Evolution of brain ontogenetic growth under ecological challenges

Mauricio González-Forero[1*], Timm Faulwasser[2†], and Laurent Lehmann[1‡]

[1]Department of Ecology and Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland

[2]Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, and Institute for Applied Computer Science, Karlsruhe Institute of Technology, 76344 Eggenstein-Leopoldshafen, Germany

[*]Mauricio.GonzalezForero@unil.ch, † Timm.Faulwasser@epfl.ch, ‡ Laurent.Lehmann@unil.ch.

Abstract

Large brains are metabolically expensive but support skills (or cognitive abilities, knowledge, information, etc.) that allow overcoming ecological and social challenges, with social challenges being thought to strongly promote large-brain evolution by causing evolutionary arms races in cognition yielding exaggerated brain sizes. We formulate a mathematical model that yields quantitative predictions of brain and body mass throughout ontogeny when individuals evolve facing ecological but no social challenges. We find that ecological challenges alone can generate adult brain and body mass of ancient human scale, showing that evolutionary arms races in cognition are not necessary for extreme brain sizes. We show that large brains are favored by intermediately challenging ecological environments where skills are moderately effective and metabolically expensive for the brain to maintain. We further show that observed correlations of cognitive abilities and brain mass can result from saturation with skill maintenance of the brain metabolic rate allocated to skills.

Introduction

Large brains use copious amounts of resources that could otherwise be allocated to reproductive function (Aiello and Wheeler, 1995, Mink *et al.*, 1981, Kuzawa *et al.*, 2014). Large brains can, however, support complex cognition giving rise to abilities such as creativity, expertise, intelligence, language, consciousness, self control, and predicting the thoughts of others (theory of mind) (Sternberg and Ben-Zeev, 2001, Shettleworth, 2010, MacLean *et al.*, 2014, Heyes and Frith, 2014). Large brains may thus entail sizable benefits in reproductive success by providing the individual with skills (or cognitive abilities, knowledge, information, etc.) to overcome ecological and social challenges. For instance, brain-generated skills may allow overcoming ecological challenges such as obtaining nutritionally rich but relatively inaccessible food (Seyfarth and Cheney, 2002, Milton, 1981, Clutton-Brock and Harvey, 1980, Barton, 1999, Kaplan *et al.*, 2000, Kaplan and Robson, 2002, Wrangham, 2009). Additionally, brain-generated skills may allow overcoming social challenges such as coordinating with or out-competing social partners, for example to hunt big game or ascend the social hierarchy (Humphrey, 1976, Byrne and Whiten, 1988, de Waal, 1998, Dunbar and Shultz, 2007). An important aspect of social challenges is that they can involve conflicts of interest among social partners, which may promote evolutionary arms races in cognition, possibly leading to exaggerated brain sizes (Humphrey, 1976, Byrne and Whiten, 1988, de Waal, 1998, Dunbar and Shultz, 2007, McNally and Jackson, 2013, Arbilly *et al.*, 2014). Yet, regardless of the selective forces for large brains, the energy needed to support them must be available in order to meet their substantial energetic demands (Aiello and Wheeler, 1995, Isler and van Schaik, 2006).

Ecological and social challenge hypotheses are often assessed by means of correlations between ecological or social variables with measurements of cognitive abilities or proxies thereof (Clutton-Brock and Harvey, 1980, Barton, 1999, Dunbar and Shultz, 2007, Dunbar, 1998, Fish and Lockwood, 2003, Taylor and van Schaik, 2007, MacLean *et al.*, 2009, Allen and Kay, 2012, MacLean *et al.*, 2014, Shettleworth, 2010, MacLean *et al.*, 2013, Benson-Amram *et al.*, 2016). For instance, in primates, diet breath correlates with self control (MacLean *et al.*, 2014) and group size correlates with neocortex ratio (Dunbar and Shultz, 2007); yet, diet quality has failed to correlate with endocranial volume in New World monkeys (Allen and Kay, 2012) and group size has failed to correlate with problem-solving ability in mammalian carnivores (Benson-Amram *et al.*, 2016). Ecological and social challenge hypotheses have also been evaluated with functional studies. For example, in humans, behavioral experiments have found refined cognitive skills for social rather than general (ecological) function (Herrmann *et al.*, 2007, Cosmides *et al.*, 2010), and brain imaging has identified various brain regions specialized for social interaction (Amodio and Frith, 2006, Frith, 2007). Recently, studies have more directly addressed the causes for large-brain evolution via phylogenetic analyses, artificial selection

experiments, and genomic patterns of selection (Pérez-Barbería *et al.*, 2007, Finarelli and Flynn, 2009, Babbitt *et al.*, 2010, Kotrschal *et al.*, 2013, Mathieson *et al.*, 2015). However, there is a need of testable mathematical theory guiding causal understanding of the relative contribution of ecological and social challenges to large-brain evolution (Healy and Rowe, 2007, Jones, 2015).

Here we study the possible causal contribution of ecological challenges alone to large-brain evolution by means of a mathematical model. We formulate a metabolically explicit model for the evolution of brain ontogenetic growth when individuals face ecological but no social challenges. We use the model to determine how much energy should be allocated to brain growth at each age as a result of natural selection given that overcoming ecological challenges provides energetic returns (e.g., through food procurement). By excluding social challenges, the model deliberately eliminates the possibilities of evolutionary arms races in allocation to brain growth, and thus serves as a baseline for understanding brain growth evolution. We derive the model in terms of measurable parameters using the approach of West *et al.* (2001). In particular, the model incorporates parameters measuring the mass-specific metabolic costs of brain growth and maintenance, which capture the relatively large metabolic expense of the brain. These parameters can be measured empirically, and are likely to differ among species given different brain structures and efficiencies. Once parameterized with values obtained from data, the model yields quantitative predictions for brain and body mass throughout ontogeny under the assumption that individuals evolved under ecological challenges alone.

A defining feature of the model is that it assumes that some of the brain's energetic consumption is due to acquisition and maintenance of skills (or cognitive abilities, knowledge, information, etc.). In particular, we focus on skills that allow extracting energy from the environment (Schniter *et al.*, 2015). Our approach builds on previous models considering brain (physical embodied capital) and skill (functional embodied capital) as part of the individual's embodied capital invested in fitness (Kaplan and Robson, 2002). It also accounts for the notion that information gained and maintained by the brain during ontogeny should be explicitly considered when attempting to understand brain evolution (Boyd and Richerson, 1985, Shettleworth, 2010, van Schaik and Burkart, 2011). Then, given that the brain consumes energy to gain and maintain skills, our model allows to predict how much an individual should grow its brain to obtain the energetic returns from skills. By feeding the model with parameter values for modern humans (i.e., *Homo sapiens*), we find that the model can correctly predict various major modern human life history stages as well as adult body and brain mass of ancient human scale (i.e., of late *Homo erectus* and Neanderthals). These findings show that ecological challenges alone can generate extreme brain sizes despite the absence of evolutionary arms races in cognition.

Model description

We consider a population with overlapping generations and measure the age of individuals in continuous time (Charlesworth, 1980). To keep the model tractable, we assume random mating, constant population size, and constant environment. We focus on females throughout and assume that at each age, each individual faces ecological challenges of energy extraction from the non-social environment (e.g., cracking a nut, hunting-gathering, or lighting a fire to cook). We assume that some of the energetic consumption of the brain is due to acquisition (learning) and maintenance (memory) of energy-extraction skills. We assume that each individual uses the energy-extraction skills it has at a given age to extract energy. In accordance with our aim of building a model without social challenges, we further assume that energy extraction is done individually, but is facilitated early in life by parental or alloparental care. For simplicity, we assume that (allo)parental care has fertility but no survival costs. The individual can use the energy extracted in growth, maintenance, and reproduction. We define growth metabolic rate as the heat released due to body growth by the resting individual per unit time at each age. We further define a tissue's growth schedule as the fraction of growth metabolic rate due to the growth of that tissue. We let growth schedules be evolvable traits, and by making further standard life history assumptions (Mylius and Diekmann, 1995, Dieckmann *et al.*, 2006), we identify evolutionarily stable growth schedules (ESGS) for each tissue by using optimal control theory.

Energy and mass

We partition the mass of an individual into three types of tissues: brain tissue, reproductive tissue, and the remainder which we refer to as somatic tissue. The mass of tissue i of a representative individual at age a is $x_i(a)$, and we use $i \in \{\mathrm{b},\mathrm{r},\mathrm{s}\}$ for brain, reproductive, and somatic tissue, respectively. The resting metabolic rate of the individual at age a is $B_{\mathrm{rest}}(a)$, which is the heat released by the resting individual per unit time at age a. An average mass unit of tissue i of the resting individual releases an amount of heat B_i per unit time, which for simplicity we assume constant with respect to age. Hence, the maintenance metabolic rate at age a, which is the heat released by the resting individual per unit time for maintaining its existing mass, is $B_{\mathrm{maint}}(a) = \sum_{i\in\{\mathrm{b},\mathrm{r},\mathrm{s}\}} x_i(a)B_i$. Then, the growth metabolic rate is $B_{\mathrm{rest}}(a) - B_{\mathrm{maint}}(a)$, which gives the amount of heat released by the resting individual per unit time for producing new tissue. The fraction of the growth metabolic rate allocated to tissue i at age a is the tissue's growth schedule $u_i(a)$. We ask what is the growth schedule of each tissue at each age as a result of natural selection, so we take the growth schedules $u_i(a)$ of the three tissues as the evolvable traits.

Producing an average mass unit of tissue i releases as heat an amount of energy E_i, which for

simplicity we also assume constant with respect to age. Building on the metabolic model of West *et al.* (2001), we show in the Supporting Information (SI, §1.1–1.3) that the growth rate of the mass of tissue $i \in \{\mathrm{b}, \mathrm{r}, \mathrm{s}\}$ is

$$\dot{x}_i(a) = u_i(a)\left(\frac{B_{\mathrm{rest}}(a) - B_{\mathrm{maint}}(a)}{E_i}\right), \tag{1}$$

where $\dot{x}_i(a)$ denotes the derivative of $x_i(a)$ with respect to age. Equation (1) is a general equation describing how growth schedules $u_i(a)$ specify tissue growth rates.

Skill

We let the individual have a number $x_{\mathrm{k}}(a)$ of energy-extraction skills at age a. We assume that a fraction ν_{k} of the brain metabolic rate is due to the energetic expense incurred by the brain for acquiring (learning) and maintaining (memory) energy-extraction skills. We also assume that the brain releases as heat an amount of energy E_{k} for gaining an average skill (learning cost) and an amount B_{k} per unit time for maintaining an average skill (memory cost). We also assume ν_{k}, E_{k}, and B_{k} to be constant. The growth rate of energy-extraction skills (see SI §1.4 for derivation) is then

$$\dot{x}_{\mathrm{k}}(a) = \frac{\nu_{\mathrm{k}} M_{\mathrm{brain}}(a) - x_{\mathrm{k}}(a) B_{\mathrm{k}}}{E_{\mathrm{k}}}, \tag{2}$$

where

$$M_{\mathrm{brain}}(a) = x_{\mathrm{b}}(a) B_{\mathrm{b}} + \dot{x}_{\mathrm{b}}(a) E_{\mathrm{b}} \tag{3}$$

is the brain metabolic rate at age a (i.e., the energy released as heat by the brain per unit time with the individual at rest) which consists of the heat released for brain tissue maintenance $[x_{\mathrm{b}}(a) B_{\mathrm{b}}]$ and growth $[\dot{x}_{\mathrm{b}}(a) E_{\mathrm{b}}]$. Equation (2) is also a general equation capturing the link of brain with skill; it is general in that, for example, (2) is not restricted to energy-extraction skills (given that ν_{k} is accordingly reinterpreted). In analogy with (1), the first term in the numerator of (2) gives the heat released due to energetic input for skill growth whereas the second term gives the heat released for skill maintenance.

Skill function

Finally, we specify how skills allow for energy extraction. We denote the probability of energy extraction at age a as $p(x_{\mathrm{k}}(a))$, defined as the ratio of the amount of energy extracted per unit time at age a over that extracted if the individual is maximally successful at energy extraction. We assume that $p(x_{\mathrm{k}}(a))$ depends on skill number but is independent of body mass. Given the empirical relationship of resting metabolic rate and body mass as a power law (Kleiber, 1961, Peters, 1983, Sears *et al.*, 2012), which for humans also holds ontogenetically to a good approximation (Fig. S4), we show in the SI

(§1.5) that resting metabolic rate takes the form

$$B_{\text{rest}}(a) = K\, p(x_{\text{k}}(a))\, x_{\text{T}}(a)^{\beta}, \tag{4}$$

where β is a scaling coefficient, K is a constant independent of body mass, and body mass is $x_{\text{T}}(a) = \sum_{i\in\{\text{b,r,s}\}} x_i(a)$. Equation (4) captures the notion that energy extraction gives the individual energy that it can use to grow or maintain its different tissues.

We consider energy extraction at age a as a contest against the environment. We thus let the probability of energy extraction $p(x_{\text{k}}(a))$ take the form of a contest success function (Hirshleifer, 1995, Skaperdas, 1996):

$$p(x_{\text{k}}(a)) = \frac{c(x_{\text{k}}(a))}{d(a) + c(x_{\text{k}}(a))}, \tag{5}$$

which we assume increases with the number $x_{\text{k}}(a)$ of energy-extraction skills, and depends on two terms. First, the probability of energy extraction depends on the difficulty of energy-extraction at age a, measured by $d(a)$. Thus, the higher $d(a)$, the more challenging energy extraction is and the more energy-extraction skills the individual must have to obtain resources. We let $d(a) = \alpha - \varphi(a)$, where α is the environmental difficulty and $\varphi(a)$ is the facilitation of energy-extraction due to (allo)parental care. We let this facilitation be an exponentially decreasing function with age, $\varphi(a) = \varphi_0 \mathrm{e}^{-\varphi_{\text{r}} a}$, and we ignore the increased resting metabolic rate caused by gestation and lactation (Pontzer, 2015).

Second, the probability of energy extraction depends on the individual's competence at energy extraction, denoted by $c(x_{\text{k}}(a))$. We consider two cases that are standard in contest models: (1) a power function $c_1(x_{\text{k}}(a)) = (x_{\text{k}}(a))^{\gamma}$, so the probability of energy extraction $p(x_{\text{k}}(a))$ is a contest success function in ratio form (power competence); and (2) an exponential function $c_2(x_{\text{k}}(a)) = \left(\mathrm{e}^{x_{\text{k}}(a)}\right)^{\gamma}$ so the probability of energy extraction is in difference form (exponential competence) (Hirshleifer, 1995, Skaperdas, 1996). In both cases, the parameter γ describes the effectiveness of skills at energy extraction. Thus, with $\gamma = 0$, skills are ineffective while with increasing γ fewer skills are needed to extract energy. In general, competence $c(x_{\text{k}}(a))$ represents features of the individual (e.g., how increasing skill changes efficiency in information processing by the brain), and of the environment (e.g., how adding the skill of caching nuts to that of cracking nuts changes energy extraction efficiency). For a given skill effectiveness (γ), exponential competence assumes a steeper increase in competence with increasing skill number than power competence.

Evolutionary invasion analysis

Under standard life history assumptions, if an evolutionary equilibrium is reached, natural selection maintains the population at this equilibrium where evolutionarily stable growth schedules (ESGS) maximize the individual's lifetime number of offspring assuming that population density is regulated through fertility (Mylius and Diekmann, 1995, Dieckmann *et al.*, 2006, see also Lande, 1982).

Using equations (1)–(5), we seek the ESGS, denoted by $u_i^*(a)$, which yield the optimal tissue mass denoted by $x_i^*(a)$. With this aim, we obtain the ESGS by solving this maximization problem, for which we employ optimal control theory (SI §2-4) and the software GPOPS (Patterson and Rao, 2014) for numerical solutions. For simplicity, we assume that mortality is constant, and so the brain only affects fitness through fertility. Assuming that part of the heat released by reproductive tissue is due to offspring cell production and that (allo)parental care only entails fertility costs, we let fertility be proportional to the mass of reproductive tissue (see SI §1.6 and Chang *et al.*, 1998).

The model depends on 21 parameters that affect the ESGS, and they measure (P1) tissue mass in the newborn, (P2) tissue metabolism (i.e., metabolic costs of tissue maintenance and growth), (P3) demography, (P4) skill of the newborn, (P5) skill metabolism (i.e., metabolic costs of memory and learning), (P6) (allo)parental care, and (P7) contest success (SI §5). We use published data for modern human females to estimate 13 parameters that affect the ESGS that are readily estimated from available data (P1-P3) (SI §5,6; Table S2). These parameters include the brain and body metabolic costs, and with these parameters fixed, the model can only generate a vastly narrower set of outcomes. The remaining 8 parameters (P4-P7) are not readily estimated from available data, so for them we identify by trial-and-error benchmark values that yield a model output in agreement with observed body and brain mass data for modern human females. The benchmark values are thus different with power (Table S3) and exponential (Table S4) competence. We first present the numerical results for the two sets of benchmark parameter values and then the results when these benchmark parameter values vary (see SI for further details and computer code).

Results

Predicted life history stages: childhood, adolescence, and adulthood

The resulting ESGS divide the individual's lifespan in three broad stages: (1) a "childhood" stage, defined as the stage lasting from birth to a_{m} years of age and during which allocation to growth of reproductive tissue is zero; (2) an "adolescence" stage, defined as the stage lasting from a_{m} to a_{a} years of age and during which there is simultaneous allocation to growth of somatic and reproductive tissue; and (3) an "adulthood" stage, defined as the stage lasting from a_{a} to the end of the individual's reproductive career and during which all growth allocation is to reproductive tissue (Fig. 1a). These life stages are obtained with either power or exponential competence (Fig. 1a,e). Note that the ages at "menarche" a_{m} and adulthood a_{a} are not parameters but an output of the model.

The obtained childhood stage, which is the only stage where there is brain growth, is further subdivided in three periods: (1a) "early childhood", defined here as the earliest childhood period with

pure allocation to somatic growth; (1b) "mid childhood", defined here as the childhood period where there is simultaneous allocation to somatic and brain growth; and (1c) "preadolescence", defined here as the latest childhood period of pure somatic growth. Hence, brain growth occurs exclusively during "mid childhood". This result disagrees with observation as the obtained absence of allocation to brain growth during early childhood does not occur in humans. This discrepancy may be an inaccuracy arising because the approximation of resting metabolic rate by a power law of body mass which we use in the model (West *et al.*, 2001) underestimates resting metabolic rate, and thus growth metabolic rate, during early childhood (Fig. S4). The period we refer to here as mid childhood then lasts from the obtained age a_{b0} of brain growth onset to the obtained age a_b of brain growth arrest (Fig. 1a).

With the exception of the age of brain growth onset, the predicted timing of childhood, adolescence, and adulthood closely follows that observed in modern humans with competence being either a power or an exponential function of skill number (Table 1). Note that measurement units (i.e., years, kg, and MJ), excepting skill units, are not arbitrary as they result from the units of the parameter values estimated from empirical data (Table S2). Hence, the model correctly predicts major stages of human life history with accurate timing, with the exception of brain growth allocation during early childhood (Table 1).

Body and brain mass through ontogeny

The ESGS generate the following predicted body and brain mass throughout ontogeny. For total body mass, there is fast growth during early childhood, followed by slow growth during mid childhood, a growth spurt during preadolescence, slow growth during adolescence, and no growth during adulthood, all closely following the observed pattern (Fig. 1b). The slow growth during mid childhood results from the simultaneous allocation to somatic and brain growth and from the decreasing growth metabolic rate due to the increasing energetic costs of brain maintenance (Fig. 1c). The growth spurt during adolescence arises because (1) all growth metabolic rate is allocated to inexpensive somatic growth, and (2) growth metabolic rate increases due to increased metabolic rate caused by increasing, inexpensive-to-maintain somatic mass (Fig. 1c). The slow growth during adolescence is due to simultaneous somatic and reproductive growth, and to the elevated costs of reproductive tissue maintenance (Fig. 1c). These growth patterns result in two major peaks in growth metabolic rate (Fig. 1c). While the first peak in growth metabolic rate is made possible by (allo)parental care, the second peak is made possible by the individual's own skills (Fig. S7d). After the onset of adulthood at a_a, growth metabolic rate is virtually depleted and allocation to growth has essentially no effect on tissue growth (Fig. 1c).

Whereas predicted body growth patterns are qualitatively similar with either power or exponential competence, they differ quantitatively (Fig. 1b,f). With power competence, the predicted body mass is quantitatively nearly identical to that observed in modern humans throughout life (Fig. 1b). In contrast, with exponential competence, the predicted body mass is larger throughout life than that of modern human females (Fig. 1f).

Regarding brain mass, the model predicts it to have the following growth pattern. During early childhood, brain mass remains static, in contrast to the observed pattern (Fig. 1d). During mid childhood, brain mass initially grows quickly, then it slows down slightly, and finally grows quickly again before brain growth arrest at the onset of preadolescence (Fig. 1d). Predicted brain growth is thus delayed by the obtained early-childhood period relative to the observed brain growth in modern humans (Fig. 1d). As previously stated, this delay in predicted brain growth may be an inaccuracy arising from the underestimation of resting metabolic rate during early childhood by the power law of body mass.

Predicted brain growth patterns are also qualitatively similar but quantitatively different with power and exponential competence (Fig. 1d,h). Adult brain mass is predicted to be smaller or larger than that observed in modern human females depending on whether competence is respectively a power or an exponential function (Fig. 1d,h). Remarkably, considering body and brain mass together, the predicted adult body and brain mass can match those observed in late *H. erectus* if competence is a power function (Fig. 1b,d). In contrast, the predicted adult body and brain mass can match those of Neanderthals if competence is an exponential function (Fig. 1f,h). Consequently, the encephalization quotient (EQ, which is the ratio of observed adult brain mass over expected adult brain mass for a given body mass) is larger with exponential competence for the parameter values used (Table 1).

