the juvenile stage. Adult females reproduce regularly, and they survive each breeding cycle with probability S_i. Offspring that reach the juvenile stage, either remain in the local patch with

probability $1 - d_i$ or disperse with probability d_i. Disperses can suffer in-transit hazards such that only a fraction $1 - k$ survive dispersal.

We assume two types of social dynamics: social immobility and social mobility. Under the social immobility scenario, once females acquire a rank they keep it until they die, and therefore they experience neither upward nor downward social mobility. Under the social mobility scenario, females can experience upward social mobility when females in higher ranks die. Under this scenario, although females can improve their rank during the course of their lifetime, the relative rank of surviving females remains constant.

Under social immobility, surviving juveniles compete for the breeding places left vacant by the decreases mothers from which they inherit their social rank. Under social mobility, surviving adult mothers occupy the upper most social ranks, and therefore only the bottom most ranks remain available for juveniles, who compete for these breeding places. Irrespective of the social mobility scenario, juveniles who fail to obtain a breeding place die, while successful juveniles become adult breeders.

Appendix B.2. Reproductive effort

We assume a trade-off between the fecundity and survival of a mother. Thus, mothers who invest more in reproductive effort pay a survival cost. In particular, we assume that the fecundity of a mother is given by $f_i = z_i^{½}$ while her survival is given by $S_i = \sigma_i (1 - z_i)^{½}$, where z_i denotes her reproductive effort. The parameter σ_i

= $1 - \varepsilon_i$ denotes the extrinsic survivorship of rank-*i* mothers, where ε_i is the extrinsic mortality.

We are interested in finding the rank-dependent optimal reproductive effort strategy, which we denote by z_i^*. Optimal reproductive effort strategies are the values of reproductive effort z_i^* at which natural selection favours neither a slight increase nor a slight decrease in trait value (Otto and Day 2007).

Appendix B.3. Total mortality, ageing and extrinsic mortality

The total mortality rate of an individual is given by $M_i = 1 - S_i$. We partition total mortality into two components: extrinsic and intrinsic mortality. Thus, intrinsic mortality, denoted by ι_i, is given by $\iota_i = M_i - \varepsilon_i$. We assume that intrinsic mortality provides a proxy for ageing rate, and therefore $\alpha_i = \iota_i$, where α_i is the ageing rate of a rank-*i* individual.

Figure Captions

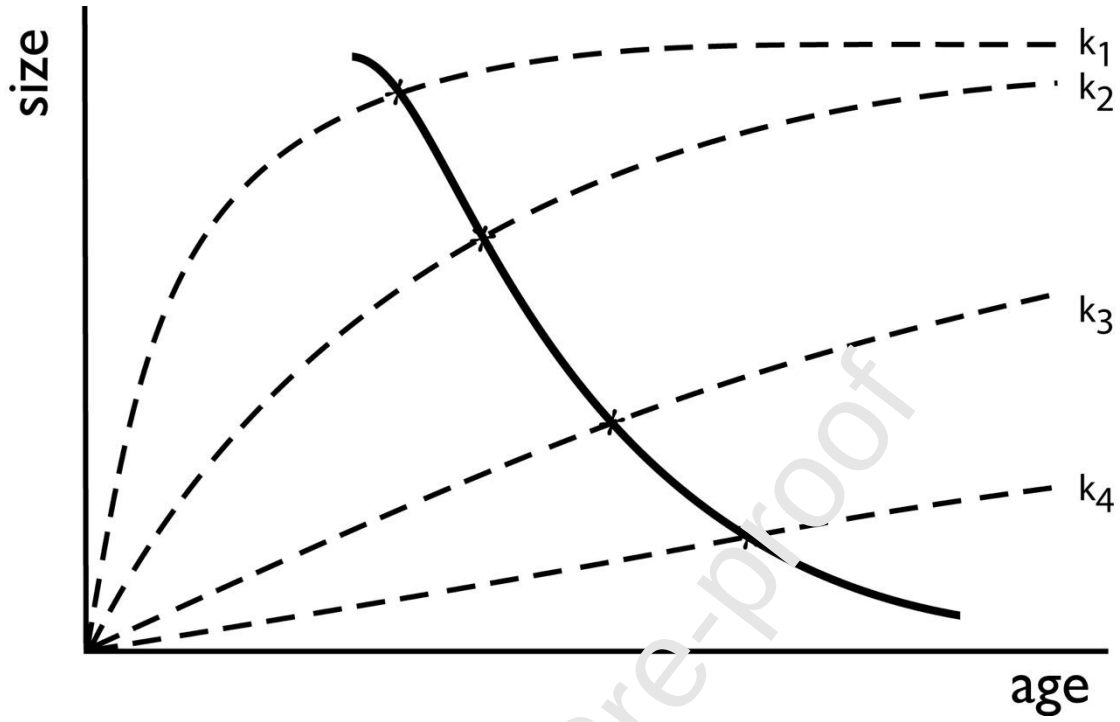
Fig. 1.- The predicted reaction norm for age and size at maturity is the thick black line. The dotted lines represent growth from fast (k_1) to slow (k_4). Such a maturation reaction norm is an evolved rule of thumb for when to mature conditionally, based on the environments encountered. Redrawn from (Stearns and Koella 1986).