Skills through ontogeny

The obtained ESGS predict the following patterns for energy-extraction skills throughout ontogeny. For the scenario in Fig. 1, the individual gains most skills during childhood and adolescence, skill number continues to increase after brain growth arrest, and skill number plateaus in adulthood (Fig. 2). That is, skill growth is determinate, in agreement with empirical observations (Fig. 2). Yet, if memory cost B_k is substantially lower, skill number can continue to increase throughout the individual's reproductive career (i.e., skill growth is then indeterminate; Fig. S8e) [see equation (2)]. Nevertheless, in that case, the agreement between predicted and observed body and brain mass throughout ontogeny is substantially reduced (Fig. S8b,c).

When skill growth is determinate, the model predicts adult skill number to be proportional to adult brain mass. In particular, with determinate skill growth, the number of skills that is asymptoti-

cally achieved [from equation (2) setting $\dot{x}_k(a) = 0$ and $u_b^*(a) = 0$] is

$$\hat{x}_k = \upsilon_k \frac{B_b}{B_k} x_b^*(a_a), \tag{6}$$

where $\hat{x}_k$ is the asymptotic skill number, $x_b^*(a_a)$ is the adult brain mass, υ_k is the fraction of brain metabolic rate allocated to energy-extraction skills, and B_b is the brain mass-specific maintenance cost. The requirement for skill growth to be determinate is that the brain metabolic rate allocated to skills [$\upsilon_k M_{brain}(a)$] becomes saturated with skill maintenance [$x_k(a)B_k$] within the individual's reproductive career [equation (2)]. Hence, adult skill number is proportional to adult brain mass in the model because of saturation with skill maintenance of the brain metabolic rate allocated to skills and because adult brain metabolic rate is found to be proportional to adult brain mass [setting $\dot{x}_b(a_a) = 0$ in equation (3) yields $M_{brain}(a_a) = x_b(a_a)B_b$]. Weak correlations between cognitive ability and brain mass have been identified across taxa including humans (Andreasen *et al.*, 1993, Deaner *et al.*, 2007, MacLean *et al.*, 2014, Pietschniga *et al.*, 2015, Benson-Amram *et al.*, 2016). Since skills are here broadly understood to include cognitive abilities (provided parameters are suitably reinterpreted), this result provides an explanation for these correlations in terms of saturation of brain metabolic rate with skill (cognitive ability) maintenance.

We now vary parameter values to assess what factors favor a large brain at adulthood.

A large brain is favored by intermediate environmental difficulty, moderate skill effectiveness, and costly memory

A larger adult brain mass is favored by an increasingly challenging environment [increasing α; equation (5)], but is *disfavored* by an exceedingly challenging environment (Fig. 3a). Environmental difficulty favors a larger brain because more skills are needed for energy extraction [equation (5)], and from equation (2) more skills can be gained by increasing brain metabolic rate in turn by increasing brain mass. Thus, a large brain is favored to energetically support skill growth in a challenging environment. However, with exceedingly challenging environments, the individual is favored to reproduce early without substantial body or brain growth because it fails to gain enough skills to maintain its body mass as (allo)parental care decreases with age (Fig. S12).

A larger adult brain is favored by moderately effective skills. When skills are ineffective at energy extraction [$\gamma \to 0$; equation (5)], the brain entails little fitness benefit and fails to grow in which case the individual also reproduces without substantially growing (Fig. 3b). When skill effectiveness (γ) crosses a threshold value, the fitness effect of brain becomes large enough that the brain becomes favored to grow. Yet, as skill effectiveness increases further and thus fewer skills are needed for energy extraction, a smaller brain supports enough skill growth, so the optimal adult brain mass *decreases* with skill effectiveness (Fig. 3b). Hence, adult brain mass is largest with moderately effective skills.

A larger brain is also favored by skills that are increasingly expensive for the brain to maintain (costly memory, increasing B_k), but exceedingly costly memory prevents body and brain growth (Fig. 3c). Costly memory favors a large brain because then a larger brain mass is required to energetically support skill growth [equation (2)]. If memory is exceedingly costly, skills fail to grow and energy extraction is unsuccessful, causing the individual to reproduce without substantial growth (Fig. 3c).

Factors favoring a large EQ and high skill

A large EQ and high adult skill number are generally favored by the same factors that favor a large adult brain. However, the memory cost has a particularly strong effect favoring a large EQ because it simultaneously favors increased brain and reduced body mass (Fig. 3c,f). In contrast to its effect on EQ, increasing memory cost *disfavors* a high adult skill number (Fig. 3f). That is, a higher EQ attained by increasing memory costs is accompained by a *decrease* in skill number (Fig. 3c,f). The factors that favor a large brain, large EQ, and high skill are similar with either power or exponential competence (Fig. 3 and Figs. S15,S16). Importantly, although with the estimated parameter values the model can recover modern human growth patterns yielding adult body and brain mass of ancient humans, our exploration of the parameters that were not estimated from data suggests that the model cannot recover modern human growth patterns yielding adult body and brain mass of *modern* humans.

Discussion

Our model shows that ecological challenges alone can be sufficient, and that evolutionary arms races in cognition are not necessary, to generate major human life history stages as well as adult brain and body mass of ancient human scale. We find that the brain is favored to grow to energetically support skill growth, and thus a large brain is favored when simultaneously (1) competence at energy extraction has a steep dependence on skill number, (2) many skills are needed for energy extraction due to environmental difficulty and moderate skill effectiveness, and (3) skills are expensive for the brain to maintain but are still necessary for energy extraction.

While the model considers ecological challenges alone and so evolutionary arms races in cognition do not take place, the model can recover body and brain mass of ancient human scale. Predicted encephalization can match that of late *H. erectus* with competence being a power function of skills, and that of Neanderthals with competence as an exponential function. These results call for empirical assessment of the probability of energy extraction versus skill number (or cognitive ability, knowledge, etc.) to allow for increasingly accurate predictions (Jia *et al.*, 2013). Similarly, use of parameter values for non-human taxa would allow to determine the model's ability to predict diverse

life histories and brain growth patterns (Moses *et al.*, 2008), offering a means to assess the explanatory potential of ecological challenges for large-brain evolution across taxa.

The model also provides an explanation for observed inter- and intraspecific correlations between adult cognitive ability and brain mass across taxa including humans (Andreasen *et al.*, 1993, Deaner *et al.*, 2007, MacLean *et al.*, 2014, Pietschniga *et al.*, 2015, Benson-Amram *et al.*, 2016). The explanation is the saturation with skill maintenance of the brain metabolic rate allocated to skills during the individual's lifespan [equation (6)]. The proportionality arises because the adult brain metabolic rate is found to be proportional to brain mass. This explanation follows from a general equation for the learning rate of skills [equation (2)] that is based on metabolic considerations (West *et al.*, 2001) without making assumptions about skill function; yet, this equation assumes that the fraction of brain metabolic rate allocated to the skills of interest (υ_k) is independent of brain mass (and similarly for B_b and B_k). The model further predicts that additional variation in correlations between cognitive ability and brain mass can be explained by variation in maintenance costs of brain and skill, and by variation in brain metabolic rate allocation to skill [equation (6)]. However, the model indicates that adult skill number and brain mass need not be correlated since saturation with skill maintenance of the brain metabolic rate allocated to skills may not occur during the individual's lifespan, for example if memory is inexpensive, so skill number increases throughout life (Fig. S8e).

Predicted adult brain mass and skill have non-monotonic relationships with their predictor variables (Fig. 3 and Figs. S15,S16). Consequently, conflicting inferences can be drawn if predictor variables are evaluated only on their low or high ends. For instance, increasingly challenging environments favor large brains up to a point, so that exceedingly challenging environments disfavor large brains. Thus, on the low end of environmental difficulty, the prediction that increasingly challenging environments favor large brains is consistent with ecological challenge hypotheses (Kaplan and Robson, 2002, Kaplan *et al.*, 2000); yet, on the high end of environmental difficulty, the prediction that increasingly challenging environments disfavor large brains is consistent with constraint hypotheses according to which facilitation of environmental challenge favors larger brains (Austad and Fischer, 1994, Kaplan and Robson, 2002, Hintze *et al.*, 2015). Counter-intuitively on first encounter, the finding that moderately effective skills are most conducive to a large brain and high skill is simply a consequence of the need of more skills when their effectiveness decreases (Fig. 3b). Regarding memory cost, the strong effect of memory cost on favoring a high EQ at first glance suggests that a larger EQ than the observed in modern humans is possible if memory were costlier (see dashed lines in Fig. 3e). However, such larger memory costs cause a substantial delay in body and brain growth, and the resulting growth patterns are inconsistent with those of modern humans (Figs. S9–S11).

Although our model does not include numerous details relevant to humans including social chal-

lenges and social learning, our results are relevant for a set of hypotheses for human-brain evolution. In particular, food processing (e.g., mechanically with stone tools or by cooking) has previously been advanced as a determinant factor in human-brain evolution as it increases energy and nutrient availability from otherwise relatively inaccessible sources (Wrangham, 2009, Zink and Lieberman, 2016). Evidence of human fire control has been inconclusive for early dates (1.5 mya, associated with early *H. erectus* in South Africa), while being more secure for more recent dates (800 kya, associated with *H. erectus* in Israel) and abundant for yet more recent times (130 kya, associated with Neanderthals and *H. sapiens* throughout the Old World) (Klein, 2009). Unambiguous evidence of fire deep inside a South African cave associated to *H. erectus* has been identified for sediments dated to 1 mya (Berna *et al.*, 2012). Regarding mechanical processing, "many of the oldest stone tools bear traces of being used to slice meat" (1.5 mya in Kenya; Zink and Lieberman, 2016, Keeley and Toth, 1981) and experimental evidence shows that meat slicing and vegetable pounding substantially reduce chewing effort (Zink and Lieberman, 2016). Food processing relates to our results not only in that it constitutes an ecological rather than a social challenge, but also in that it may help satisfy at least two of the three key conditions identified for large-brain evolution listed in the first paragraph of the Discussion. First, a shift in food-processing technology (e.g., from primarily mechanical to cooking) may create a steeper relationship between energy-extraction skills and competence by substantially facilitating energy extraction (relating to condition 1). Second, food processing (e.g., by building the required tools or lighting a fire) is a challenging feat to learn and may often fail (relating to condition 2). Yet, there are scant data allowing to judge the metabolic expense for the brain to maintain tool-making or fire-control skills (condition 3). Our results thus indicate that food processing may well have been a key causal factor in human brain expansion. Also, although we did not consider social aspects in our model, the steepness of competence with respect to skill may increase with social learning as well. Social learning can facilitate the acquisition of adaptive skills (Boyd and Richerson, 1985, van Schaik and Burkart, 2011), and skills increasing the steepness of competence with respect to skill could be particularly adaptive. In this case, sociality could favor high encephalization in the absence of cognitive arms races (van Schaik and Burkart, 2011).

Despite considering ecological challenges alone and additional simplifying assumptions, our model accurately predicts major stages of human life history while simultaneously recovering adult brain and body mass of ancient human scale. The model identifies various ecological drivers of large-brain evolution, in particular steep competence with respect to skill, intermediate environmental difficulty, moderate skill effectiveness, and costly memory. As we did not consider social challenges, our model cannot refute or support social challenge hypotheses. However, our results show that when the various factors favoring large brains co-occur, ecological challenges alone can be sufficient

410 to explain major aspects of human life history and large-brain evolution.

411 Acknowledgments

412 We thank Tadeusz J. Kawecki for helpful discussion. This work was funded by Swiss NSF grant PP00P3-
413 146340 to LL.

414 References

415 Aiello, L.C. and Wheeler, P. (1995). The expensive-tissue hypothesis: the brain and the digestive sys-
416 tem in human and primate evolution. *Curr. Anthropol.*, **36**, 199–221.

417 Allen, K.L. and Kay, R.F. (2012). Dietary quality and encephalization in platyrrhine primates. *Proc. R.*
418 *Soc. B*, **279**, 715–721.

419 Amodio, D.M. and Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cogni-
420 tion. *Nat. Rev. Neurosci.*, **7**, 268–277.

421 Andreasen, N.C., Flaum, M., Swayze II, V., O'Leary, D.S., Alliger, R., Cohen, G. et al (1993). Intelligence
422 and brain structure in normal individuals. *Am. J. Psychiatry*, **150**, 130–134.

423 Arbilly, M., Weissman, D.B., Feldman, M.W. and Grodzinskic, U. (2014). An arms race between pro-
424 ducers and scroungers can drive the evolution of social cognition. *Behav. Ecol.*, **25**, 487–495.

425 Austad, S.N. and Fischer, K.E. (1994). Primate longevity: Its place in the mammalian scheme. *Am. J.*
426 *Primatol.*, **28**, 251–261.

427 Babbitt, C.C., Warner, L.R., Fedrigo, O., Wall, C.E. and Wray, G.A. (2010). Genomic signatures of diet-
428 related shifts during human origins. *Proc. R. Soc. B*, **278**, 961–969.

429 Barton, R.A. (1999). The evolutionary ecology of the primate brain. In P. C. Lee, editor, *Comparative*
430 *Primate Socioecology*, pages 167–203. Cambridge Unilv. Press.

431 Benson-Amram, S., Dantzer, B., Stricker, G., Swanson, E.M. and Holekamp, K.E. (2016). Brain
432 size predicts problem-solving ability in mammalian carnivores. *Proc. Natl. Acad. Sci. USA*,
433 **doi:10.1073/pnas.1505913113**.

434 Berna, F., Goldberg, P., Horwitz, L.K., Brink, J., Holt, S., Bamford, M. et al (2012). Microstratigraphic
435 evidence of in situ fire in the Acheulean strata of Wonderwerk Cave, Northern Cape province, South
436 Africa. *Proc. Natl. Acad. Sci. USA*, **109**, E1215–E1220.

437 Boyd, R. and Richerson, P.J. (1985). *Culture and the Evolutionary Process*. Univ. Chicago Press.

438 Byrne, R. and Whiten, A., editors (1988). *Machiavellian Intelligence*. Oxford Univ. Press.

439 Chang, M.Y., Chiang, C.H., Hsieh, T.T., Soong, Y.K. and Hsu, K.H. (1998). Use of the antral follicle
440 count to predict the outcome of assisted reproductive technologies. *Fertil. Steril.*, **69**, 505–510.

441 Charlesworth, B. (1980). *Evolution in age-structured populations*. Cambridge Univ. Press.

442 Clutton-Brock, T.H. and Harvey, P.H. (1980). Primates, brains and ecology. *J. Zool.*, **190**, 309–323.

443 Cosmides, L., Barrett, C. and Tooby, J. (2010). Adaptive specializations, social exchange, and the
444 evolution of human intelligence. *Proc. Natl. Acad. Sci. USA*, **107**, 9007–9014. Supplement 2.

445 de Waal, F. (1998). *Chimpanzee Politics*. Johns Hopkins Univ. Press, 2nd edition.

446 Deaner, R.O., Isler, K., Burkart, J. and van Schaik, C. (2007). Overall brain size, and not encephaliza-
447 tion quotient, best predicts cognitive ability across non-human primates. *Brain Behav. Evol.*, **70**,
448 115–124.

449 Dieckmann, U., Heino, M. and Parvinen, K. (2006). The adaptive dynamics of function-valued traits.
450 *J. Theor. Biol.*, **241**, 370–389.

451 Dunbar, R.I.M. (1998). The social brain hypothesis. *Evol. Anthropol.*, **6**, 178–190.

452 Dunbar, R.I.M. and Shultz, S. (2007). Evolution in the social brain. *Science*, **317**, 1344–1347.

453 Finarelli, J.A. and Flynn, J.J. (2009). Brain-size evolution and sociality in Carnivora. *Proc. Natl. Acad.*
454 *Sci. USA*, **106**, 9345–9349.

455 Fish, J.L. and Lockwood, C.A. (2003). Dietary constraints on encephalization in primates. *Am. J. Phys.*
456 *Anthropol.*, **120**, 171–181.

457 Frith, C.D. (2007). The social brain? *Phil. Trans. R. Soc. B*, **362**, 671–678.

458 Froehle, A.W. and Churchill, S.E. (2009). Energetic competition between Neandertals and anatomi-
459 cally modern humans. *PaleoAnthropology*, pages 96–116.

460 Gluckman, P.D. and Hanson, M.A. (2006). Evolution, development and timing of puberty. *Trends*
461 *Endocrinol. Metab.*, **17**, 7–12.

462 Healy, S.D. and Rowe, C. (2007). A critique of comparative studies of brain size. *Proc. R. Soc. B*, **274**,
463 453–464.

Henry, P.I., Morelli, G.A. and Tronick, E.Z. (2005). Child caretakers of among the Efé foragers in the Ituri forest. In B. S. Hewlett and M. E. Lamb, editors, *Hunter-Gatherer Childhoods*, pages 191–213. Transaction.

Herrmann, E., Call, J., Hernández-Lloreda, M.V., Hare, B. and Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science*, **317**, 1360–1366.

Heyes, C.M. and Frith, C.D. (2014). The cultural evolution of mind reading. *Science*, **344**, 1243091.

Hintze, A., Phillips, N. and Hertwig, R. (2015). The Janus face of Darwinian competition. *Sci. Rep.*, **5**, 13662.

Hirshleifer, J. (1995). Theorizing about conflict. In K. Hartley and T. Sandler, editors, *Handbook of Defense and Economics*, volume 1, chapter 7, pages 165–189. Elsevier.

Humphrey, N.K. (1976). The social function of the intellect. In P. P. G. Bateson and R. A. Hinde, editors, *Growing Points in Ethology*, pages 303–317. Cambridge Univ. Press.

Isler, K. and van Schaik, C.P. (2006). Metabolic costs of brain size evolution. *Biol. Lett.*, **2**, 557–560.

Jia, H., Skaperdas, S. and Vaidya, S. (2013). Contest functions: theoretical foundations and issues in estimation. *Int. J. Ind. Organ.*, **31**, 211–222.

Jones, J.H. (2015). Resource transfers and human life-history evolution. *Annu. Rev. Anthropol.*, **44**, 513–531.

Kaplan, H. and Robson, A.J. (2002). The emergence of humans: the coevolution of intelligence and longevity with intergenerational transfers. *Proc. Natl. Acad. Sci. USA*, **99**, 10221–10226.

Kaplan, H., Hill, K., Lancaster, J. and Hurtado, A.M. (2000). A theory of human life history evolution: diet, intelligence, and longevity. *Evol. Anthropol.*, **9**, 156–185.

Keeley, L.H. and Toth, N. (1981). Microwear polishes on early stone tools from Koobi Fora, Kenya. *Nature*, **293**, 464–465.

Kleiber, M. (1961). *The Fire of Life*. Wiley.

Klein, R.G. (2009). *The Human Career: Human Biological and Cultural Origins*. The Univ. of Chicago Press, 3rd edition.

491 Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I. et al (2013). Artificial
492 selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain.
493 *Curr. Biol.*, **23**, 168–171.

494 Kuzawa, C.W., Chugani, H.T., Grossman, L.I., Lipovich, L., Muzik, O., Hof, P.R. et al (2014). Metabolic
495 costs and evolutionary implications of human brain development. *Proc. Nat. Acad. Sci. USA*, **111**,
496 13010–13015.

497 Lande, R. (1982). A quantitative genetic theory of life history evolution. *Ecology*, **63**, 607–615.

498 MacLean, E.L., Barrickman, N.L., Johnson, E.M. and Wall, C.E. (2009). Sociality, ecology, and relative
499 brain size in lemurs. *J. Hum. Evol.*, **56**, 471–478.

500 MacLean, E.L., Sandel, A.A., Bray, J., Oldenkamp, R.E., Reddy, R.B. and Hare, B.A. (2013). Group size
501 predicts social but not nonsocial cognition in lemurs. *PLOS ONE*, **8**, e66359.

502 MacLean, E.L., Hare, B., Nunn, C.L., Addessi, E., Amici, F., Anderson, R.C. et al (2014). The evolution
503 of self-control. *Proc. Natl. Acad. Sci. USA*, **111**, E2140–E2148.

504 Martin, R.D. (1981). Relative brain size and basal metabolic rate in terrestrial vertebrates. *Nature*,
505 **293**, 57–60.

506 Mathieson, I., Lazaridis, I., Rohland, N., Mallick, S., Patterson, N., Roodenberg, S.A. et al (2015).
507 Genome-wide patterns of selection in 230 ancient Eurasians. *Nature*, **528**, 499–503.

508 McHenry, H.M. (1994). Tempo and mode in human evolution. *Proc. Natl. Acad. Sci. USA*, **91**, 6780–
509 6786.

510 McNally, L. and Jackson, A.L. (2013). Cooperation creates selection for tactical deception. *Proc. R.*
511 *Soc. B*, **280**, 20130699.

512 Milton, K. (1981). Distribution patterns of tropical plant foods as an evolutionary stimulus to primate
513 mental development. *Am. Anthropol.*, **83**, 534–548.

514 Mink, J.W., Blumenschine, R.J. and Adams, D.B. (1981). Ratio of central nervous system to body
515 metabolism in vertebrates: its constancy and functional basis. *Am. J. Physiol.*, **241**, R203–R212.

516 Moses, M.E., Hou, C., Woodruff, W.H., West, G.B., Nekola, J.C., Zuo, W. et al (2008). Revisiting a model
517 of ontogenetic growth: estimating model parameters from theory and data. *Am. Nat.*, **171**, 632–
518 645.

519 Mylius, S.D. and Diekmann, O. (1995). On evolutionarily stable life histories, optimization and the
520 need to be specific about density dependence. *Oikos*, **74**, 218–224.

521 Patterson, M.A. and Rao, A.V. (2014). GPOPS-II: A MATLAB software for solving multiple-phase op-
522 timal control problems using hp-adaptive gaussian quadrature collocation methods and sparse
523 nonlinear programming. *ACM Trans. Math. Softw.*, **41**, 1–37.

524 Pérez-Barbería, F.J., Schultz, S. and Dunbar, R.I.M. (2007). Evidence for coevolution of sociality and
525 relative brain size in three orders of mammals. *Evolution*, **61**, 2811–2821.

526 Peters, R.H. (1983). *The Ecological Implications of Body Size*. Cambridge Univ. Press, Cambridge, UK.

527 Pietschniga, J., Penke, L., Wicherts, J.M., Zeiler, M. and Voracek, M. (2015). Meta-analysis of associ-
528 ations between human brain volume and intelligence differences: How strong are they and what
529 do they mean? *Neurosci. Biobehav. Rev.*, **57**, 411–432.

530 Pontzer, H. (2015). Energy expenditure in humans and other primates: A new synthesis. *Annu. Rev.*
531 *Anthropol.*, **44**, 169–187.

532 Ruff, C.B., Trinkaus, E. and Holliday, T.W. (1997). Body mass and encephalization in Pleistocene
533 *Homo*. *Nature*, **387**, 173–176.

534 Schniter, E., Gurven, M., Kaplan, H.S., Wilcox, N.T. and Hooper, P.L. (2015). Skill ontogeny among
535 Tsimane forager-horticulturalists. *Am. J. Phys. Anthrop.*, **158**, 3–18.

536 Sears, K.E., Kerkhoff, A.J., Messerman, A. and Itagaki, H. (2012). Ontogenetic scaling of metabolism,
537 growth, and assimilation: Testing metabolic scaling theory with *Manduca sexta* larvae. *Physiol.*
538 *Biochem. Zool.*, **85**, 159–173.

539 Seyfarth, R.M. and Cheney, D.L. (2002). What are big brains for? *Proc. Natl. Acad. Sci. USA*, **99**,
540 4141–4142.

541 Shettleworth, S.J. (2010). *Cognition, Evolution, and Behavior*. Oxford Univ. Press, 2nd edition.

542 Skaperdas, S. (1996). Contest success functions. *Econ. Theory*, **7**, 283–290.

543 Sternberg, R.J. and Ben-Zeev, T. (2001). *Complex Cognition*. Oxford Univ. Press.

544 Taylor, A.B. and van Schaik, C.P. (2007). Variation in brain size and ecology in *Pongo*. *J. Hum. Evol.*,
545 **52**, 59–71.

546 van Schaik, C.P. and Burkart, J.M. (2011). Social learning and evolution: the cultural intelligence
547 hypothesis. *Phil. Trans. R. Soc. B*, **366**, 1008–1016.

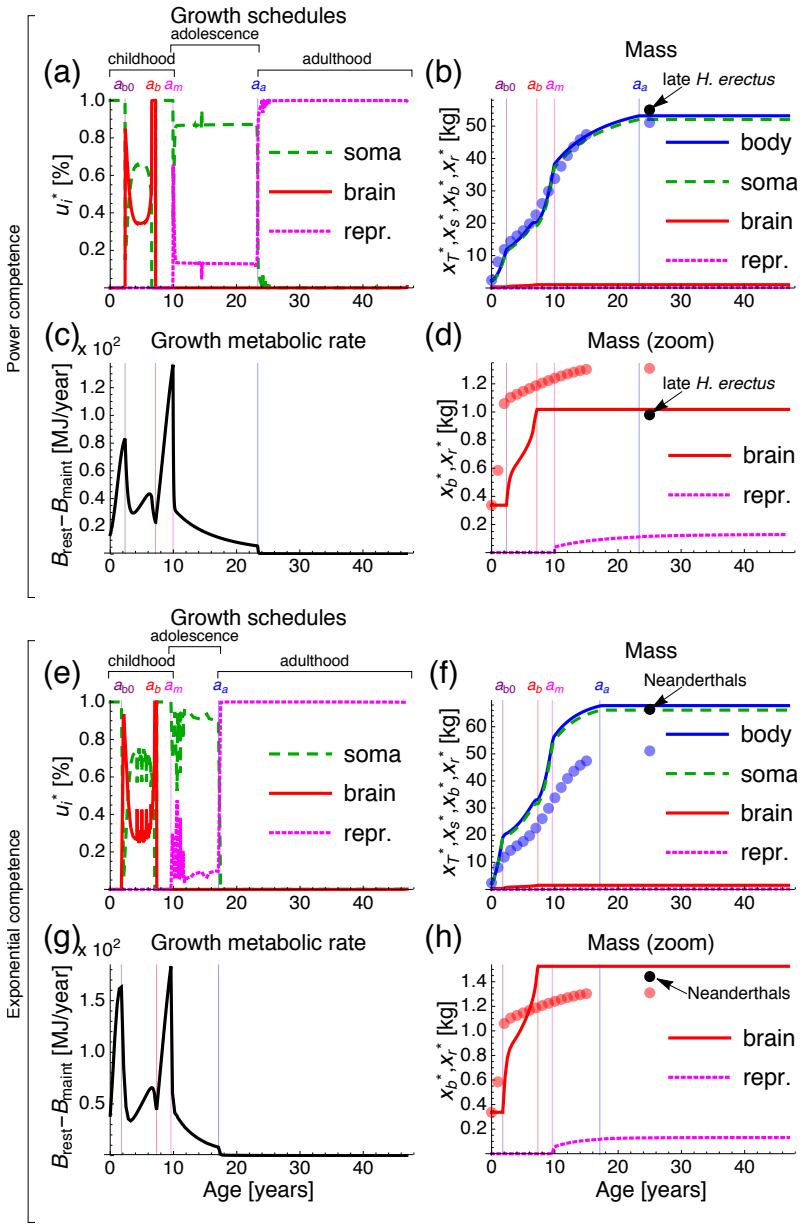
West, G.B., Brown, J.H. and Enquist, B.J. (2001). A general model for ontogenetic growth. *Nature*, **413**, 628–631.

Wrangham, R. (2009). *Catching Fire.* Basic Books.

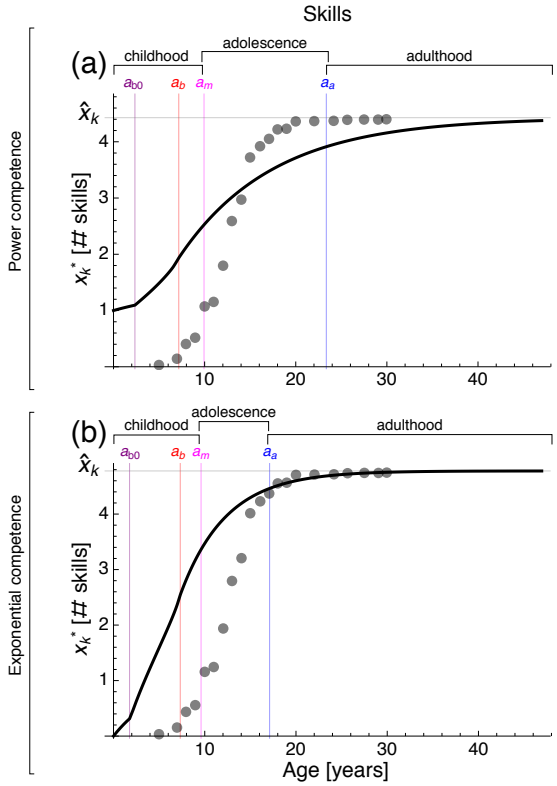
Zink, K.D. and Lieberman, D.E. (2016). Impact of meat and Lower Palaeolithic food processing techniques on chewing in human. *Nature*, **531**, 500–503.

Table 1: Life history predictions. Predicted values use competence as a power or exponential function (PC and EC) with their respective benchmark parameter values. Observed values are those in three *Homo* species. Predictions and observations with the same color (blue or red) agree. *Observed adult body mass in females and adult brain mass averaged across sexes for both late *H. erectus* (McHenry, 1994) and Neanderthals (Froehle and Churchill, 2009, Ruff *et al.*, 1997). For *H. sapiens* all values are for females: age at menarche (Gluckman and Hanson, 2006), adulthood (Henry *et al.*, 2005), brain growth onset and arrest (Kuzawa *et al.*, 2014), adult body mass (Kuzawa *et al.*, 2014), and adult brain mass (Kuzawa *et al.*, 2014). †Encephalization quotient, calculated as $\mathrm{EQ} = x_\mathrm{b}(a_\mathrm{a}) / \left[11.22 \times 10^{-3} x_\mathrm{T}(a_\mathrm{a})^{0.76}\right]$ (mass in kg) (Martin, 1981).

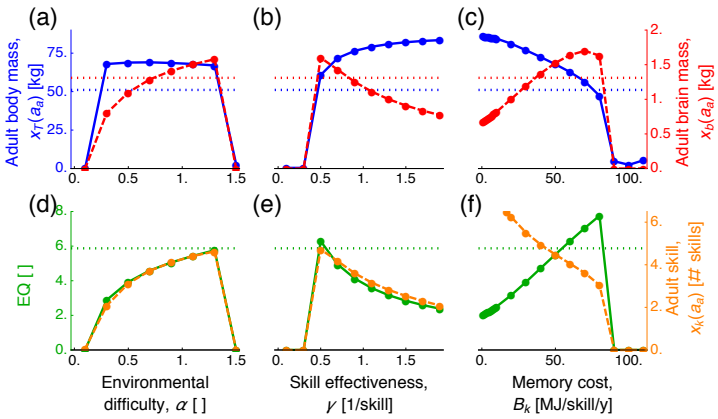
		Predicted with		Observed in*		
		PC	EC	late *H. erectus*	Neanderthals	*H. sapiens*
Age at:	Menarche, a_m [y]	9.94	9.70			7–13
	Adulthood, a_a [y]	23.37	17.33			≈17
	Brain growth onset, $a_{\mathrm{b}0}$ [y]	2.36	1.81			0
	Brain growth arrest, a_b [y]	7.19	7.34			≈17
	Adult body mass, [kg]	53.19	67.79	55	66.4	51.1
	Adult brain mass, [kg]	1.02	1.53	0.98	1.44	1.31
	EQ†, [ ]	4.43	5.52	4.15	5.30	5.87



564 Figure 1: Ecological challenges alone can generate modern human life history stages and ancient
565 human body and brain sizes. Lines are model's predictions and large dots are observations. Results
566 with (a-d) power and (e-h) exponential competence. (a,e) Predicted growth schedules vs. age. (c,g)
567 Growth metabolic rate vs. age. (b,f) Predicted body and tissue mass vs. age. (d,h) Predicted brain
568 and reproductive mass vs. age. Dots and lines with the same color are respectively the observed
569 and predicted values in modern human females (Kuzawa *et al.*, 2014). Black dots are the observed
570 (b,f) adult female body mass and (d,h) adult sex-averaged brain mass, either for late *H. erectus* or
571 Neanderthals (Table 1). Jitter in growth schedules (a,e) is due to negligible numerical error (Fig. S3).



572 Figure 2: Predicted skill ontogeny plateaus before the end of the individual's reproductive career.
573 Lines are the predicted number of skills vs. age with power (a) and exponential (b) competence for
574 the results in Fig 1. Dots are the observed cumulative distribution of self-reported acquisition ages
575 of food production skills in female Tsimane horticulturalists (Schniter *et al.*, 2015) multiplied by our
576 $\hat{x}_k$. However, note that the observed skills in Tsimane include socially learned skills which we do not
577 consider explicitly in the model.



578 Figure 3: Large adult brain mass and EQ are favored by environmental difficulty, moderate skill ef-
579 fectiveness, and costly memory. Plots are the predicted adult body and brain mass, EQ, and skill vs.
580 parameter values with exponential competence. a-c show adult body mass (blue) and adult brain
581 mass (red). d-f show adult EQ (green) and skill (orange). Vertical axes are in different scales. Dashed
582 horizontal lines are the observed values in modern human females (Kuzawa *et al.*, 2014).

583

Supporting Information for: Evolution of brain ontogenetic growth under ecological challenges

Mauricio González-Forero, Timm Faulwasser, and Laurent Lehmann

Contents

1 Model **3**
1.1 Tracking resting metabolic rate . . . 3
1.2 Energy use . . . 3
1.3 Tissue mass . . . 4
1.4 Skills . . . 5
1.5 Energy acquisition . . . 5
1.6 Fitness and evolving traits . . . 6
1.7 Model summary . . . 6

2 Optimal control problem **8**
2.1 Problem statement . . . 8
2.2 The Pontryagin Maximum Principle . . . 9

3 Analytical results **11**

4 Derivation of analytical results **15**
4.1 Singular controls for regime BS: $\sigma_s > 0$, $\sigma_b > 0$, and $\sigma_s = \sigma_b$ . . . 15
4.2 Singular controls for regime BR: $\sigma_s < 0$ and $\sigma_b = 0$ . . . 16
4.3 Singular controls for regime RS: $\sigma_s = 0$ and $\sigma_b < 0$ . . . 17
4.4 Singular controls for regime BRS: $\sigma_s = \sigma_b = 0$ . . . 18

5 Parameter values **21**

6 Estimation of parameter values **22**
6.1 Values for B_i for $i \in \{b, r, s\}$ . . . 22
6.2 Values for E_i for $i \in \{b, r, s\}$ . . . 22
6.3 Values for K and β . . . 23
6.4 Values for f_0, μ, and τ . . . 24

7 Supplementary results **25**

7.1 Brain metabolic rate through ontogeny 25

7.2 Mass of reproductive tissue 26

7.3 Effect of the absence of (allo)parental care 26

7.4 Indeterminate skill growth with inexpensive memory 27

7.5 Large, yet inconsistent-with-data encephalization with exceedingly expensive memory 28

7.6 Reproduction without growth and body collapse for certain parameter values 30

7.7 A large brain is also favored by small metabolic costs of learning, few innate skills, and intermediate allocation of brain metabolic rate to skills 32

References **33**

1 Model

In this section we derive the equations of the model presented in the main text [equations (1–4)] and formulate the evolutionary question. This question gives rise to an optimal control problem that we describe in section 2.1.

1.1 Tracking resting metabolic rate

Life history models generally study the allocation of an individual's energy budget to different functions (Kozłowski, 1992). Consequently, parameters in life history models refer to complete components of the energy budget (e.g., assimilated energy (Ziółko and Kozłowski, 1983)). In practice, it is easier to measure heat release (metabolic rates) (Blaxter, 1989). Hence, in order to facilitate parameter measurement, we follow the approach of West *et al.* (2001) to formulate our life history model in terms of resting metabolic rate allocation rather than energy budget allocation. Thus, in the model, we track how resting metabolic rate is due to growth and maintenance of different tissues, in particular the brain.

We start from the partition of the individual's energy budget used by Hou *et al.* (2008) which divides the energy budget (assimilation rate) into heat released at rest (resting metabolic rate) and the remainder (see Blaxter (1989) for details into why this partition is correct). The amount of energy used per unit time by an individual is its assimilation rate. Part of this energy per unit time is stored in the body (S) and the rest is the total metabolic rate which is the energy released as heat per unit time after use. Part of the total metabolic rate is the resting metabolic rate B_{rest} and the rest is the energy released as heat per unit time due to activity B_{act}. In turn, part of the resting metabolic rate is due to maintenance of existing biomass B_{maint}, and the rest is due to production of new biomass B_{syn}. We refer to B_{syn} as the growth metabolic rate. This partitioning is illustrated in Fig. S1. We formulate our model in terms of allocation of resting metabolic rate B_{rest} to maintenance and growth of the different tissues.

1.2 Energy use

Suppose that an individual of age a has a number $N_i(a)$ of cells of type i, for $i \in \{\mathrm{b},\mathrm{r},\mathrm{s}\}$ corresponding to brain, reproductive, and (the remainder) somatic cells, respectively. Assume that an average cell of type i in the resting body releases as heat an amount of energy B_{ci} per unit time. Hence, the total amount of energy released as heat per unit time by existing cells in the resting individual is

$$B_{\mathrm{maint}}(a) = N_{\mathrm{b}}(a)B_{\mathrm{cb}} + N_{\mathrm{r}}(a)B_{\mathrm{cr}} + N_{\mathrm{s}}(a)B_{\mathrm{cs}}, \tag{S1}$$

which gives the part of resting metabolic rate due to body mass maintenance (Hou *et al.*, 2008).

Assume that producing a new average cell of type i releases as heat an amount of energy E_{ci}. Hence, the total amount of energy released as heat per unit time by the resting individual due to production of new cells is

$$B_{\mathrm{syn}}(a) = \dot{N}_{\mathrm{b}}(a)E_{\mathrm{cb}} + \dot{N}_{\mathrm{r}}(a)E_{\mathrm{cr}} + \dot{N}_{\mathrm{s}}(a)E_{\mathrm{cs}}, \tag{S2}$$

which gives the rate of heat release in biosynthesis (Hou *et al.*, 2008), and we call it the growth metabolic rate.

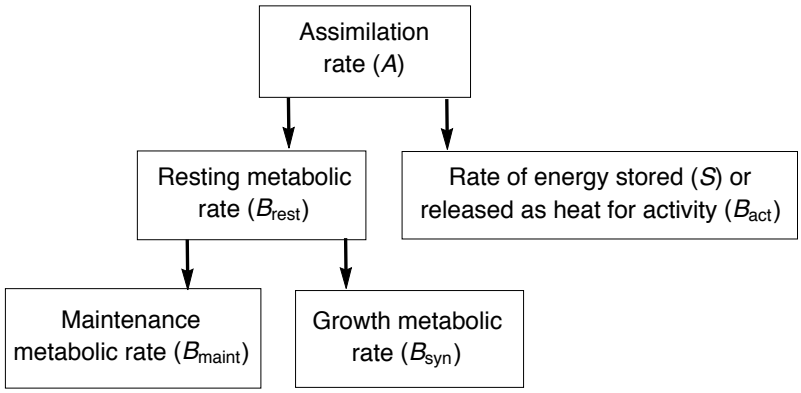


Figure S1: Relation of resting metabolic rate to assimilation rate. Modified from Hou *et al.* (2008).

From (S2), we have that

$$\dot{N}_i(a)E_{ci} = u_i(a)B_{\text{syn}}(a), \tag{S3}$$

for $i \in \{\text{b},\text{r},\text{s}\}$, where $u_i(a)$ is the fraction of growth metabolic rate due to production of new type-i cells [summing over all cell types returns (S2)].

Adding the expressions above, the total amount of energy released as heat by the resting individual per unit time is

$$B_{\text{rest}}(a) = B_{\text{maint}}(a) + B_{\text{syn}}(a). \tag{S4}$$

1.3 Tissue mass

Let the mass of an average cell of type i be x_{ci} for $i \in \{\text{b},\text{r},\text{s}\}$. Then, the mass of tissue i at age a is

$$x_i(a) = x_{ci}N_i(a), \tag{S5}$$

and hence, using (S3), we have that

$$\begin{aligned} \dot{x}_i(a) &= x_{ci}\dot{N}_i(a) \\ &= \frac{x_{ci}}{E_{ci}}u_i(a)B_{\text{syn}}(a). \end{aligned} \tag{S6}$$

Defining $E_i = E_{ci}/x_{ci}$, this gives

$$\dot{x}_i(a) = u_i(a)\frac{B_{\text{syn}}(a)}{E_i} \tag{S7}$$

for $i \in \{\text{b},\text{r},\text{s}\}$. From (S4), we then have equation (1) of the main text where from (S1)

$$B_{\text{maint}}(a) = x_{\text{b}}(a)B_{\text{b}} + x_{\text{r}}(a)B_{\text{r}} + x_{\text{s}}(a)B_{\text{s}} \tag{S8}$$

and $B_i = B_{ci}/x_{ci}$. We will denote body mass at age a as $x_T(a) = x_b(a) + x_r(a) + x_s(a)$.

1.4 Skills

We consider that some of the brain metabolic rate is to acquiring and maintaining energy-extraction skills. We assume that the individual at age a has a number $x_k(a)$ of energy-extraction skills. From energy conservation and (S1) and (S2), the brain metabolic rate must equal $M_{brain}(a) = x_b(a)B_b + \dot{x}_b(a)E_b$. We thus let υ_k be the fraction of brain metabolic rate that is due to acquiring and maintaining energy-extraction skills (or brain's allocation to energy-extraction skills). Suppose that the brain releases as heat an amount of energy E_k for acquiring an average energy-extraction skill (learning cost). Similarly, assume that the brain releases as heat an amount of energy B_k per unit time for maintaining an average energy-extraction skill (memory cost). Hence, from energy conservation,

$$x_k(a)B_k + \dot{x}_k(a)E_k = \upsilon_k\left[x_b(a)B_b + \dot{x}_b(a)E_b\right]. \tag{S9}$$

Rearranging, we have

$$\dot{x}_k(a) = \frac{\upsilon_k\left[x_b(a)B_b + \dot{x}_b(a)E_b\right] - x_k(a)B_k}{E_k}, \tag{S10}$$

which is equation (2) in the main text. [A similar reasoning can be used to derive (S7), not in terms of allocation to tissue growth $u_i(a)$, but in terms of allocation to tissue growth *and* maintenance $\upsilon_i(a)$.]

1.5 Energy acquisition

We now derive an expression that specifies how energy extraction affects fitness in the model. To that end, we assume that at age a the individual obtains an amount of energy $E(x_k(a))$ per unit time from the environment, which we assume depends on skill $x_k(a)$ (and possibly body mass). The quantity $E(x_k(a))$ is thus the individual's energetic production per unit time at age a. Let $E_{max}(a)$ be the amount of energy that the individual obtains from the environment per unit time at age a if it is maximally successful at energy extraction (which also possibly depends on body mass). Let us use $x \equiv y$ to denote that x is defined as y. Then, we define the probability of energy extraction at age t as the normalized production per unit time at age a:

$$p(x_k(a)) \equiv \frac{E(x_k(a))}{E_{max}(a)}. \tag{S11}$$

We also define the ratio of resting metabolic rate and energy obtained per unit time as

$$q(x_k(a)) \equiv \frac{B_{rest}(a)}{E(x_k(a))} \tag{S12}$$

and, motivated by (S12), the quantity

$$B_{rest,max}(a) \equiv q(x_k(a))\,E_{max}(a) \tag{S13a}$$

$$= \frac{B_{rest}(a)}{p(x_k(a))}. \tag{S13b}$$

From (S13b), we have that

$$B_{rest}(a) = p(x_k(a))\,B_{rest,max}(a). \tag{S14}$$

Consequently, $B_{rest,max}(a)$ gives the resting metabolic rate when the individual is maximally successful at energy extraction. Adult resting metabolic rate typically scales with adult body mass as a power law across all

living systems (Kleiber, 1932, 1961; Peters, 1983; Schmidt-Nielsen, 1984), and also ontogenetically in humans to a good approximation (Fig. S4; see also Sears *et al.* (2012)). Hence, assuming that $p(x_k(a))$ is independent of body mass, we assume that

$$B_{\mathrm{rest,max}}(a) = K x_T(a)^{\beta}, \tag{S15}$$

where K is a constant independent of body mass. Equation (S14) then becomes equation (4) in the main text.