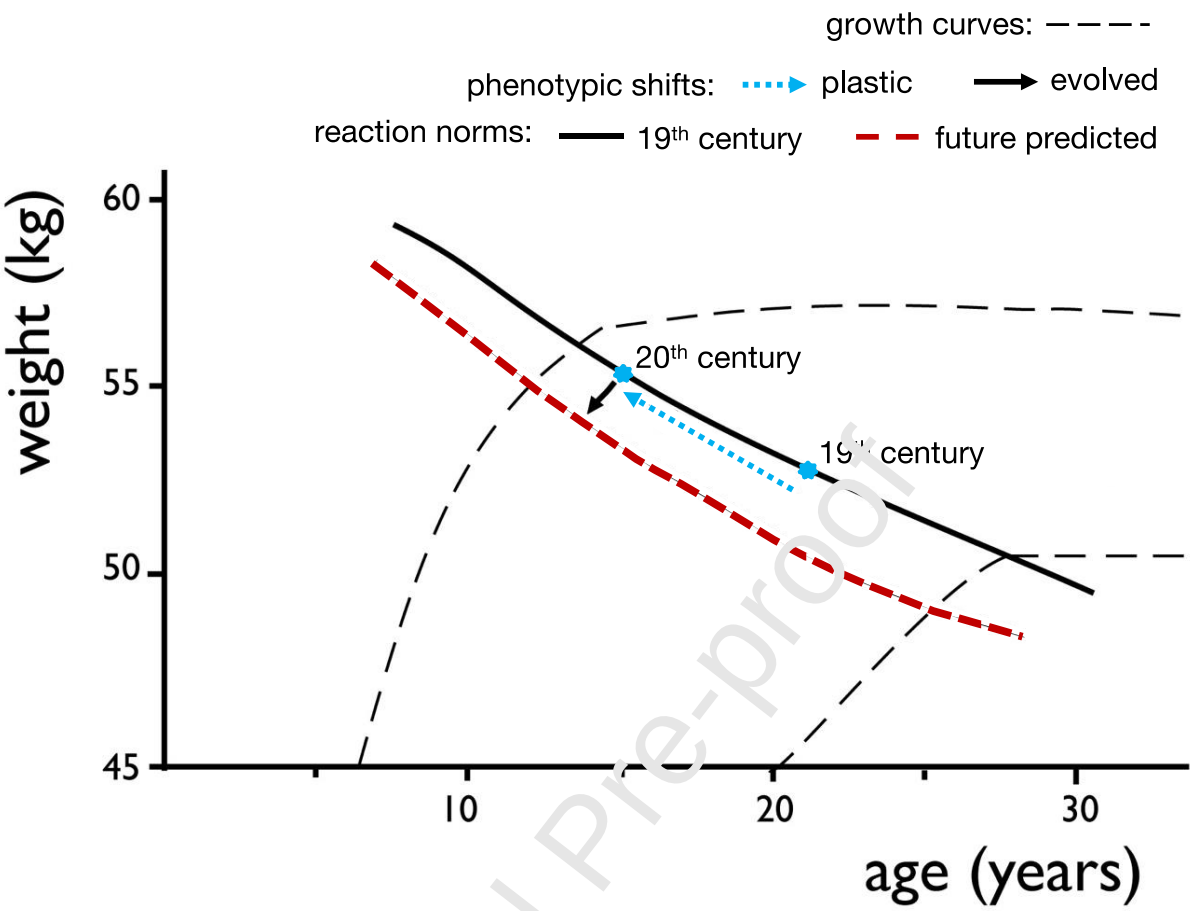
Fig. 2.- Optimal reaction norms for age and size at maturity in human females evoked by different growth rates (thin dashed lines). The upper, solid black reaction norm depicts the plastic shift from the 19th to the 20th century (blue dotted line), when better nutrition and growth caused women to mature larger and younger. The lower reaction norm (think red dashes) depicts the predicted evolved shift of the entire curve down and to the left. Redrawn from (Stearns and Koella 1986).