1.6 Fitness and evolving traits

We consider the growth schedules $u_i(a)$ for $i \in \{\mathrm{b}, \mathrm{r}, \mathrm{s}\}$ as evolving traits, and we make assumptions (see below) that imply that these schedules attain evolutionarily stable values (Lande, 1982; Mylius and Diekmann, 1995). To obtain evolutionarily stable growth schedules we need a fitness measure. To obtain this measure, we consider a randomly mating population of constant size, with overlapping generations, where the environment is constant, and where the age of individuals is measured in continuous time. We assume that the probability $l(a)$ that a newborn survives to age t satisfies

$$\dot{l}(a) = -\mu l(a) \tag{S16}$$

where μ is the mortality rate. For simplicity, we take mortality rate as constant.

We obtain a measure of fertility as follows. We partition the mass-specific resting metabolic rate of reproductive tissue B_r into a component due to maintenance of reproductive tissue itself B_{ra} and a component due to production of offspring cells B_{ro}. That is, $B_r = B_{ra} + B_{ro}$ (note that B_{ro} is not part of E_r because the latter refers to the production of mother's cells). Let $\dot{N}_o(a)$ be the number of offspring cells produced by the individual per unit time at age t. Hence, the number of offspring cells produced is given by $\dot{N}_o(a) = C_1 B_{ra} N_r(a)$ for some constant C_1. Then, we assume that fertility, defined as the number of offspring produced per unit time at age a, is

$$f(a) = C_2 \dot{N}_o(a) = C_3 N_r(a) = f_0 x_r(a), \tag{S17}$$

where C_2, C_3, and f_0 are proportionality constants defined in the absence of density dependence competition. We also assume that costs of parental or alloparental care are included in f_0. Fertility is then proportional to the mass of reproductive tissue (King and Roughgarden, 1982).

From (S16)–(S17), the individual's lifetime number of offspring produced in the absence of density-dependent competition (Mylius and Diekmann, 1995) is then given by

$$R_0 = \int_0^{\tau} l(a) f(a) \mathrm{d}a, \tag{S18}$$

where τ is an age after which the individual no longer reproduces. With additional standard assumptions, evolutionarily stable growth schedules in the population of constant size regulated through fertility must maximize R_0 (Mylius and Diekmann, 1995), and so we take R_0 as a fitness (objective) function that is maximized by the evolving growth schedules $u_i(a)$ at an evolutionary equilibrium.

1.7 Model summary

Our model specifies the ontogenetic dynamics of the brain, reproductive, and somatic tissue mass, x_b, x_r and x_s, and of the number of energy-extraction skills x_k of the individual. The dynamics of these four state vari-

ables is expressed in terms of the growth schedules $u_i(a)$ that we take as evolving traits and of 22 parameters: namely, 11 tissue- and skill-metabolism parameters (K, β, ν_{k}, and B_i and E_i for $i \in \{\mathrm{b},\mathrm{r},\mathrm{s},\mathrm{k}\}$); 3 demographic parameters (f_0, μ, and τ); 2 contest success parameters (α and γ); 2 (allo)parental care parameters (φ_0 and φ_{r}); and 4 newborn tissue mass and newborn skill parameters [$x_i(0)$ for $i \in \{\mathrm{b},\mathrm{r},\mathrm{s},\mathrm{k}\}$]. Parameter f_0 only displaces the objective vertically and thus has no effect on the optimal growth schedules.

We now formulate the optimal control problem posed by our evolutionary model and later describe how we estimated parameter values from empirical data.

2 Optimal control problem

2.1 Problem statement

The maximization of R_0 by the growth schedules $u_i(a)$ for all $a \in [0,\tau]$ [or $u(\cdot)$ for short] poses an optimal control problem (King and Roughgarden, 1982; Iwasa and Roughgarden, 1984; Perrin, 1992; Irie and Iwasa, 2005; Sydsæter *et al.*, 2008). In the terminology of optimal control theory, we have the control variables

$$\mathbf{u}(a) = (u_\mathrm{b}(a), u_\mathrm{r}(a), u_\mathrm{s}(a)) \in [0,1] \quad \text{subject to} \quad u_\mathrm{b}(\cdot) + u_\mathrm{r}(\cdot) + u_\mathrm{s}(\cdot) = 1, \tag{S19a}$$

and the state variables

$$\mathbf{x}(a) = (x_\mathrm{b}(a), x_\mathrm{r}(a), x_\mathrm{s}(a), x_\mathrm{k}(a)) \geq 0. \tag{S19b}$$

For readability, we will suppress the argument in $\mathbf{u}(a)$ and $\mathbf{x}(a)$, and write $\mathbf{u}$ and $\mathbf{x}$.

We then have the optimal control problem

$$\max_{\mathbf{u}(\cdot)} R_0, \tag{S19c}$$

where from (S16)–(S18)

$$R_0 = f_0 \int_0^\tau \mathrm{e}^{-\mu a} x_\mathrm{r} \mathrm{d}a, \tag{S19d}$$

subject to the dynamic constraints

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{u}, \mathbf{x}, a), \tag{S19e}$$

with

$$g_i(\mathbf{u}, \mathbf{x}, a) = e_i u_i B_\mathrm{syn}(\mathbf{x}, a) \quad \text{for } i \in \{\mathrm{b}, \mathrm{r}, \mathrm{s}\} \tag{S19f}$$

$$g_\mathrm{k}(\mathbf{u}, \mathbf{x}, a) = d_1 \left[x_\mathrm{b} B_\mathrm{b} + u_\mathrm{b} B_\mathrm{syn}(\mathbf{x}, a) \right] - d_2 x_\mathrm{k}, \tag{S19g}$$

which are obtained from (S7) and (S10), where $e_i = 1/E_i$, $d_1 = \nu_\mathrm{k}/E_\mathrm{k}$, and $d_2 = B_\mathrm{k}/E_\mathrm{k}$. From (S4), (S8), (S14), and (S15), we have that growth metabolic rate is

$$B_\mathrm{syn}(\mathbf{x}, a) = K p(x_\mathrm{k}, a) x_\mathrm{T}^\beta - B_\mathrm{b} x_\mathrm{b} - B_\mathrm{r} x_\mathrm{r} - B_\mathrm{s} x_\mathrm{s}, \tag{S19h}$$

where body mass is

$$x_\mathrm{T} = x_\mathrm{b} + x_\mathrm{r} + x_\mathrm{s}, \tag{S19i}$$

and, from (5) in the main text, the probability of energy extraction at age a is

$$p(x_\mathrm{k}, a) = \frac{c(x_\mathrm{k})}{\alpha - \varphi_0 \mathrm{e}^{-\varphi_\mathrm{r} a} + c(x_\mathrm{k})}, \tag{S19j}$$

where competence at energy extraction is

$$c(x_\mathrm{k}) = \begin{cases} x_\mathrm{k}^\gamma & \text{(power competence)} \\ \mathrm{e}^{\gamma x_\mathrm{k}} & \text{(exponential competence).} \end{cases} \tag{S19k}$$

Finally, the initial conditions of (S19e) are

$$x_i(0) = x_{i0} \quad \text{for all } i \tag{S19l}$$

and we do not consider any terminal conditions for (S19e) .

2.2 The Pontryagin Maximum Principle

Necessary first-order conditions for maximizing the objective R_0 with respect to the controls throughout t are given by the Pontryagin maximum principle (Bryson, Jr. and Ho, 1975; Kamien and Schwartz, 2012; Sydsæter *et al.*, 2008). The Pontryagin maximum principle states that if $(\mathbf{u}^*, \mathbf{x}^*)$ is a solution to the optimal control problem (S19), then an associated function, the Hamiltonian, is maximized with respect to the controls when evaluated at $(\mathbf{u}^*, \mathbf{x}^*)$. The Hamiltonian for problem (S19) is

$$H(\mathbf{u}, \mathbf{x}, \boldsymbol{\lambda}, a) = f_0 \mathrm{e}^{-\mu a} x_\mathrm{r} + \sum_{i \in \{\mathrm{b,r,s,k}\}} \lambda_i g_i(\mathbf{u}, \mathbf{x}, a), \tag{S20}$$

where λ_i is the costate variable associated to state variable i and $\boldsymbol{\lambda}$ is the vector of costates. Here we also drop the argument of $\lambda_i(a)$ and write simply λ_i. A costate variable gives the marginal value of the corresponding state variable; that is, it is the effect on the maximized objective (fitness) for a marginal change in the corresponding state variable (Dorfman, 1969). Thus, we now proceed to maximize the Hamiltonian to obtain candidate optimal controls $\mathbf{u}^*$ that satisfy these necessary conditions for optimality.

Due to the constraint $u_\mathrm{b} + u_\mathrm{r} + u_\mathrm{s} = 1$, we set $u_\mathrm{r} = 1 - u_\mathrm{b} - u_\mathrm{s}$ and only two controls must be determined: u_b^* and u_s^*. Using (S19f) and (S19g), collecting for B_syn in (S20), and evaluating at $\mathbf{x} = \mathbf{x}^*$ we have

$$H(\mathbf{u}, \mathbf{x}^*, \boldsymbol{\lambda}, a) = f_0 \mathrm{e}^{-\mu a} x_\mathrm{r}^* + B_\mathrm{syn}(\mathbf{x}^*, a)\phi(\mathbf{u}, \lambda) + \lambda_\mathrm{k}\xi(\mathbf{x}^*), \tag{S21}$$

where

$$\phi(\mathbf{u}, \boldsymbol{\lambda}) = u_\mathrm{b}\sigma_\mathrm{b} + u_\mathrm{s}\sigma_\mathrm{s} + e_\mathrm{r}\lambda_\mathrm{r} \tag{S22a}$$

$$\xi(\mathbf{x}^*) = d_1 x_\mathrm{b}^* B_\mathrm{b} - d_2 x_\mathrm{k}^* \tag{S22b}$$

and

$$\sigma_\mathrm{b}(\boldsymbol{\lambda}) = e_\mathrm{b}\lambda_\mathrm{b} - e_\mathrm{r}\lambda_\mathrm{r} + d_1\lambda_\mathrm{k} \tag{S23a}$$

$$\sigma_\mathrm{s}(\boldsymbol{\lambda}) = e_\mathrm{s}\lambda_\mathrm{s} - e_\mathrm{r}\lambda_\mathrm{r}. \tag{S23b}$$

We thus seek to maximize (S21) with respect to $\mathbf{u} = (u_\mathrm{b}, u_\mathrm{s})$.

The derivatives of the Hamiltonian (S21) with respect to the two controls (u_b, u_s) are [see equation (10) on p. 126 of Kamien and Schwartz (2012)]

$$\left.\frac{\partial H(\mathbf{u}, \mathbf{x}^*, \boldsymbol{\lambda}, a)}{\partial u_i}\right|_{\mathbf{u}=\mathbf{u}^*} = B_\mathrm{syn}\sigma_i \qquad \text{for } i \in \{\mathrm{b,s}\}. \tag{S24}$$

If $B_\mathrm{syn} > 0$, then the Hamiltonian is maximized with respect to u_b and u_s depending on the signs of the switching functions σ_i and, because of the constraint that $u_\mathrm{b} + u_\mathrm{s} \leq 1$, also depending on the sign of the difference

$$\sigma_\mathrm{s} - \sigma_\mathrm{b} = e_\mathrm{s}\lambda_\mathrm{s} - e_\mathrm{b}\lambda_\mathrm{b} - d_1\lambda_\mathrm{k}. \tag{S25}$$

By definition, the costates satisfy [see equation (7) on p. 126 of Kamien and Schwartz (2012)]

$$\dot{\lambda}_i = -\left.\frac{\partial H(\mathbf{u}^*, \mathbf{x}, \boldsymbol{\lambda}, a)}{\partial x_i}\right|_{\mathbf{x}=\mathbf{x}^*} \qquad \text{for } i \in \{\mathrm{b,r,s,k}\} \tag{S26a}$$

$$\lambda_i(\tau) = 0. \tag{S26b}$$

185 Hence, the dynamical equations of the costates are

$$\dot{\lambda}_{\mathrm{b}} = -\left(\phi\psi_{\mathrm{b}} + \lambda_{\mathrm{k}} d_1 B_{\mathrm{b}}\right) \tag{S27a}$$

$$\dot{\lambda}_{\mathrm{r}} = -\left(\phi\psi_{\mathrm{r}} + f_0 \mathrm{e}^{-\mu a}\right) \tag{S27b}$$

$$\dot{\lambda}_{\mathrm{s}} = -\phi\psi_{\mathrm{s}} \tag{S27c}$$

$$\dot{\lambda}_{\mathrm{k}} = -\left(\phi\psi_{\mathrm{k}} - \lambda_{\mathrm{k}} d_2\right), \tag{S27d}$$

186 evaluated at $(\mathbf{x}^*, \mathbf{u}^*)$, where we define

$$\psi_i(\mathbf{x}^*, a) = \left.\frac{\partial B_{\mathrm{syn}}}{\partial x_i}\right|_{\mathbf{x}=\mathbf{x}^*} \tag{S28}$$

187 for $i \in \{\mathrm{b}, \mathrm{r}, \mathrm{s}, \mathrm{k}\}$. Note that the marginal returns on energy extraction from increasing skill and skill synergy are

188 respectively

$$\frac{\partial p}{\partial x_{\mathrm{k}}} = p(1-p)\frac{\mathrm{d}\ln c(x_{\mathrm{k}})}{\mathrm{d}x_{\mathrm{k}}} \tag{S29a}$$

$$= p(1-p)\frac{\gamma}{\delta(x_{\mathrm{k}})} \tag{S29b}$$

$$\frac{\partial^2 p}{\partial x_{\mathrm{k}}^2} = p(1-p)\left[\frac{\mathrm{d}^2\ln c(x_{\mathrm{k}})}{\mathrm{d}x_{\mathrm{k}}^2} + (1-2p)\left(\frac{\mathrm{d}\ln c(x_{\mathrm{k}})}{\mathrm{d}x_{\mathrm{k}}}\right)^2\right] \tag{S29c}$$

$$= p(1-p)\frac{\gamma}{\delta(x_{\mathrm{k}})^2}\left[\gamma(1-2p) - \hat{\delta}\right], \tag{S29d}$$

189 where

$$\delta(x_{\mathrm{k}}^*) = \begin{cases} x_{\mathrm{k}}^* & \text{for } c(x_{\mathrm{k}}) = x_{\mathrm{k}}^{\gamma} \\ 1 & \text{for } c(x_{\mathrm{k}}) = \mathrm{e}^{\gamma x_{\mathrm{k}}}, \end{cases} \tag{S30a}$$

$$\hat{\delta} = \begin{cases} 1 & \text{if } c(x_{\mathrm{k}}) = x_{\mathrm{k}}^{\gamma} \\ 0 & \text{if } c(x_{\mathrm{k}}) = \mathrm{e}^{\gamma x_{\mathrm{k}}}. \end{cases} \tag{S30b}$$

190 Hence,

$$\psi_i(\mathbf{x}^*, a) = \psi(\mathbf{x}^*, a) - B_i \qquad \text{for } i \in \{\mathrm{b}, \mathrm{r}, \mathrm{s}\} \tag{S31a}$$

$$\psi_{\mathrm{k}}(\mathbf{x}^*, a) = K\frac{\partial p}{\partial x_{\mathrm{k}}} x_{\mathrm{T}}(\mathbf{x}^*)^{\beta} \tag{S31b}$$

$$= K\gamma p(x_{\mathrm{k}}^*, a)[1 - p(x_{\mathrm{k}}^*, a)]\frac{x_{\mathrm{T}}(\mathbf{x}^*)^{\beta}}{\delta(x_{\mathrm{k}}^*)}, \tag{S31c}$$

191 whereby

$$\psi(\mathbf{x}^*, a) = K\beta p(x_{\mathrm{k}}^*, a) x_{\mathrm{T}}(\mathbf{x}^*)^{\beta-1}. \tag{S31d}$$

3 Analytical results

We present the analytical results for the candidate optimal controls in this section, and their derivations in section 4. In these two sections, we assume that growth metabolic rate is positive; that is, $B_{\text{syn}}(\mathbf{x}^*, t) > 0$.

The Hamiltonian of the optimal control problem (S19) is affine (or, less rigorously, linear) in the controls [equation (S21)]. Since we assume that $B_{\text{syn}}(\mathbf{x}^*, a) > 0$, the sign of the derivative of the Hamiltonian with respect to u_s or u_b is given by the sign of the two switching functions σ_s and σ_b [equations (S23)]. If σ_i is negative, the Hamiltonian is maximized when $u_i^* = 0$. If σ_i is positive and the other switching function, denoted by $\sigma_{i'}$, is negative, then the Hamiltonian is maximized when $(u_i^*, u_{i'}^*) = (1, 0)$. If both σ_i and $\sigma_{i'}$ are positive, because of the constraint that $u_\text{s}^* + u_\text{b}^* \leq 1$, the Hamiltonian is maximized when $(u_i^*, u_{i'}^*) = (1, 0)$ if and only if $\sigma_i > \sigma_{i'}$. If σ_i is zero and $\sigma_{i'}$ is positive, then the Hamiltonian is maximized when $(u_i^*, u_{i'}^*) = (0, 1)$. If σ_i is zero and $\sigma_{i'}$ is negative, then the Hamiltonian is maximized when $u_{i'}^* = 0$ but the Hamiltonian is independent of u_i. In this case, the candidate optimal control $u_i^* = \hat{u}_i$ is called a singular arc and must be determined by another method (Bryson, Jr. and Ho, 1975). If both σ_s and σ_b are zero, the Hamiltonian is independent of both controls and the candidate optimal controls are the singular arcs $(u_\text{s}^*, u_\text{b}^*) = (\hat{u}_\text{s}, \hat{u}_\text{s})$. Finally, if both σ_s and σ_b are positive and equal, then both u_s^* and u_b^* are positive and maximal given the constraint $u_\text{s}^* + u_\text{b}^* \leq 1$, so $(u_\text{s}^*, u_\text{b}^*) = (1 - \hat{u}_\text{b}, \hat{u}_\text{b})$.

Together, these cases show that there are seven possible growth regimes (Table S1). Regimes B, R, and S involve pure growth of one of the three tissues, whereas regimes BS, BR, RS, and BRS are singular arcs where at least two tissues grow simultaneously. These regimes occur as indicated in Table S1 depending on the sign of both the switching functions and their difference. Numerical illustration of these regimes is given in Fig. S2.

Regime	Tissues growing	Candidate optimal controls	Sign of switching functions
		$(u_\text{s}^*, u_\text{b}^*)$	$\text{sign}(\sigma_\text{s}, \sigma_\text{b}, \sigma_\text{s} - \sigma_\text{b})$
R	Reproductive	$(0,0)$	$(-,-,\cdot)$
B	Brain	$(0,1)$	$(-,+,\cdot)$, $(+,+,-)$, $(0,+,\cdot)$
S	Soma	$(1,0)$	$(+,-,\cdot)$, $(+,+,+)$, $(+,0,\cdot)$
BS	Brain and soma	$(1-\hat{u}_\text{b}, \hat{u}_\text{b})$	$(+,+,0)$
BR	Brain and reproductive	$(0, \hat{u}_\text{b})$	$(-,0,\cdot)$
RS	Reproductive and soma	$(\hat{u}_\text{s}, 0)$	$(0,-,\cdot)$
BRS	Brain, reproductive, and soma	$(\hat{u}_\text{s}, \hat{u}_\text{b})$	$(0,0,\cdot)$

Table S1: Growth regimes. Four regimes are singular arcs. Note that $u_\text{r}^* = 1 - u_\text{s}^* - u_\text{b}^*$. The "$\cdot$" means any sign.

For simplicity of presentation in the remainder of section 3 and 4, we will explicitly write the arguments of a function only when defining the function and will suppress their writing elsewhere, except in a few places where it is useful to recall them.

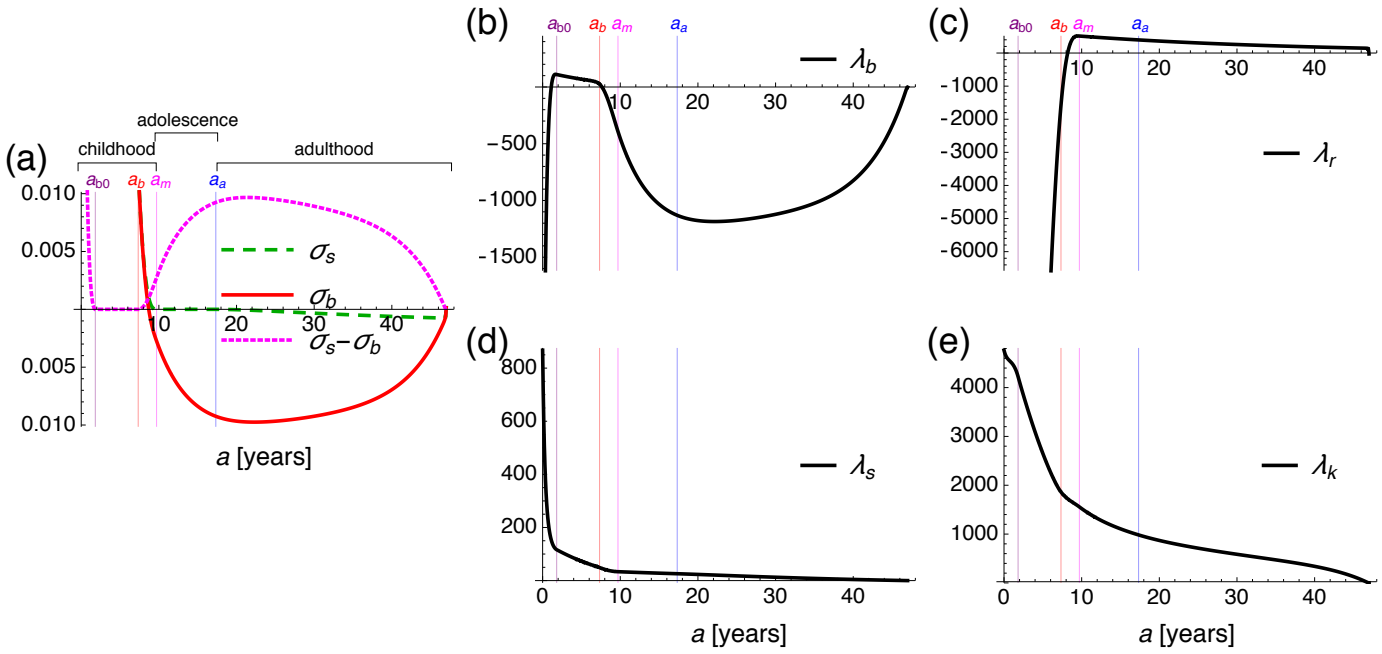


Figure S2: Switching functions and costates for the process in Fig. 1. GPOPS yields the costates λ_i using a direct approach rather than the Pontryagin maximum principle (Patterson and Rao, 2014). The switching functions σ_i are calculated using (S23).

In section 4 we show that for the singular arcs and assuming the denominators are non-zero, the candidate optimal controls are

$$\text{Regime BS:} \quad \hat{u}_{\mathrm{b}}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{\rho_{\mathrm{sk}} - \chi^{\mathrm{sb}}_{/\mathrm{sk}}}{\chi^{\mathrm{sb}}_{\mathrm{bsk}}} \tag{S32a}$$

$$\text{Regime BR:} \quad \hat{u}_{\mathrm{b}}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{\rho_{\mathrm{rk}} - \chi^{\mathrm{br}}_{/\mathrm{rk}}}{\chi^{\mathrm{br}}_{\mathrm{brk}}} \tag{S32b}$$

$$\text{Regime RS:} \quad \hat{u}_{\mathrm{s}}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{\rho_{\mathrm{rs}} - \chi^{\mathrm{sr}}_{/\mathrm{r}/}}{\chi^{\mathrm{sr}}_{\mathrm{sr}/}} \tag{S32c}$$

$$\text{Regime BRS:} \quad \hat{u}_{\mathrm{s}}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{(\rho_{\mathrm{rs}} - \chi^{\mathrm{sr}}_{/\mathrm{sk}})\chi^{\mathrm{br}}_{\mathrm{brk}} - (\rho_{\mathrm{rk}} - \chi^{\mathrm{br}}_{/\mathrm{rk}})\chi^{\mathrm{sr}}_{\mathrm{br}/}}{\chi^{\mathrm{sr}}_{\mathrm{sr}/}\chi^{\mathrm{br}}_{\mathrm{brk}} - \chi^{\mathrm{sr}}_{\mathrm{br}/}\chi^{\mathrm{br}}_{\mathrm{srk}}} \tag{S32d}$$

$$\hat{u}_{\mathrm{b}}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{(\rho_{\mathrm{rk}} - \chi^{\mathrm{br}}_{/\mathrm{rk}})\chi^{\mathrm{sr}}_{\mathrm{sr}/} - (\rho_{\mathrm{rs}} - \chi^{\mathrm{sr}}_{/\mathrm{r}/})\chi^{\mathrm{br}}_{\mathrm{srk}}}{\chi^{\mathrm{sr}}_{\mathrm{sr}/}\chi^{\mathrm{br}}_{\mathrm{brk}} - \chi^{\mathrm{sr}}_{\mathrm{br}/}\chi^{\mathrm{br}}_{\mathrm{srk}}}. \tag{S32e}$$

Here we have

$$\chi^{lm}_{ijk}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{e_j \lambda_j}{x_{\mathrm{T}}}\left[\psi \omega_{ij}(e_l - e_m) + \theta_0 \psi_{\mathrm{k}} \eta_{ij} d_1\right] \tag{S33a}$$

$$\rho_{jk}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \theta_1\Big(\theta_2 d_1 d_2 \lambda_{\mathrm{k}}(e_{\mathrm{b}} B_{\mathrm{b}} - d_2) \tag{S33b}$$

$$+ e_j \lambda_j \left\{e_j \psi_j \left[e_j \psi_j - \theta_2(e_{\mathrm{b}}\psi_{\mathrm{b}} + d_1 \psi_{\mathrm{k}}) - \hat{\theta}_2 e_{\mathrm{s}} \psi_{\mathrm{s}}\right] - \theta_2 d_1 \psi_{\mathrm{k}}(e_{\mathrm{b}} B_{\mathrm{b}} - d_2)\right\} \tag{S33c}$$

$$+ \theta_3 f_0 \mathrm{e}^{-\mu a}\left[\mu + (e_j \psi_j - \theta_2 e_{\mathrm{b}} \psi_{\mathrm{b}} - \hat{\theta}_2 e_{\mathrm{s}} \psi_{\mathrm{s}})\right]\Big), \tag{S33d}$$

for $i, j, k, l, m \in \{\mathrm{b}, \mathrm{r}, \mathrm{s}, \mathrm{k}\}$, and a subscript "/" in χ^{lm}_{ijk} in (S32) denotes a removed subscript. In turn, functions defining the χ^{lm}_{ijk}'s and ρ_{jk}'s functions are

$$\omega_{\mathrm{sr}}(\mathbf{x}^*, a) = B_{\mathrm{syn}}(\mathbf{x}^*, a)(\beta - 1)(e_{\mathrm{s}} - e_{\mathrm{r}}) \tag{S34a}$$

$$\omega_{\mathrm{br}}(\mathbf{x}^*, a) = B_{\mathrm{syn}}(\mathbf{x}^*, a)\left[(\beta - 1)(e_{\mathrm{b}} - e_{\mathrm{r}}) + \gamma d_1 \frac{x_{\mathrm{T}}(\mathbf{x}^*)}{\delta(x^*_{\mathrm{k}})}(1 - p(x^*_{\mathrm{k}}, a))\right] \tag{S34b}$$

$$\omega_{\mathrm{r}}(\mathbf{x}^*, a) = B_{\mathrm{syn}}(\mathbf{x}^*, a)(\beta - 1)e_{\mathrm{r}} + \gamma \frac{x_{\mathrm{T}}(\mathbf{x}^*)}{\delta(x^*_{\mathrm{k}})}\xi(\mathbf{x}^*)(1 - p(x^*_{\mathrm{k}}, a)) - \frac{x_{\mathrm{T}}(\mathbf{x}^*)}{c(x^*_{\mathrm{k}})} p(x^*_{\mathrm{k}}, a)\varphi_{\mathrm{r}}\varphi(a) \tag{S34c}$$

$$\eta_{\mathrm{sr}}(\mathbf{x}^*, a) = \beta B_{\mathrm{syn}}(\mathbf{x}^*, a)(e_{\mathrm{s}} - e_{\mathrm{r}}) \tag{S34d}$$

$$\eta_{\mathrm{br}}(\mathbf{x}^*, a) = B_{\mathrm{syn}}(\mathbf{x}^*, a)\left\{\beta(e_{\mathrm{b}} - e_{\mathrm{r}}) + \frac{x_{\mathrm{T}}(\mathbf{x}^*)}{\delta(x^*_{\mathrm{k}})} d_1 \left[\gamma(1 - 2p(x^*_{\mathrm{k}}, a)) - \hat{\delta}\right]\right\} \tag{S34e}$$

$$\eta_{\mathrm{r}}(\mathbf{x}^*, a) = \beta B_{\mathrm{syn}}(\mathbf{x}^*, a)e_{\mathrm{r}} + \frac{x_{\mathrm{T}}(\mathbf{x}^*)}{\delta(x^*_{\mathrm{k}})}\xi(\mathbf{x}^*)\left[\gamma(1 - 2p(x^*_{\mathrm{k}}, a)) - \hat{\delta}\right] + x_{\mathrm{T}}(\mathbf{x}^*)\varphi_{\mathrm{r}}\varphi(a)\left(\frac{1}{\alpha - \varphi(a)} - 2\frac{p(x^*_{\mathrm{k}}, a)}{c(x^*_{\mathrm{k}})}\right), \tag{S34f}$$

where

$$\omega_{\mathrm{bs}}(\mathbf{x}^*, a) = \omega_{\mathrm{br}} - \omega_{\mathrm{sr}} \tag{S35a}$$

$$\omega_{\mathrm{s}}(\mathbf{x}^*, a) = \omega_{\mathrm{sr}} + \omega_{\mathrm{r}} \tag{S35b}$$

$$\eta_{\mathrm{bs}}(\mathbf{x}^*, a) = \eta_{\mathrm{br}} - \eta_{\mathrm{sr}} \tag{S35c}$$

$$\eta_{\mathrm{s}}(\mathbf{x}^*, a) = \eta_{\mathrm{sr}} + \eta_{\mathrm{r}}. \tag{S35d}$$

Finally, to complete the specification of (S32), we have

$$\theta_0 = \begin{cases} 0, & \text{if } (l,m) = (\mathrm{s},\mathrm{r}) \\ -1, & \text{if } (l,m) = (\mathrm{s},\mathrm{b}) \\ 1, & \text{otherwise} \end{cases}; \tag{S36a}$$

$$\theta_1 = \begin{cases} 1, & \text{if } (j,k) = (\mathrm{s},\mathrm{k}) \\ -1, & \text{if } (j,k) = (\mathrm{r},\mathrm{k}),(\mathrm{r},\mathrm{s}) \end{cases}; \quad \theta_2 = \begin{cases} 1, & \text{if } (j,k) = (\mathrm{s},\mathrm{k}),(\mathrm{r},\mathrm{k}) \\ 0, & \text{if } (j,k) = (\mathrm{r},\mathrm{s}) \end{cases} \tag{S36b}$$

$$\hat{\theta}_2 = \begin{cases} 0, & \text{if } (j,k) = (\mathrm{s},\mathrm{k}),(\mathrm{r},\mathrm{k}) \\ 1, & \text{if } (j,k) = (\mathrm{r},\mathrm{s}) \end{cases}; \quad \theta_3 = \begin{cases} 0, & \text{if } (j,k) = (\mathrm{s},\mathrm{k}) \\ 1, & \text{if } (j,k) = (\mathrm{r},\mathrm{k}),(\mathrm{r},\mathrm{s}). \end{cases} \tag{S36c}$$

The analytical solutions for the candidate optimal controls given by Table S1 and (S32) are functions of the candidate optimal states $\mathbf{x}^*$ and costates $\boldsymbol{\lambda}$, which we have not specified analytically. To assess if these analytical candidate optimal controls are indeed optimal, we compare them to optimal controls found numerically by GPOPS (Patterson and Rao, 2014) (Fig. 1a,e). GPOPS uses a direct approach to solve optimal control problems by iterating varying controls and determining which improves maximization of the objective (Patterson and Rao, 2014), rather than the indirect approach of the Pontryagin maximum principle via necessary conditions for optimality (see Diehl *et al.* (2006) for a comparison of direct and indirect solution approaches to optimal control problems). From the numerical solutions given by GPOPS, we obtain optimal states and their costates which are part of the output given by GPOPS (Fig. S2b-e). Feeding these numerically obtained optimal states and costates to the expressions for the analytical candidate optimal control, we plot in Fig. S3 the analytical solutions for the candidate optimal controls given by Table S1 and (S32). Comparison with Fig. 1a,e shows that the analytical candidate optimal controls closely follow the controls found numerically by GPOPS.

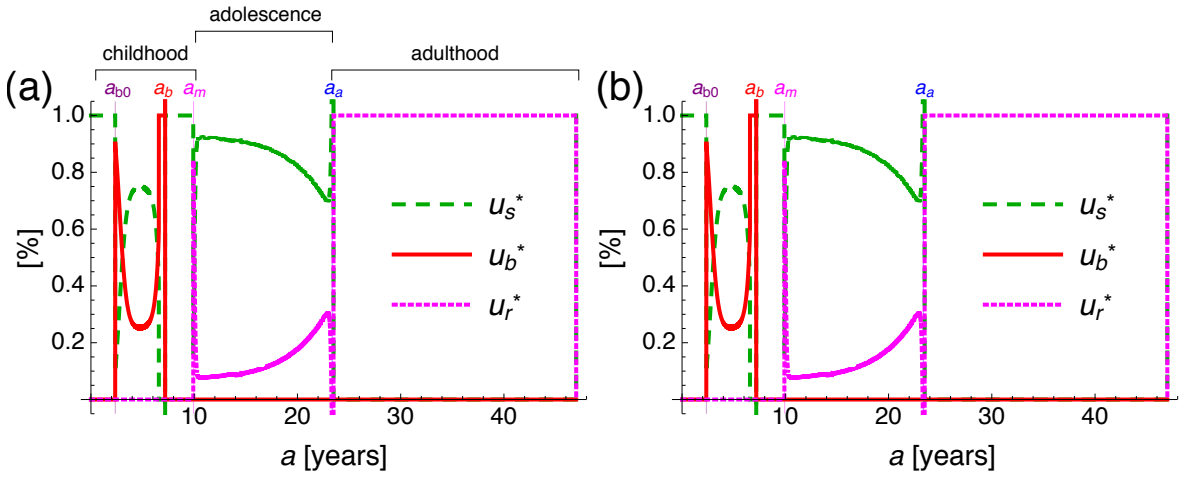


Figure S3: Plots of the analytically found candidate optimal controls. (a) is for the power competence case in Fig. 1a-d. (b) is for the exponential competence case in Fig. 1e-h. Near the switching points between regimes $(a_{\mathrm{b0}}, a_{\mathrm{b}}, a_{\mathrm{m}}, a_{\mathrm{a}})$, the analytically found controls can be greater than one or smaller than zero, possibly due to negligible numerical error in the location of the switching points.

4 Derivation of analytical results

Here we derive the expressions for $\hat{u}_\mathrm{b}(\mathbf{x}^*,\boldsymbol{\lambda},t)$ and $\hat{u}_\mathrm{s}(\mathbf{x}^*,\boldsymbol{\lambda},t)$ during the singular arcs given by (S32). To do so, we make use of the well-known result, according to which $\hat{u}_\mathrm{b}$ and $\hat{u}_\mathrm{s}$ can be obtained from the age derivatives of the switching functions up to some even, but not odd, order (Kelley *et al.*, 1967). Note that during singular arcs, either $\sigma_i = 0$ for some i or the difference $\sigma_\mathrm{s} - \sigma_\mathrm{b} = 0$, and hence their age derivatives also equal zero during the singular arcs. We thus obtain the singular controls by taking second age derivatives of these functions, but before doing so, we obtain expressions that will be useful.

By differentiating (S31d) and (S31c) with respect to age, we obtain

$$\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi}{x_\mathrm{T}}\left(u_\mathrm{s}^*\omega_\mathrm{sr} + u_\mathrm{b}^*\omega_\mathrm{br} + \omega_\mathrm{r}\right) \tag{S37a}$$

$$\dot{\psi}_\mathrm{k}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi_\mathrm{k}}{x_\mathrm{T}}\left(u_\mathrm{s}^*\eta_\mathrm{sr} + u_\mathrm{b}^*\eta_\mathrm{br} + \eta_\mathrm{r}\right). \tag{S37b}$$

From (S27), taking the second age derivatives for the costates and noting that $\dot{\psi}_i = \dot{\psi}$ for $i \in \{\mathrm{b},\mathrm{r},\mathrm{s}\}$, we find

$$\ddot{\lambda}_\mathrm{b} = -\left(\phi\dot{\psi} + \dot{\phi}\psi_\mathrm{b} + \dot{\lambda}_\mathrm{k} d_1 B_\mathrm{b}\right) \tag{S38a}$$

$$\ddot{\lambda}_\mathrm{r} = -\left[\phi\dot{\psi} + \dot{\phi}\psi_\mathrm{r} - f_0\mu\mathrm{e}^{-\mu a}\right] \tag{S38b}$$

$$\ddot{\lambda}_\mathrm{s} = -\left(\phi\dot{\psi} + \dot{\phi}\psi_\mathrm{s}\right) \tag{S38c}$$

$$\ddot{\lambda}_\mathrm{k} = -\left(\phi\dot{\psi}_\mathrm{k} + \dot{\phi}\psi_\mathrm{k} - \dot{\lambda}_\mathrm{k} d_2\right). \tag{S38d}$$

4.1 Singular controls for regime BS: $\sigma_\mathrm{s} > 0$, $\sigma_\mathrm{b} > 0$, and $\sigma_\mathrm{s} = \sigma_\mathrm{b}$

We now obtain the singular controls for growth regime BS. The procedure is essentially the same for growth regimes BR, RS, and BRS.

For regime BS, we have the singular arc where $(u_\mathrm{b}^*, u_\mathrm{s}^*) = (\hat{u}_\mathrm{b}, 1-\hat{u}_\mathrm{b})$ and $\sigma_\mathrm{s} = \sigma_\mathrm{b}$. Hence, from (S22a), during regime BS the variable ϕ in the Hamiltonian (S21) is no longer an explicit function of the controls:

$$\begin{aligned}\phi(\boldsymbol{\lambda}) &= (1-\hat{u}_\mathrm{b})\sigma_\mathrm{s} + \hat{u}_\mathrm{b}\sigma_\mathrm{s} + e_\mathrm{r}\lambda_\mathrm{r} \\ &= \sigma_\mathrm{s} + e_\mathrm{r}\lambda_\mathrm{r} \\ &= e_\mathrm{s}\lambda_\mathrm{s}.\end{aligned} \tag{S39a}$$

From (S37), we also have the simplifications

$$\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi}{x_\mathrm{T}}\left(\hat{u}_\mathrm{b}\omega_\mathrm{bs} + \omega_\mathrm{s}\right) \tag{S39b}$$

$$\dot{\psi}_\mathrm{k} = (\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi_\mathrm{k}}{x_\mathrm{T}}\left(\hat{u}_\mathrm{b}\eta_\mathrm{bs} + \eta_\mathrm{s}\right). \tag{S39c}$$

Since $\sigma_\mathrm{s} - \sigma_\mathrm{b} = 0$, we have that $\ddot{\sigma}_\mathrm{s} - \ddot{\sigma}_\mathrm{b} = 0$, which using (S25), (S38), and (S39) becomes

$$e_\mathrm{s}\ddot{\lambda}_\mathrm{s} - e_\mathrm{b}\ddot{\lambda}_\mathrm{b} - d_1\ddot{\lambda}_\mathrm{k} = 0 \tag{S40a}$$

$$-e_\mathrm{s}\left(\phi\dot{\psi} + \dot{\phi}\psi_\mathrm{s}\right) + e_\mathrm{b}\left(\phi\dot{\psi} + \dot{\phi}\psi_\mathrm{b} + \dot{\lambda}_\mathrm{k} d_1 B_\mathrm{b}\right) + d_1\left(\phi\dot{\psi}_\mathrm{k} + \dot{\phi}\psi_\mathrm{k} - \dot{\lambda}_\mathrm{k} d_2\right) = 0 \tag{S40b}$$

$$\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a)\phi(\boldsymbol{\lambda})(e_\mathrm{b} - e_\mathrm{s}) + \dot{\psi}_\mathrm{k}(\mathbf{u}^*,\mathbf{x}^*,a)\phi(\boldsymbol{\lambda})d_1 + \rho_\mathrm{sk}(\mathbf{x}^*,\boldsymbol{\lambda},a) = 0, \tag{S40c}$$

where

$$\rho_{\mathrm{sk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \dot{\lambda}_{\mathrm{k}} d_1 \left(e_{\mathrm{b}} B_{\mathrm{b}} - d_2\right) - \dot{\phi}\left(e_{\mathrm{s}}\psi_{\mathrm{s}} - e_{\mathrm{b}}\psi_{\mathrm{b}} - d_1\psi_{\mathrm{k}}\right) \tag{S41a}$$

$$= d_1 d_2 \lambda_{\mathrm{k}}(e_{\mathrm{b}} B_{\mathrm{b}} - d_2) + e_{\mathrm{s}}\lambda_{\mathrm{s}}\left[e_{\mathrm{s}}\psi_{\mathrm{s}}(e_{\mathrm{s}}\psi_{\mathrm{s}} - e_{\mathrm{b}}\psi_{\mathrm{b}} - d_1\psi_{\mathrm{k}}) - d_1\psi_{\mathrm{k}}(e_{\mathrm{b}}B_{\mathrm{b}} - d_2)\right]. \tag{S41b}$$

Here $\dot{\lambda}_{\mathrm{k}}$ during the singular arc BS is similarly not an explicit function of the controls.