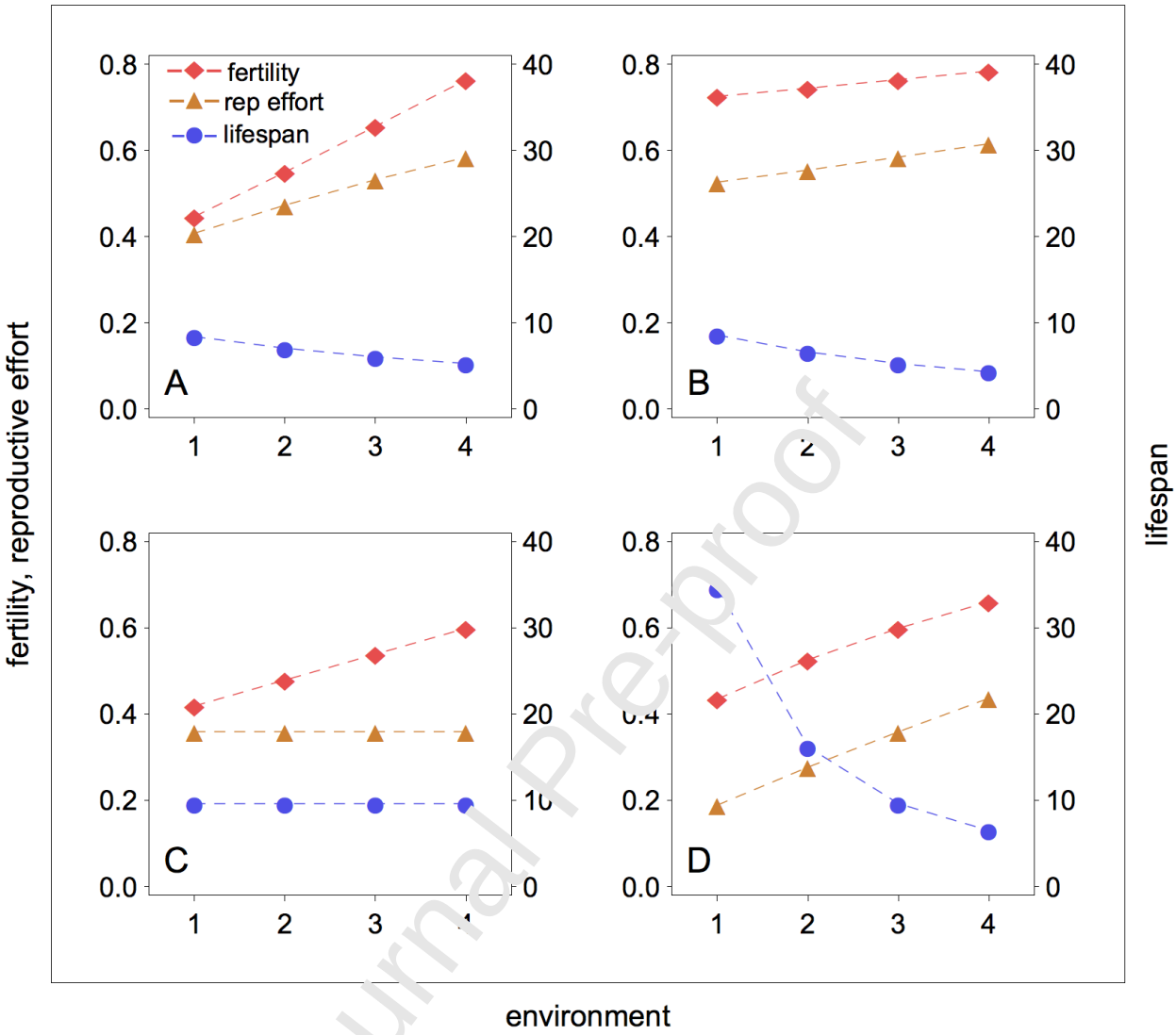
Fig. 3.- Reproductive effort, fertility rate and expected lifespan as a function of an environmental gradient (from 1 to 4; see Appendix for details). In the absence of spatial structure (panels A and B), a nutritional gradient (panel A) and a gradient in maternal extrinsic mortality (panel B) both lead to an inverse correlation between fertility rate and lifespan. When we consider spatial structure (panels C and D), these predictions change. In particular, maternal nutritional state has no impact on expected lifespan (panel C). In panels A and C, the environmental gradient is defined by $s_1 = 0.7$, $s_2 = 0.8$, $s_3 = 0.9$, and $s_4 = 1.0$; while in panels B and D, the environmental gradient is defined by $\sigma_1 = 0.90$, $\sigma_2 = 0.85$, $\sigma_3 = 0.80$, and $\sigma_4 = 0.75$, where s_i is the nutritional state of the mother, and σ_i the instantaneous survivorship.