In (S40c), only $\dot{\psi}$ and $\dot{\psi}_{\mathrm{k}}$ are functions of $\mathbf{u}^*$. Expanding these terms in (S40c), we obtain an affine equation in the singular control $\hat{u}_{\mathrm{b}}$:

$$\left[\frac{\psi}{x_{\mathrm{T}}}\left(\hat{u}_{\mathrm{b}}\omega_{\mathrm{bs}} + \omega_{\mathrm{s}}\right)\right]\phi(e_{\mathrm{b}} - e_{\mathrm{s}}) + \left[\frac{\psi_{\mathrm{k}}}{x_{\mathrm{T}}}\left(\hat{u}_{\mathrm{b}}\eta_{\mathrm{bs}} + \eta_{\mathrm{s}}\right)\right]\phi d_1 + \rho_{\mathrm{sk}} = 0 \tag{S42a}$$

$$-\hat{u}_{\mathrm{b}}\zeta_{\mathrm{bsk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) + \zeta_{\mathrm{sk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = 0, \tag{S42b}$$

where

$$\zeta_{\mathrm{bsk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \frac{\phi}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{bs}}(e_{\mathrm{s}} - e_{\mathrm{b}}) - \psi_{\mathrm{k}}\eta_{\mathrm{bs}}d_1\right]$$

$$= \frac{e_{\mathrm{s}}\lambda_{\mathrm{s}}}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{bs}}(e_{\mathrm{s}} - e_{\mathrm{b}}) - \psi_{\mathrm{k}}\eta_{\mathrm{bs}}d_1\right] \tag{S43a}$$

$$\zeta_{\mathrm{sk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \rho_{\mathrm{sk}} - \frac{\phi}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{s}}(e_{\mathrm{s}} - e_{\mathrm{b}}) - \psi_{\mathrm{k}}\eta_{\mathrm{s}}d_1\right]$$

$$= \rho_{\mathrm{sk}} - \frac{e_{\mathrm{s}}\lambda_{\mathrm{s}}}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{s}}(e_{\mathrm{s}} - e_{\mathrm{b}}) - \psi_{\mathrm{k}}\eta_{\mathrm{s}}d_1\right]. \tag{S43b}$$

Therefore, assuming that $\zeta_{\mathrm{bsk}} \neq 0$, the singular control for regime BS is

$$\hat{u}_{\mathrm{b}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \frac{\zeta_{\mathrm{sk}}}{\zeta_{\mathrm{bsk}}}. \tag{S44}$$

4.2 Singular controls for regime BR: $\sigma_{\mathrm{s}} < 0$ and $\sigma_{\mathrm{b}} = 0$

For regime BR, we have that $(u_{\mathrm{b}}^*, u_{\mathrm{s}}^*) = (\hat{u}_{\mathrm{b}}, 0)$. Hence, from (S22a), during regime BR the variable ϕ is no longer an explicit function of the controls:

$$\phi(\boldsymbol{\lambda}) = 0 \times \sigma_{\mathrm{s}} + \hat{u}_b \times 0 + e_{\mathrm{r}}\lambda_{\mathrm{r}}$$

$$= e_{\mathrm{r}}\lambda_{\mathrm{r}}. \tag{S45a}$$

From (S37), we have the simplifications

$$\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{b}}\omega_{\mathrm{br}} + \omega_{\mathrm{r}}) \tag{S45b}$$

$$\dot{\psi}_{\mathrm{k}}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi_{\mathrm{k}}}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{b}}\eta_{\mathrm{br}} + \eta_{\mathrm{r}}). \tag{S45c}$$

From $\sigma_{\mathrm{b}} = 0$, we have that $\ddot{\sigma}_{\mathrm{b}} = 0$, which becomes

$$e_{\mathrm{b}}\ddot{\lambda}_{\mathrm{b}} - e_{\mathrm{r}}\ddot{\lambda}_{\mathrm{r}} + d_1\ddot{\lambda}_{\mathrm{k}} = 0 \tag{S46a}$$

$$-e_{\mathrm{b}}\left(\phi\dot{\psi} + \dot{\phi}\psi_{\mathrm{b}} + \dot{\lambda}_{\mathrm{k}}d_1 B_{\mathrm{b}}\right) + e_{\mathrm{r}}\left(\phi\dot{\psi} + \dot{\phi}\psi_{\mathrm{r}} - f_0\mu\mathrm{e}^{-\mu a}\right) - d_1\left(\phi\dot{\psi}_{\mathrm{k}} + \dot{\phi}\psi_{\mathrm{k}} - \dot{\lambda}_{\mathrm{k}}d_2\right) = 0 \tag{S46b}$$

$$-\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a)\phi(\boldsymbol{\lambda})(e_{\mathrm{b}} - e_{\mathrm{r}}) - \dot{\psi}_{\mathrm{k}}(\mathbf{u}^*,\mathbf{x}^*,a)\phi(\boldsymbol{\lambda})d_1 + \rho_{\mathrm{rk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = 0, \tag{S46c}$$

where

$$\rho_{\mathrm{rk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = -\dot{\lambda}_{\mathrm{k}} d_1 \left(e_{\mathrm{b}} B_{\mathrm{b}} - d_2\right) + \dot{\phi}\left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{b}}\psi_{\mathrm{b}} - d_1\psi_{\mathrm{k}}\right) - e_{\mathrm{r}} f_0 \mu \mathrm{e}^{-\mu a} \tag{S47a}$$

$$= -d_1 d_2 \lambda_{\mathrm{k}}\left(e_{\mathrm{b}} B_{\mathrm{b}} - d_2\right) - e_{\mathrm{r}}\lambda_{\mathrm{r}}\left[e_{\mathrm{r}}\psi_{\mathrm{r}}\left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{b}}\psi_{\mathrm{b}} - d_1\psi_{\mathrm{k}}\right) - d_1\psi_{\mathrm{k}}\left(e_{\mathrm{b}}B_{\mathrm{b}} - d_2\right)\right]$$

$$- e_{\mathrm{r}} f_0 \mathrm{e}^{-\mu a}\left[\mu + \left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{b}}\psi_{\mathrm{b}} - d_1\psi_{\mathrm{k}}\right)\right]. \tag{S47b}$$

Again, in (S46c), only $\dot{\psi}$ and $\dot{\psi}_{\mathrm{k}}$ are functions of $\mathbf{u}^*$. Expanding these terms in (S46c), we similarly obtain an affine equation in the singular control $\hat{u}_{\mathrm{b}}$:

$$-\left[\frac{\psi}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{b}}\omega_{\mathrm{br}} + \omega_{\mathrm{r}})\right]\phi(e_{\mathrm{b}} - e_{\mathrm{r}}) - \left[\frac{\psi_{\mathrm{k}}}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{b}}\eta_{\mathrm{br}} + \eta_{\mathrm{r}})\right]\phi d_1 + \rho_{\mathrm{rk}} = 0 \tag{S48a}$$

$$-\hat{u}_{\mathrm{b}}\zeta_{\mathrm{brk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) + \zeta_{\mathrm{rk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = 0, \tag{S48b}$$

where

$$\zeta_{\mathrm{brk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \frac{\phi}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{br}}(e_{\mathrm{b}} - e_{\mathrm{r}}) + \psi_{\mathrm{k}}\eta_{\mathrm{br}} d_1\right]$$

$$= \frac{e_{\mathrm{r}}\lambda_{\mathrm{r}}}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{br}}(e_{\mathrm{b}} - e_{\mathrm{r}}) + \psi_{\mathrm{k}}\eta_{\mathrm{br}} d_1\right] \tag{S49a}$$

$$\zeta_{\mathrm{rk}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \rho_{\mathrm{rk}} - \frac{\phi}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{r}}(e_{\mathrm{b}} - e_{\mathrm{r}}) + \psi_{\mathrm{k}}\eta_{\mathrm{r}} d_1\right]$$

$$= \rho_{\mathrm{rk}} - \frac{e_{\mathrm{r}}\lambda_{\mathrm{r}}}{x_{\mathrm{T}}}\left[\psi\omega_{\mathrm{r}}(e_{\mathrm{b}} - e_{\mathrm{r}}) + \psi_{\mathrm{k}}\eta_{\mathrm{r}} d_1\right]. \tag{S49b}$$

Therefore, assuming that $\zeta_{\mathrm{brk}} \neq 0$, the singular control for regime BR is

$$\hat{u}_{\mathrm{b}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \frac{\zeta_{\mathrm{rk}}}{\zeta_{\mathrm{brk}}}. \tag{S50}$$

4.3 Singular controls for regime RS: $\sigma_{\mathrm{s}} = 0$ and $\sigma_{\mathrm{b}} < 0$

For regime RS, we have that $(u_{\mathrm{b}}^*, u_{\mathrm{s}}^*) = (0, \hat{u}_{\mathrm{s}})$. Hence, during regime RS the variable ϕ is again no longer an explicit function of the controls:

$$\phi(\boldsymbol{\lambda}) \equiv \hat{u}_{\mathrm{s}} \times 0 + 0 \times \sigma_{\mathrm{b}} + e_{\mathrm{r}}\lambda_{\mathrm{r}}$$

$$= e_{\mathrm{r}}\lambda_{\mathrm{r}}. \tag{S51a}$$

We have the simplifications

$$\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{s}}\omega_{\mathrm{sr}} + \omega_{\mathrm{r}}) \tag{S51b}$$

$$\dot{\psi}_{\mathrm{k}}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi_{\mathrm{k}}}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{s}}\eta_{\mathrm{sr}} + \eta_{\mathrm{r}}). \tag{S51c}$$

From $\sigma_{\mathrm{s}} = 0$, we have that $\ddot{\sigma}_{\mathrm{s}} = 0$, which becomes

$$e_{\mathrm{s}}\ddot{\lambda}_{\mathrm{s}} - e_{\mathrm{r}}\ddot{\lambda}_{\mathrm{r}} = 0 \tag{S52a}$$

$$-e_{\mathrm{s}}\left(\phi\dot{\psi} + \dot{\phi}\psi_{\mathrm{s}}\right) + e_{\mathrm{r}}\left(\phi\dot{\psi} + \dot{\phi}\psi_{\mathrm{r}} - f_0\mu\mathrm{e}^{-\mu a}\right) = 0 \tag{S52b}$$

$$-\dot{\psi}(\mathbf{u},\mathbf{x},a)\phi(\boldsymbol{\lambda})(e_{\mathrm{s}} - e_{\mathrm{r}}) + \rho_{\mathrm{rs}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = 0, \tag{S52c}$$

where

$$\begin{aligned}\rho_{\mathrm{rs}}(\mathbf{x}^*,\boldsymbol{\lambda},a) &= \dot{\phi}\left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{s}}\psi_{\mathrm{s}}\right) - e_{\mathrm{r}} f_0 \mu \mathrm{e}^{-\mu a} \\ &= -e_{\mathrm{r}}\lambda_{\mathrm{r}}\left[e_{\mathrm{r}}\psi_{\mathrm{r}}\left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{s}}\psi_{\mathrm{s}}\right)\right] - e_{\mathrm{r}} f_0 \mathrm{e}^{-\mu a}\left[\mu + \left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{s}}\psi_{\mathrm{s}}\right)\right]. \end{aligned} \tag{S53}$$

Once again, only $\dot{\psi}$ is a function of $\mathbf{u}^*$ in (S52c). Expanding this term in (S52c), we obtain an affine equation in the singular control $\hat{u}_{\mathrm{s}}$:

$$-\left[\frac{\psi}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{s}}\omega_{\mathrm{sr}} + \omega_{\mathrm{r}})\right]\phi(e_{\mathrm{s}} - e_{\mathrm{r}}) + \rho_{\mathrm{r}} = 0 \tag{S54a}$$

$$-\hat{u}_{\mathrm{s}}\zeta_{\mathrm{sr}}(\mathbf{x}^*,\boldsymbol{\lambda},a) + \zeta_{\mathrm{r}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = 0, \tag{S54b}$$

where we define

$$\begin{aligned}\zeta_{\mathrm{sr}}(\mathbf{x}^*,\boldsymbol{\lambda},a) &= \phi\frac{\psi}{x_{\mathrm{T}}}\omega_{\mathrm{sr}}(e_{\mathrm{s}} - e_{\mathrm{r}}) \\ &= \frac{e_{\mathrm{r}}\lambda_{\mathrm{r}}}{x_{\mathrm{T}}}\psi\omega_{\mathrm{sr}}(e_{\mathrm{s}} - e_{\mathrm{r}}) \end{aligned} \tag{S55a}$$

$$\begin{aligned}\zeta_{\mathrm{r}}(\mathbf{x}^*,\boldsymbol{\lambda},a) &= \rho_{\mathrm{rs}} - \phi\frac{\psi}{x_{\mathrm{T}}}\omega_{\mathrm{r}}(e_{\mathrm{s}} - e_{\mathrm{r}}) \\ &= \rho_{\mathrm{rs}} - \frac{e_{\mathrm{r}}\lambda_{\mathrm{r}}}{x_{\mathrm{T}}}\psi\omega_{\mathrm{r}}(e_{\mathrm{s}} - e_{\mathrm{r}}). \end{aligned} \tag{S55b}$$

Therefore, assuming that $\zeta_{\mathrm{sr}} \neq 0$, the singular control for regime RS is

$$\hat{u}_{\mathrm{s}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = \frac{\zeta_{\mathrm{r}}}{\zeta_{\mathrm{sr}}}. \tag{S56}$$

4.4 Singular controls for regime BRS: $\sigma_{\mathrm{s}} = \sigma_{\mathrm{b}} = 0$

For regime BRS, we have that $(u_{\mathrm{b}}^*, u_{\mathrm{s}}^*) = (\hat{u}_{\mathrm{b}}, \hat{u}_{\mathrm{s}})$. As before, the variable ϕ is no longer an explicit function of the controls:

$$\begin{aligned}\phi(\boldsymbol{\lambda}) &= \hat{u}_{\mathrm{s}} \times 0 + \hat{u}_{\mathrm{b}} \times 0 + e_{\mathrm{r}}\lambda_{\mathrm{r}} \\ &= e_{\mathrm{r}}\lambda_{\mathrm{r}}. \end{aligned} \tag{S57a}$$

Similarly, we have the simplifications

$$\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi}{x_{\mathrm{T}}}(\hat{u}_{\mathrm{s}}\omega_{\mathrm{sr}} + \hat{u}_{\mathrm{b}}\omega_{\mathrm{br}} + \omega_{\mathrm{r}}) \tag{S57b}$$

$$\dot{\psi}_{\mathrm{k}}(\mathbf{u}^*,\mathbf{x}^*,a) = \frac{\psi_{\mathrm{k}}}{x_{\mathrm{T}}}\left(\hat{u}_{\mathrm{s}}\eta_{\mathrm{sr}} + \hat{u}_{\mathrm{b}}\eta_{\mathrm{br}} + \eta_{\mathrm{r}}\right). \tag{S57c}$$

From $\sigma_{\mathrm{s}} = 0$, we have that $\ddot{\sigma}_{\mathrm{s}} = 0$, which is

$$e_{\mathrm{s}}\ddot{\lambda}_{\mathrm{s}} - e_{\mathrm{r}}\ddot{\lambda}_{\mathrm{r}} = 0 \tag{S58a}$$

$$-e_{\mathrm{s}}\left(\phi\dot{\psi} + \dot{\phi}\psi_{\mathrm{s}}\right) + e_{\mathrm{r}}\left[\phi\dot{\psi} + \dot{\phi}\psi_{\mathrm{r}} - f_0\mu\mathrm{e}^{-\mu a}\right] = 0 \tag{S58b}$$

$$\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a)\phi(\boldsymbol{\lambda})(e_{\mathrm{r}} - e_{\mathrm{s}}) + \rho_{\mathrm{rs}}(\mathbf{x}^*,\boldsymbol{\lambda},a) = 0, \tag{S58c}$$

where as before

$$\begin{aligned}\rho_{\mathrm{rs}}(\mathbf{x}^*,\boldsymbol{\lambda},a) &= \dot{\phi}\left(e_{\mathrm{r}}\psi_r - e_{\mathrm{s}}\psi_{\mathrm{s}}\right) - e_{\mathrm{r}} f_0 \mu \mathrm{e}^{-\mu a} \\ &= -e_{\mathrm{r}}\lambda_{\mathrm{r}}\left[e_{\mathrm{r}}\psi_{\mathrm{r}}\left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{s}}\psi_{\mathrm{s}}\right)\right] - e_{\mathrm{r}} f_0 \mathrm{e}^{-\mu a}\left[\mu + \left(e_{\mathrm{r}}\psi_{\mathrm{r}} - e_{\mathrm{s}}\psi_{\mathrm{s}}\right)\right]. \end{aligned} \tag{S59}$$

291 Expanding $\dot{\psi}$ in (S58c), we obtain an affine equation in the two controls $\hat{u}_\mathrm{s}$ and $\hat{u}_\mathrm{b}$:

$$\left[\frac{\psi}{x_\mathrm{T}}\left(\hat{u}_\mathrm{s}\omega_\mathrm{sr}+\hat{u}_\mathrm{b}\omega_\mathrm{br}+\omega_\mathrm{r}\right)\right]\phi(e_\mathrm{r}-e_\mathrm{s})+\rho_\mathrm{rs}=0 \tag{S60a}$$

$$-\hat{u}_\mathrm{s}\zeta_\mathrm{sr}(\mathbf{x}^*,\boldsymbol{\lambda},a)-\hat{u}_\mathrm{b}\zeta_\mathrm{br}(\mathbf{x}^*,\boldsymbol{\lambda},a)+\zeta_\mathrm{r}(\mathbf{x}^*,\boldsymbol{\lambda},a)=0, \tag{S60b}$$

292 where

$$\begin{aligned}\zeta_\mathrm{sr}(\mathbf{x}^*,\boldsymbol{\lambda},a)&=-\phi\frac{\psi}{x_\mathrm{T}}\omega_\mathrm{sr}(e_\mathrm{r}-e_\mathrm{s})\\&=\frac{e_\mathrm{r}\lambda_\mathrm{r}}{x_\mathrm{T}}\psi\omega_\mathrm{sr}(e_\mathrm{s}-e_\mathrm{r})\end{aligned} \tag{S61a}$$

$$\begin{aligned}\zeta_\mathrm{br}(\mathbf{x}^*,\boldsymbol{\lambda},a)&=-\phi\frac{\psi}{x_\mathrm{T}}\omega_\mathrm{br}(e_\mathrm{r}-e_\mathrm{s})\\&=\frac{e_\mathrm{r}\lambda_\mathrm{r}}{x_\mathrm{T}}\psi\omega_\mathrm{br}(e_\mathrm{s}-e_\mathrm{r})\end{aligned} \tag{S61b}$$

$$\begin{aligned}\zeta_\mathrm{r}(\mathbf{x}^*,\boldsymbol{\lambda},a)&=\rho_\mathrm{rs}+\phi\frac{\psi}{x_\mathrm{T}}\omega_\mathrm{r}(e_\mathrm{r}-e_\mathrm{s})\\&=\rho_\mathrm{rs}-\frac{e_\mathrm{r}\lambda_\mathrm{r}}{x_\mathrm{T}}\psi\omega_\mathrm{r}(e_\mathrm{s}-e_\mathrm{r}).\end{aligned} \tag{S61c}$$

293 Now, from $\sigma_\mathrm{b}=0$, we have that $\ddot{\sigma}_\mathrm{b}=0$, which is

$$e_\mathrm{b}\lambda_\mathrm{b}-e_\mathrm{r}\lambda_\mathrm{r}+d_1\lambda_\mathrm{k}=0 \tag{S62a}$$

$$-e_\mathrm{b}\left(\phi\dot{\psi}+\dot{\phi}\psi_\mathrm{b}+\dot{\lambda_\mathrm{k}}d_1B_\mathrm{b}\right)+e_\mathrm{r}\left[\phi\dot{\psi}+\dot{\phi}\psi_\mathrm{r}-\mu\mathrm{e}^{-\mu a}\right]-d_1\left(\phi\dot{\psi}_\mathrm{k}+\dot{\phi}\psi_\mathrm{k}-\dot{\lambda_\mathrm{k}}d_2\right)=0 \tag{S62b}$$

$$-\dot{\psi}(\mathbf{u}^*,\mathbf{x}^*,a)\phi(\boldsymbol{\lambda})(e_\mathrm{b}-e_\mathrm{r})-\dot{\psi}_\mathrm{k}(\mathbf{u}^*,\mathbf{x}^*,a)\phi(\boldsymbol{\lambda})d_1+\rho_\mathrm{rk}(\mathbf{x}^*,\boldsymbol{\lambda},a)=0, \tag{S62c}$$

294 where as before

$$\rho_\mathrm{rk}(\mathbf{x}^*,\boldsymbol{\lambda},a)=-\dot{\lambda}_\mathrm{k}d_1\left(e_\mathrm{b}B_\mathrm{b}-d_2\right)+\dot{\phi}\left(e_\mathrm{r}\psi_r-e_\mathrm{b}\psi_\mathrm{b}-d_1\psi_\mathrm{k}\right)-e_\mathrm{r}f_0\mu\mathrm{e}^{-\mu a} \tag{S63a}$$

$$\begin{aligned}&=-d_1d_2\lambda_\mathrm{k}\left(e_\mathrm{b}B_\mathrm{b}-d_2\right)-e_\mathrm{r}\lambda_\mathrm{r}\left[e_\mathrm{r}\psi_\mathrm{r}\left(e_\mathrm{r}\psi_\mathrm{r}-e_\mathrm{b}\psi_\mathrm{b}-d_1\psi_\mathrm{k}\right)-d_1\psi_\mathrm{k}\left(e_\mathrm{b}B_\mathrm{b}-d_2\right)\right]\\&\quad-e_\mathrm{r}f_0\mathrm{e}^{-\mu a}\left[\mu+\left(e_\mathrm{r}\psi_\mathrm{r}-e_\mathrm{b}\psi_\mathrm{b}-d_1\psi_\mathrm{k}\right)\right].\end{aligned} \tag{S63b}$$

295 Expanding $\dot{\psi}$ and $\dot{\psi}_\mathrm{k}$ in (S62c), we obtain another affine equation in the two controls $\hat{u}_\mathrm{s}$ and $\hat{u}_\mathrm{b}$:

$$-\left[\frac{\psi}{x_\mathrm{T}}\left(\hat{u}_\mathrm{s}\omega_\mathrm{sr}+\hat{u}_\mathrm{b}\omega_\mathrm{br}+\omega_\mathrm{r}\right)\right]\phi(e_\mathrm{b}-e_\mathrm{r})-\left[\frac{\psi_\mathrm{k}}{x_\mathrm{T}}\left(\hat{u}_\mathrm{s}\eta_\mathrm{sr}+\hat{u}_\mathrm{b}\eta_\mathrm{br}+\eta_\mathrm{r}\right)\right]\phi d_1+\rho_\mathrm{rk}=0 \tag{S64a}$$

$$-\hat{u}_\mathrm{s}\zeta_\mathrm{srk}(\mathbf{x}^*,\boldsymbol{\lambda},a)-\hat{u}_\mathrm{b}\zeta_\mathrm{brk}(\mathbf{x}^*,\boldsymbol{\lambda},a)+\zeta_\mathrm{rk}(\mathbf{x}^*,\boldsymbol{\lambda},a)=0, \tag{S64b}$$

296 where

$$\begin{aligned}\zeta_\mathrm{srk}(\mathbf{x}^*,\boldsymbol{\lambda},a)&=\frac{\phi}{x_\mathrm{T}}\left[\psi\omega_\mathrm{sr}(e_\mathrm{b}-e_\mathrm{r})+\psi_\mathrm{k}\eta_\mathrm{sr}d_1\right]\\&=\frac{e_\mathrm{r}\lambda_\mathrm{r}}{x_\mathrm{T}}\left[\psi\omega_\mathrm{sr}(e_\mathrm{b}-e_\mathrm{r})+\psi_\mathrm{k}\eta_\mathrm{sr}d_1\right]\end{aligned} \tag{S65a}$$

$$\begin{aligned}\zeta_\mathrm{brk}(\mathbf{x}^*,\boldsymbol{\lambda},a)&=\frac{\phi}{x_\mathrm{T}}\left[\psi\omega_\mathrm{br}(e_\mathrm{b}-e_\mathrm{r})+\psi_\mathrm{k}\eta_\mathrm{br}d_1\right]\\&=\frac{e_\mathrm{r}\lambda_\mathrm{r}}{x_\mathrm{T}}\left[\psi\omega_\mathrm{br}(e_\mathrm{b}-e_\mathrm{r})+\psi_\mathrm{k}\eta_\mathrm{br}d_1\right]\end{aligned} \tag{S65b}$$

$$\begin{aligned}\zeta_\mathrm{rk}(\mathbf{x}^*,\boldsymbol{\lambda},a)&=\rho_\mathrm{rk}-\frac{\phi}{x_\mathrm{T}}\left[\psi\omega_\mathrm{r}(e_\mathrm{b}-e_\mathrm{r})+\psi_\mathrm{k}\eta_\mathrm{r}d_1\right]\\&=\rho_\mathrm{rk}-\frac{e_\mathrm{r}\lambda_\mathrm{r}}{x_\mathrm{T}}\left[\psi\omega_\mathrm{r}(e_\mathrm{b}-e_\mathrm{r})+\psi_\mathrm{k}\eta_\mathrm{r}d_1\right].\end{aligned} \tag{S65c}$$

297 Therefore, solving (S60b) and (S64b) and assuming that $\zeta_{\mathrm{sr}}\zeta_{\mathrm{brk}} - \zeta_{\mathrm{br}}\zeta_{\mathrm{srk}} \neq 0$, the singular controls for regime

298 BRS are

$$\hat{u}_{\mathrm{s}}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{\zeta_{\mathrm{r}}\zeta_{\mathrm{brk}} - \zeta_{\mathrm{br}}\zeta_{\mathrm{rk}}}{\zeta_{\mathrm{sr}}\zeta_{\mathrm{brk}} - \zeta_{\mathrm{br}}\zeta_{\mathrm{srk}}} \tag{S66a}$$

$$\hat{u}_{\mathrm{b}}(\mathbf{x}^*, \boldsymbol{\lambda}, a) = \frac{\zeta_{\mathrm{sr}}\zeta_{\mathrm{rk}} - \zeta_{\mathrm{r}}\zeta_{\mathrm{srk}}}{\zeta_{\mathrm{sr}}\zeta_{\mathrm{brk}} - \zeta_{\mathrm{br}}\zeta_{\mathrm{srk}}}. \tag{S66b}$$

5 Parameter values

Here we summarize the values of the 22 parameters used in numerical solutions. From these, 13 parameters are estimated as described in section 6 and they refer to newborn mass, tissue metabolism, and demography (Table S2). The estimates of E_i for are less accurate than those of B_i for $i \in \{\mathrm{b},\mathrm{s},\mathrm{r}\}$ as they require stronger assumptions given the available data (see Moses *et al.* (2008)). Since the parameter f_0 only displaces the objective vertically and thus has no effect on the solution, we choose its value to scale the objective R_0 (Table S2). The remaining 8 parameters refer to skill metabolism, contest success, and (allo)parental care, for which we use values that produce body and brain mass that closely approach ontogenetic modern human data. Hence, we use different benchmark values with either power (Table S3) or exponential (Table S4) competence.

Newborn mass		Tissue metabolism				Demography	
		K	$132.7281\,\frac{\mathrm{MJ}}{\mathrm{y}}\mathrm{kg}^{-\beta}$	β	0.7378		
$x_\mathrm{s}(0)$	2.0628 kg	B_s	$29.6891\,\frac{\mathrm{MJ}}{\mathrm{y}\times\mathrm{kg}}$	E_s	$12.4594\,\frac{\mathrm{MJ}}{\mathrm{kg}}$	f_0	$10\,\frac{\#\mathrm{offspring}}{\mathrm{kg}\times\mathrm{y}}$
$x_\mathrm{b}(0)$	0.3372 kg	B_b	$313.0962\,\frac{\mathrm{MJ}}{\mathrm{y}\times\mathrm{kg}}$	E_b	$123.7584\,\frac{\mathrm{MJ}}{\mathrm{kg}}$	μ	$0.034\,\frac{1}{\mathrm{y}}$
$x_\mathrm{r}(0)$	0 kg	B_r	$2697.1179\,\frac{\mathrm{MJ}}{\mathrm{y}\times\mathrm{kg}}$	E_r	$190.8196\,\frac{\mathrm{MJ}}{\mathrm{kg}}$	τ	47 y

Table S2: Estimated parameter values and f_0, which is set to an arbitrary value.

For power competence:

Skill metabolism		Contest success		(Allo)parental care	
υ_k	0.5	α	$1\ \mathrm{skill}^{\gamma}$	φ_0/α	0.6
B_k	$36\,\frac{\mathrm{MJ}}{\mathrm{y}\times\mathrm{skill}}$	γ	1.4	φ_r	$0.2\,\frac{1}{\mathrm{y}}$
E_k	$370\,\frac{\mathrm{MJ}}{\mathrm{skill}}$	$x_\mathrm{k}(0)$	1 skill		

Table S3: Benchmark parameter values with power competence. The value of φ_r yields (allo)parental care for ≈ 20 years, as observed in forager-horticulturalists (Schniter *et al.*, 2015).

For exponential competence:

Skill metabolism		Contest success		(Allo)parental care	
υ_k	0.5	α	1.15	φ_0/α	0.8
B_k	$50\,\frac{\mathrm{MJ}}{\mathrm{y}\times\mathrm{skill}}$	γ	$0.6\ \mathrm{skill}^{-1}$	φ_r	$0.2\,\frac{1}{\mathrm{y}}$
E_k	$250\,\frac{\mathrm{MJ}}{\mathrm{skill}}$	$x_\mathrm{k}(0)$	0 skill		

Table S4: Benchmark parameter values with exponential competence.

6 Estimation of parameter values

Here we describe how we obtained the parameter values in Table S2. We use ontogenetic data for modern human females published in Table S2 of Kuzawa *et al.* (2014). We denote the observed mass of tissue i at age a as $X_i(a)$ and their sum as $X_T(a)$. Thus, we set $x_s(0) = X_s(0) = 2.0628$ kg and $x_b(0) = X_b(0) = 0.3372$ kg (Kuzawa *et al.*, 2014). The count of preovulatory ovarian follicles serves as a proxy for measuring female human fertility (McGee and Hsueh, 2000), so we take reproductive cells as referring to preovulatory ovarian follicle cells and set $x_r(0) = 0$ kg. We also denote by A_a the observed age at adulthood. Hence, $X_T(A_a) = 51.1$ kg and $X_b(A_a) = 1.31$ kg (Kuzawa *et al.*, 2014). We also have that $B_{rest}(A_a) = 1243.4\ \text{kcal/day} \times 4184\ \text{J/kcal} \times 365\ \text{d/y} = 1898.8707\ \text{MJ/y}$ (Kuzawa *et al.*, 2014).

6.1 Values for B_i for $i \in \{b, r, s\}$

B_b: Let $c_1(a)$ be the ratio of glucose uptake by the brain per unit time at age a divided by the resting metabolic rate at that age. Let $c_2(a)$ be the fraction of brain glucose metabolism that is oxidative. Then, the empirically estimated brain metabolic rate at age a is the product $B_{rest}(a)c_1(a)c_2(a)$. $c_1(a)$ is obtained from Table S2 of Kuzawa *et al.* (2014) and rough estimates of $c_2(a)$ are obtained from Goyal *et al.* (2014). For adults they are $c_1(A_a) = 0.24$ and $c_2(A_a) = 0.9$ (Kuzawa *et al.*, 2014; Goyal *et al.*, 2014). Hence, we let $B_b = B_{rest}(A_a)c_1(A_a)c_2(A_a)/X_b(A_a) = 313.0962$ MJ/kg/y.

B_r: We are unaware of reports of the metabolic rate of preovulatory follicles. Thus, we use the metabolic rate of a human oocyte as a proxy. The oxygen consumption by a human oocyte is estimated to be 0.53×10^{-9} l O_2/h/oocyte (Magnusson *et al.*, 1986). Oxygen consumption can be transformed into power units by multiplying by 20.1 kJ/l O_2 (Blaxter, 1989). The mass of a mouse oocyte is 34.6 ng (Abramczuk and Sawicki, 1974). Assuming that mouse and human oocyte are of similar mass, then $B_r = 0.53 \times 10^{-9} \frac{\text{l O}_2}{\text{h} \times \text{oocyte}} \times 20.1 \frac{\text{kJ}}{\text{l O}_2} \times \frac{1\ \text{oocyte}}{34.6\ \text{ng}} \times \frac{24\ \text{h}}{1\ \text{d}} \times \frac{365\ \text{d}}{1\ \text{y}} \times \frac{10^9\ \text{ng}}{1\ \text{g}} \times \frac{1000\ \text{g}}{1\ \text{kg}} \times \frac{1\ \text{MJ}}{1000\ \text{kJ}} = 2697.1179\ \text{MJ/kg/year}$.

B_s: Adult human females have on average about 2 preovulatory follicles at any given age (Dickey *et al.*, 2002). A preovulatory follicle has an average diameter of 21.1 mm (O'Herlihy *et al.*, 1980). Approximating the follicle dry mass by the dry mass of a spherical cell with such diameter and water content of 60%, then the adult mass of reproductive tissue is $X_r(A_a) = 2\ \text{follicles} \times \frac{4}{3}\pi\left(\frac{21.1\ \text{mm}}{2}\right)^3 \times \frac{1\ \text{kg H}_2\text{O}}{10^6\ \text{mm}^3\text{H}_2\text{O}} \times \frac{0.4\ \text{kg dry mass}}{1\ \text{kg H}_2\text{O}} = 3.9349 \times 10^{-3}$ kg. Hence, $X_s(A_a) = X_T(A_a) - X_b(A_a) - X_r(A_a) = 49.7861$ kg.

Since at human adulthood there is no growth, it must be the case that $B_{rest}(A_a) = B_{maint}(A_a) = \sum_{i\in\{b,r,s\}} X_i(A_a)B_i$. Because we have that $B_{rest}(A_a) = 1898.8707$ MJ/y, it follows that $B_s = [B_{rest}(A_a) - B_b X_b(A_a) - B_r X_r(A_a)]/X_s(A_a) = 29.6891$ MJ/kg/y.

6.2 Values for E_i for $i \in \{b, r, s\}$

E_b : We have that brain metabolic rate is $M_{brain}(a) = X_b(a)B_b + \dot{X}_b(a)E_b$. Assuming that at birth most brain metabolic rate is due to brain growth, then $M_{brain}(0) \approx \dot{X}_b(0)E_b$. We also have that, $M_{brain}(0) = B_{rest}(0)c_1(0)c_2(0)$ and that $B_{rest}(0) = 166.6132$ MJ/y (Kuzawa *et al.*, 2014), $c_1(0) = 0.598$ (Kuzawa *et al.*, 2014), and $c_2(0) \approx 0.9$

(Goyal *et al.*, 2014). From the data in Kuzawa *et al.* (2014), we estimate $\dot{X}_\mathrm{b}(0) = 0.7246$ kg/y. Then, we have $E_\mathrm{b} = M_\mathrm{brain}(0)/\dot{X}_\mathrm{b}(0) = 123.7584$ MJ/kg.

E_r : We have that $B_\mathrm{syn}(a) = \sum_{i\in\{\mathrm{b,r,s}\}} \dot{X}_i(a)E_i$. We assume that shortly before adulthood most growth is reproductive. So assuming $\dot{X}_\mathrm{r}(A_\mathrm{a}-1) \neq 0$ while $\dot{X}_{i\neq\mathrm{r}}(A_\mathrm{a}-1) \approx 0$, we have that

$$E_\mathrm{r} = \frac{B_\mathrm{rest}(A_\mathrm{a}-1) - B_\mathrm{maint}(A_\mathrm{a}-1)}{\dot{X}_\mathrm{r}(A_\mathrm{a}-1)} \tag{S67a}$$

$$= \frac{B_\mathrm{rest}(A_\mathrm{a}-1) - BX_\mathrm{T}(A_\mathrm{a}-1)}{\dot{X}_\mathrm{T}(A_\mathrm{a}-1)} \tag{S67b}$$

We also have that $B_\mathrm{rest}(A_\mathrm{a}-1) = 1328.3\frac{\mathrm{kcal}}{\mathrm{d}} \times \frac{4184\,\mathrm{J}}{1\,\mathrm{kcal}} \times \frac{365\,\mathrm{d}}{1\,\mathrm{y}} = 2028.5266$ MJ/y, $X_\mathrm{T}(A_\mathrm{a}-1) = 47.4$ kg, and $\dot{X}_\mathrm{T}(A_\mathrm{a}-1) = 1.4$ kg/y (Kuzawa *et al.*, 2014). Then, $E_\mathrm{r} = 190.8196$ MJ/kg.

E_s : Again, we have that $B_\mathrm{syn}(a) = \sum_{i\in\{\mathrm{b,r,s}\}} \dot{X}_i(a)E_i$. Assuming that there is no reproductive growth at birth, then $\dot{X}_\mathrm{r}(0) = 0$ and so

$$E_\mathrm{s} = \frac{B_\mathrm{rest}(0) - B_\mathrm{maint}(0) - \dot{X}_\mathrm{b}(0)E_\mathrm{b}}{\dot{X}_\mathrm{s}(0)} \tag{S68a}$$

$$\approx \frac{B_\mathrm{rest}(0) - \dot{X}_\mathrm{b}(0)E_\mathrm{b}}{\dot{X}_\mathrm{s}(0)}, \tag{S68b}$$

assuming that at birth most resting metabolic rate is due to growth so $B_\mathrm{rest}(0) - B_\mathrm{maint}(0) \approx B_\mathrm{rest}(0)$. We have that $B_\mathrm{rest}(0) = 109.1\frac{\mathrm{kcal}}{\mathrm{d}} \times \frac{4184\,\mathrm{J}}{1\,\mathrm{kcal}} \times \frac{365\,\mathrm{d}}{1\,\mathrm{y}} = 166.6132$ MJ/y and $\dot{X}(0) = 6.9$ kg/y (Kuzawa *et al.*, 2014). Since $\dot{X}_\mathrm{s}(0) = \dot{X}(0) - \dot{X}_\mathrm{b}(0)$, then $E_\mathrm{s} = 12.4594$ MJ/kg.

6.3 Values for K and β

Using the ontogenetic (averaged) data in Table S2 of Kuzawa *et al.* (2014), where resting metabolic rate is measured in well fed individuals, we find that $B_\mathrm{rest}(a) = KX_\mathrm{T}(a)^\beta$ with $K = 132.7281\frac{\mathrm{MJ}}{\mathrm{y}}\mathrm{kg}^{-\beta}$ and $\beta = 0.7378$ ($R^2 = 0.92$) (Fig. S4).

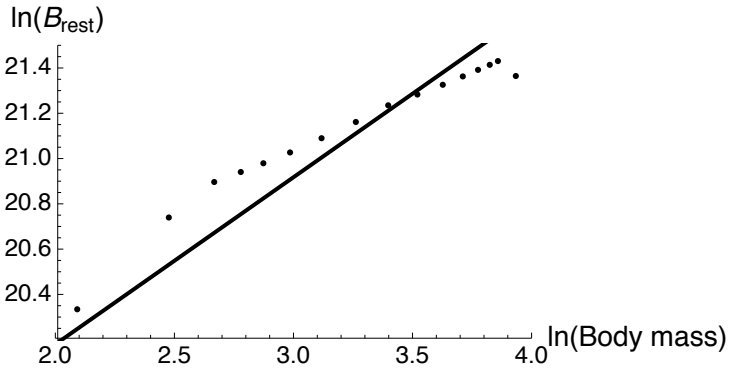


Figure S4: Power law approximation of resting metabolic rate with respect to body mass. Dots are ontogenetic values of resting metabolic rate vs. body mass in modern humans in a log-log scale (Kuzawa *et al.*, 2014). The line is the linear least square regression yielding $K = 132.7281\frac{\mathrm{MJ}}{\mathrm{y}}\mathrm{kg}^{-\beta}$ and $\beta = 0.7378$ ($R^2 = 0.92$).

6.4 Values for f_0, μ, and τ

The constant f_0 only multiplies R_0 and thus has not effect on the solution of the optimal control problem. We thus arbitrarily set it to $f_0 = 10 \frac{\#\,\text{offspring}}{\text{kg}\times\text{y}}$.

For traditional hunter-gatherers, the average life expectancy at birth is between 21 and 37 years (Gurven and Kaplan, 2007). The mid-range life expectancy is thus 29 years. With a constant mortality rate, life expectancy is $1/\mu$. We thus let $\mu = \frac{1}{29\,\text{y}} = 0.034\,\frac{1}{\text{y}}$.

For Hadza and Gainj hunter-gatherers, the average age at menopause is about 47 years (Eaton *et al.*, 1994). So, we let $\tau = 47$ years.

7 Supplementary results

7.1 Brain metabolic rate through ontogeny

With the obtained ESGS, brain metabolic rate is predicted to peak at the age of brain growth arrest, which is qualitatively consistent with recent findings for brain glucose intake (Figs. S5a,b and S6a,b; (Kuzawa *et al.*, 2014)). Brain metabolic rate and brain glucose intake are, however, not equivalent because the former refers to oxygen consumption while the latter includes non-oxidative glucose metabolism which is especially high during childhood (Kuzawa *et al.*, 2014; Goyal *et al.*, 2014). As observed with brain glucose intake (Kuzawa *et al.*, 2014), a peak in brain metabolic rate is predicted during mid childhood. The predicted small peak in brain metabolic rate results from brain growth arrest (Figs. S5b and S6b) and is enhanced by a peak in allocation to brain growth just before brain growth arrest (Figs. 1a,e). The predicted ratio of brain metabolic rate and resting metabolic rate is also qualitatively consistent with brain glucose intake in modern humans (Figs. S5c and S6c).

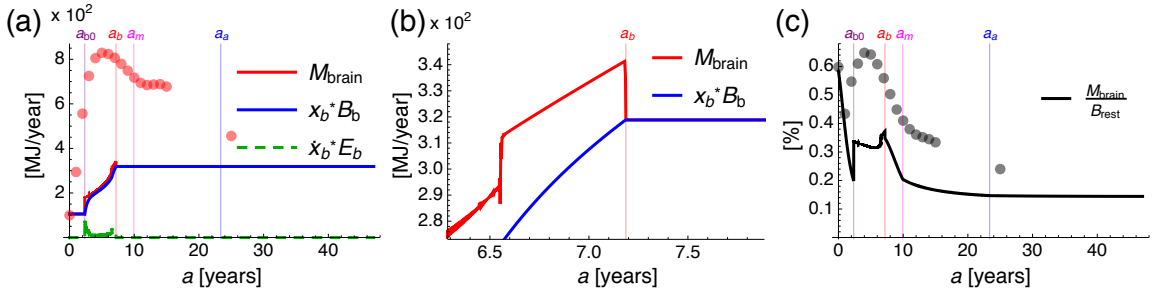


Figure S5: Predicted and observed brain metabolic patterns in humans qualitatively agree. Plots are for the scenario in Fig. 1a-d (power competence). (a) Maintenance (blue; $x_b^* B_b$), growth (green; $\dot{x}_b^* E_b$), and total (red; M_{brain}) brain metabolic rates. (b) Brain metabolic rate peaks at the age of brain growth arrest. (c) Ratio of brain metabolic rate to resting metabolic rate vs. age. Dots are (a) the energy-equivalent brain glucose intake observed in modern human females or (c) the ratio of the latter to resting metabolic rate (Kuzawa *et al.*, 2014). A similar pattern is predicted with exponential competence (Fig. S6).