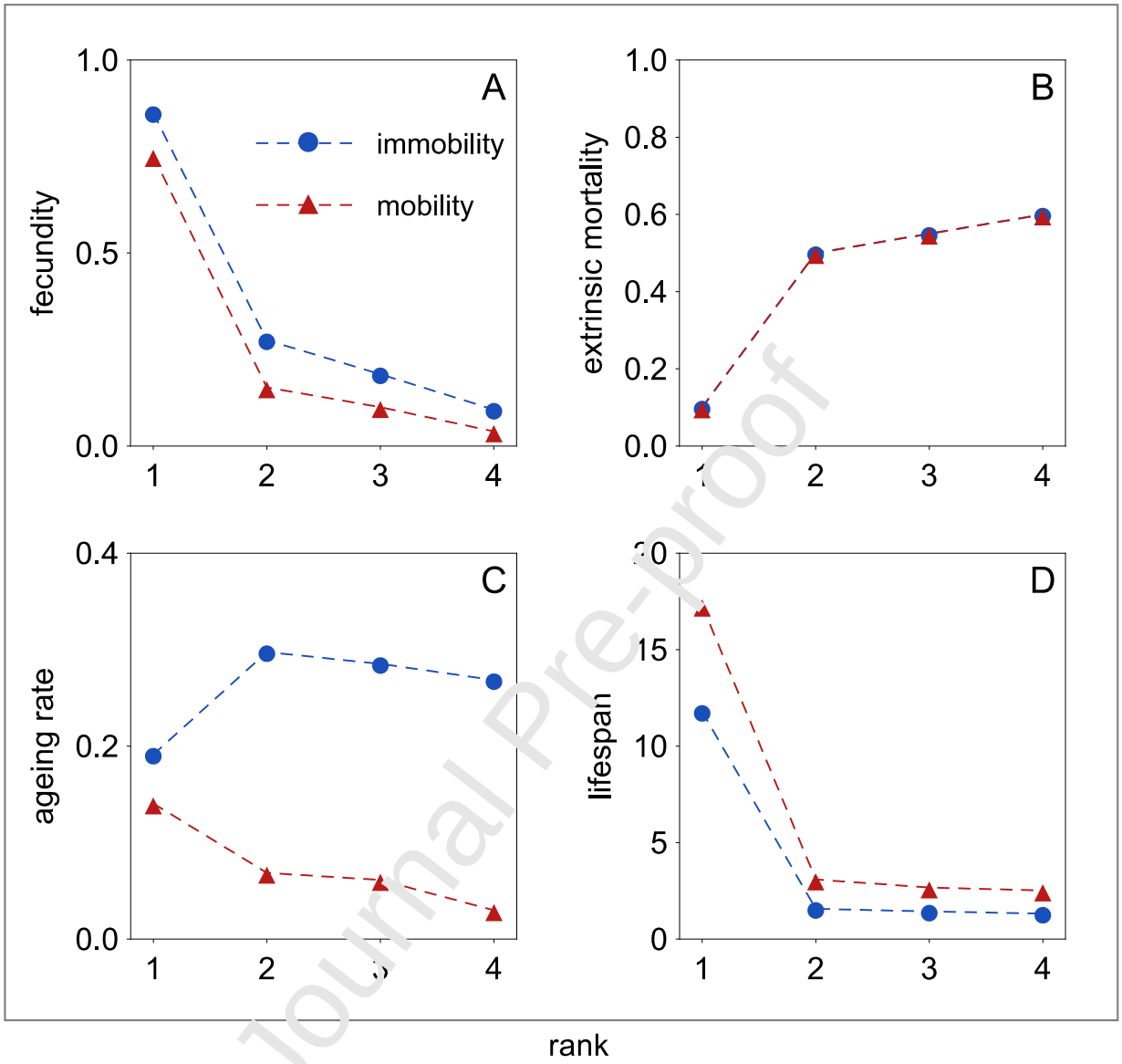
Fig. 4. Life history traits as a function of social rank for a social immobility and social mobility scenarios. Females at the top of the social hierarchy (rank-1 mothers) always produce more offspring (panel A) and have longer lifespans (panel D) than lower-rank females, irrespective of the degree of social mobility. Under social mobility, however, dominant females show higher ageing rates

(panel C) but longer lifespans. Despite their higher ageing rates, dominant females have longer lifespans because they experience lower extrinsic mortality rates (panel B). Parameter values: f_0 = {1.4, 0.3, 0.2, 0.1}, s_0 = {0.9, 0.5, 0.45, 0.4}, k = 0.5, d = 0.1.









sdsd