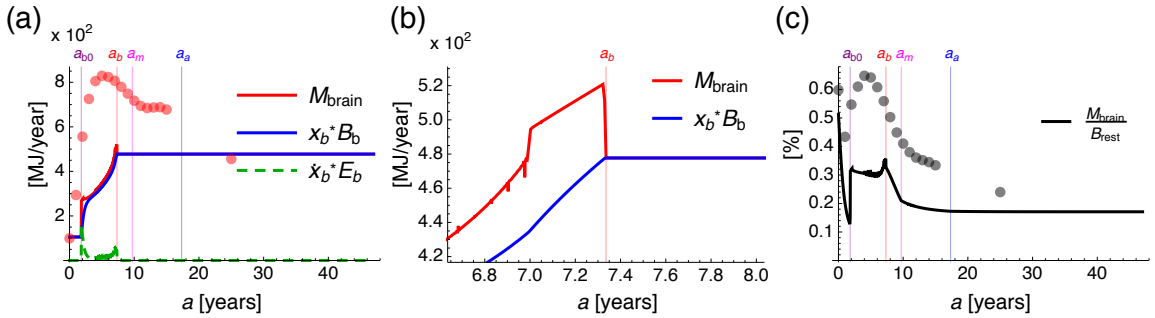


Figure S6: Predicted brain metabolic patterns with exponential competence. Plots are for the scenario in Fig. 1e-h (exponential competence). See legend of Fig. S5.

7.2 Mass of reproductive tissue

For the parameter values of Fig. 1, reproductive tissue mass remains at zero until maturity t_{m} and reaches 129 g (with power competence) or 131 g (with exponential competence) during adulthood, exceeding the 3 g we roughly estimate for human females (SI §6.1).

7.3 Effect of the absence of (allo)parental care

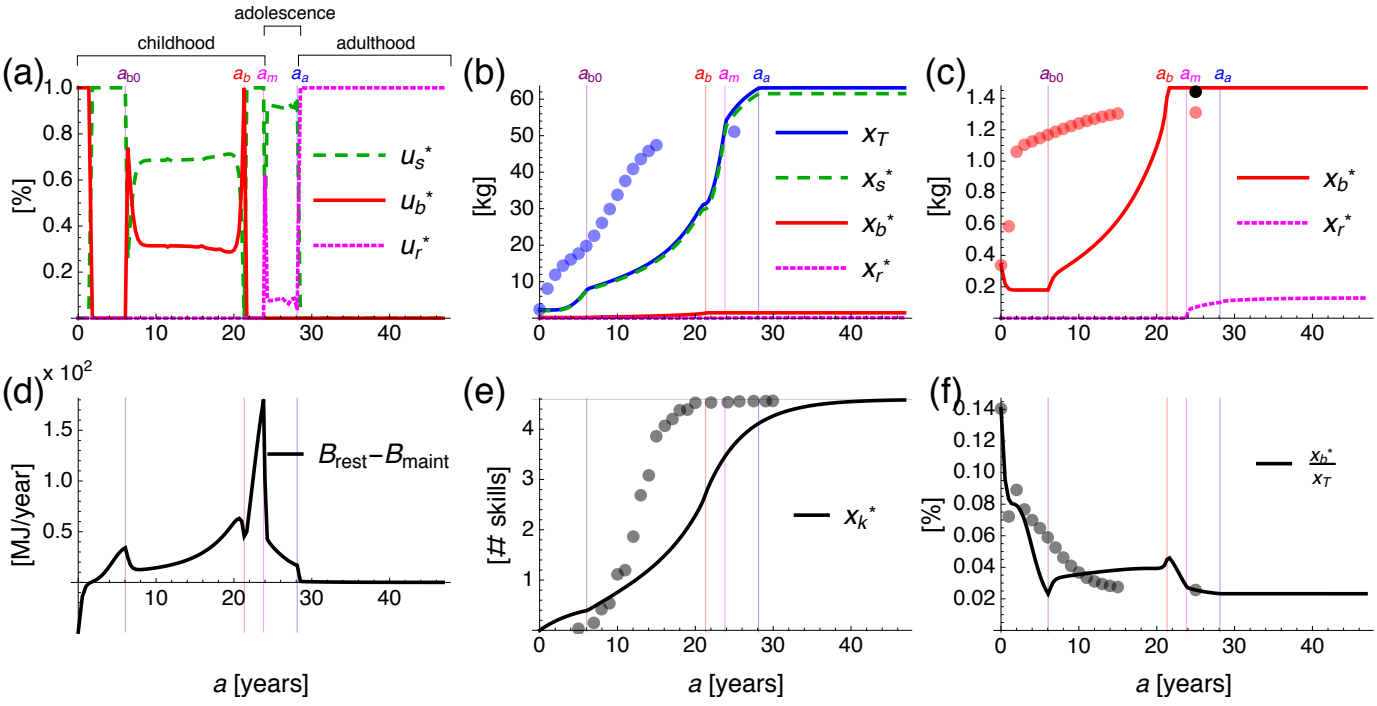


Figure S7: Effect of the absence of (allo)parental care with exponential competence. Parameters are as in Fig. 1e-h, except that here (allo)parental is absent; i.e., $\varphi_0 = 0$.

7.4 Indeterminate skill growth with inexpensive memory

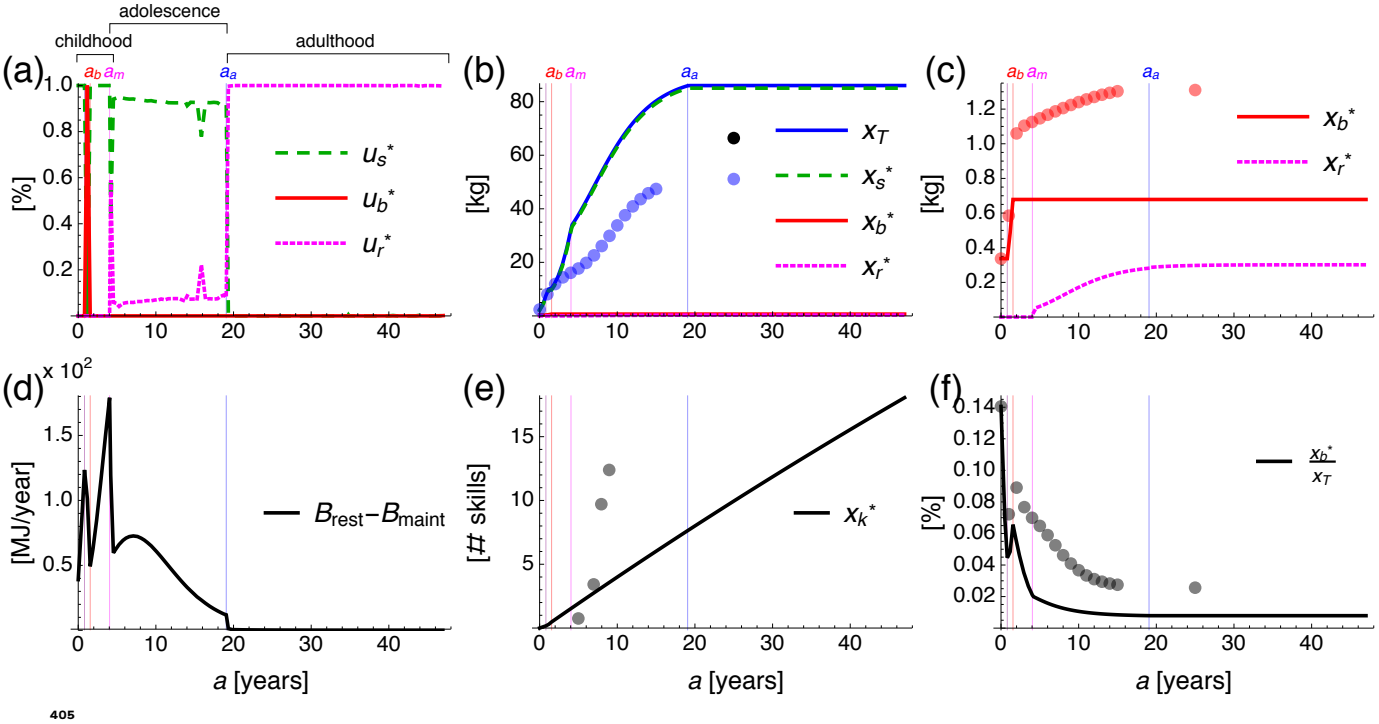


Figure S8: Indeterminate skill growth with inexpensive memory and exponential competence. Parameters are as in Fig. 1e-h, except that here $B_k = 1$ MJ/y/skill rather than $B_k = 50$ MJ/y/skill.

7.5 Large, yet inconsistent-with-data encephalization with exceedingly expensive memory

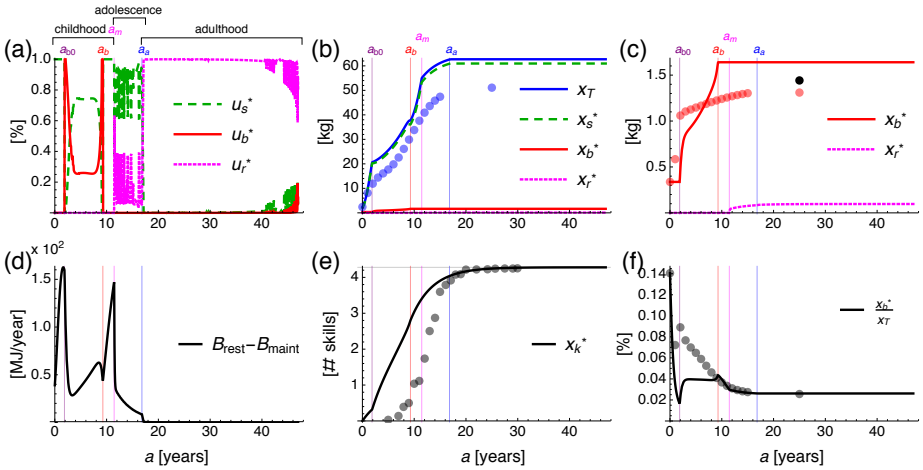


Figure S9: Larger EQ than that in Fig. 1 with exponential competence, but predicted body mass is less consistent with observation. Parameters are as in Fig. 1e-h, except that here $B_k = 60$ MJ/y/skill rather than $B_k = 50$ MJ/y/skill. Jitter in the controls indicates that the optimal control problem is computationally challenging for GPOPS (this applies to all plots in the main paper and SI).

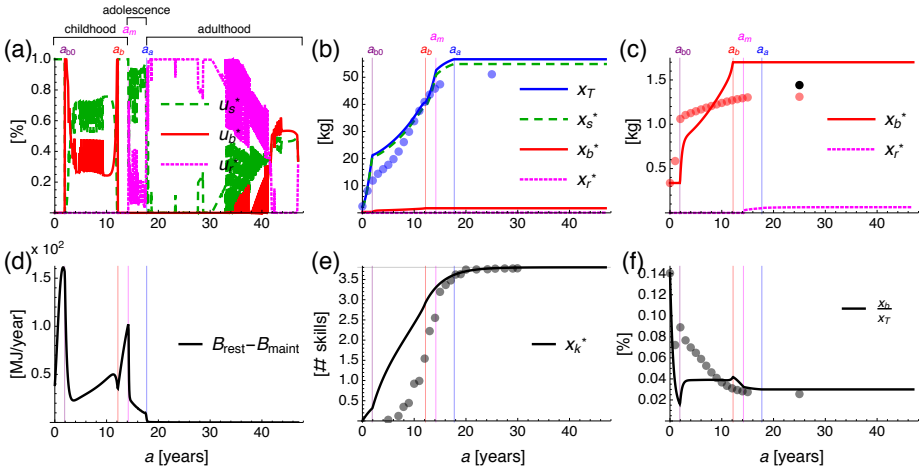


Figure S10: Larger EQ than that in Fig. 1 with exponential competence, but predicted body mass is less consistent with observation. Parameters are as in Fig. 1e-h, except that here $B_k = 70$ MJ/y/skill rather than $B_k = 50$ MJ/y/skill.

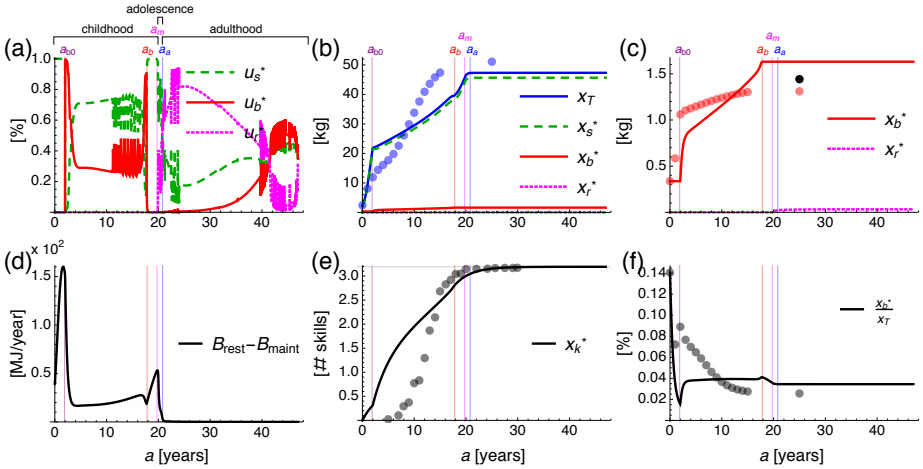


Figure S11: Larger EQ than that in Fig. 1 with exponential competence, but predicted body mass is less consistent with observation. Parameters are as in Fig. 1e-h, except that here $B_k = 80$ MJ/y/skill rather than $B_k = 50$ MJ/y/skill.

7.6 Reproduction without growth and body collapse for certain parameter values

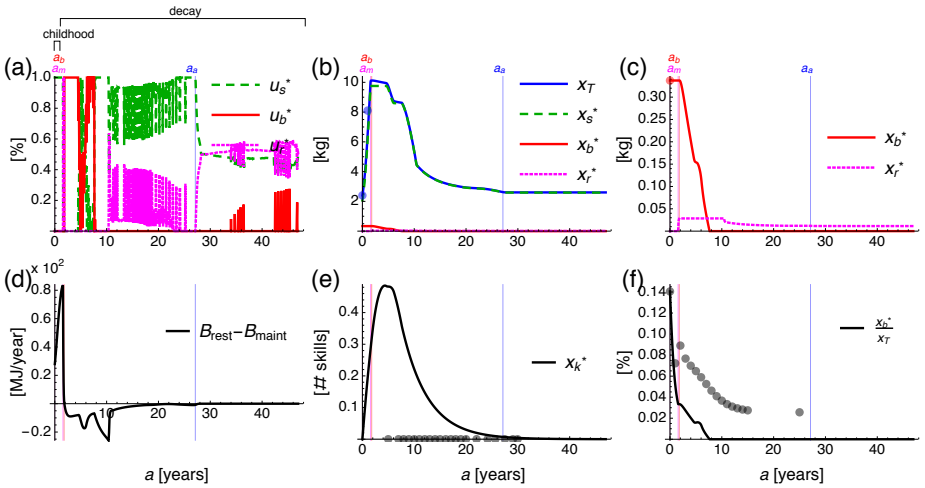


Figure S12: Reproduction without substantial growth with exponential competence when the environment is exceedingly challenging. Parameters are as in Fig. 1e-h, except that here $\alpha = 1.5$ rather than 1.15. The mass of reproductive tissue grows from 0 kg at birth, to 0.77 g at the age of $a_m \approx 6$ months, and reaches a peak of 4.64 g at $a_b \approx 8$ months. Jitter in the controls indicates that the optimal control problem is computationally challenging for GPOPS (this applies to all plots in the main paper and SI).

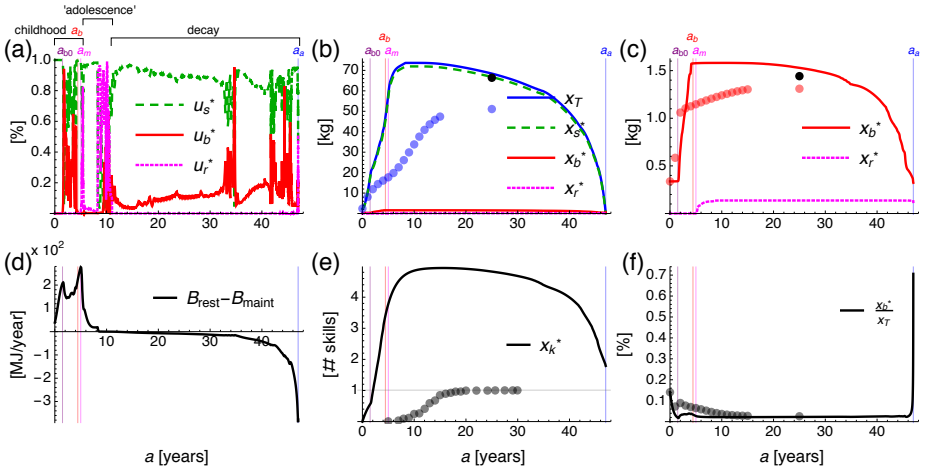


Figure S13: Brain and body collapse in adulthood with exponential competence when learning is exceedingly inexpensive. Parameters are as in Fig. 1e-h, except that here and $E_k = 100$ MJ/skill rather than 250 MJ/skill.

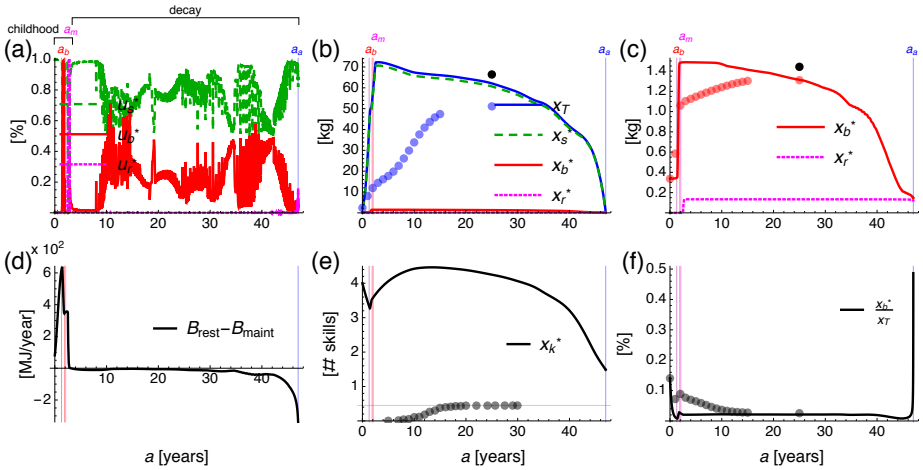


Figure S14: Brain and body collapse with exponential competence when the newborn has overly many skills. Parameters are as in Fig. 1e-h, except that here $x_k(0) = 4$ skills rather than 0.

7.7 A large brain is also favored by small metabolic costs of learning, few innate skills, and intermediate allocation of brain metabolic rate to skills

When varying the learning cost, adult brain mass is largest when learning is inexpensive but not exceedingly so (Fig. S15a). If learning is exceedingly inexpensive, the individual acquires enough skills while receiving (allo)parental care that it grows more than what it can maintain when (allo)parental care is absent. In this case, brain and body collapse during adulthood (Fig. S13). Otherwise, if learning is inexpensive but not exceedingly so, brain and body grow to levels that the individual can maintain when (allo)parental care is absent. With further increasingly expensive learning, skills grow more slowly and thus there is less growth metabolic rate at each age, yielding a decreasing adult brain mass (Fig. S15a). Yet, while small learning costs favor a larger adult brain mass, they also favor a larger adult body mass. Consequently, EQ is invariant with learning costs within the range of brain and body growth (Fig. S15d).

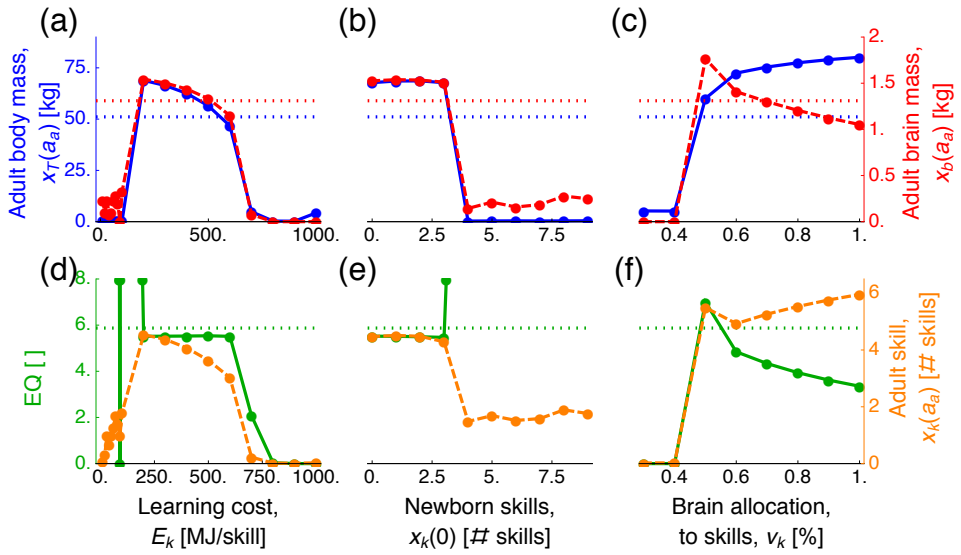


Figure S15: Predicted adult body and brain mass, EQ, and skill vs. other parameter values with exponential competence. See legend of Fig. 3. In d, jitter in EQ is due to increasing jittering in the controls when body and brain mass collapse.

When varying newborn skills, a larger adult brain mass is predicted when the newborn has fewer skills (Fig. S15b). If the newborn has overly many skills, the individual grows more during the (allo)parental care period than what it can maintain when (allo)parental care is absent, causing brain and body collapse during adulthood (Figs. S15b and S14).

Regarding allocation of brain metabolic rate to energy-extraction skills, brain mass is predicted to be larger with a decreasing, but not exceedingly, small brain allocation to skills (Fig. S15c). With an exceedingly small brain allocation to skills, the individual reproduces without substantial growth because skills grow little and the individual is unable to support itself when (allo)parental care becomes absent. Above a threshold, an increasing brain allocation to skills predicts a decreasing adult brain mass because the energetic input to skill growth is larger without the brain having to be as large [equation (A2)]. In contrast to brain mass and EQ, the predicted adult skill number increases with brain allocation to skills (Fig. S15f).

Comparative predictions with power competence are similar to those with exponential competence (Fig. S16).

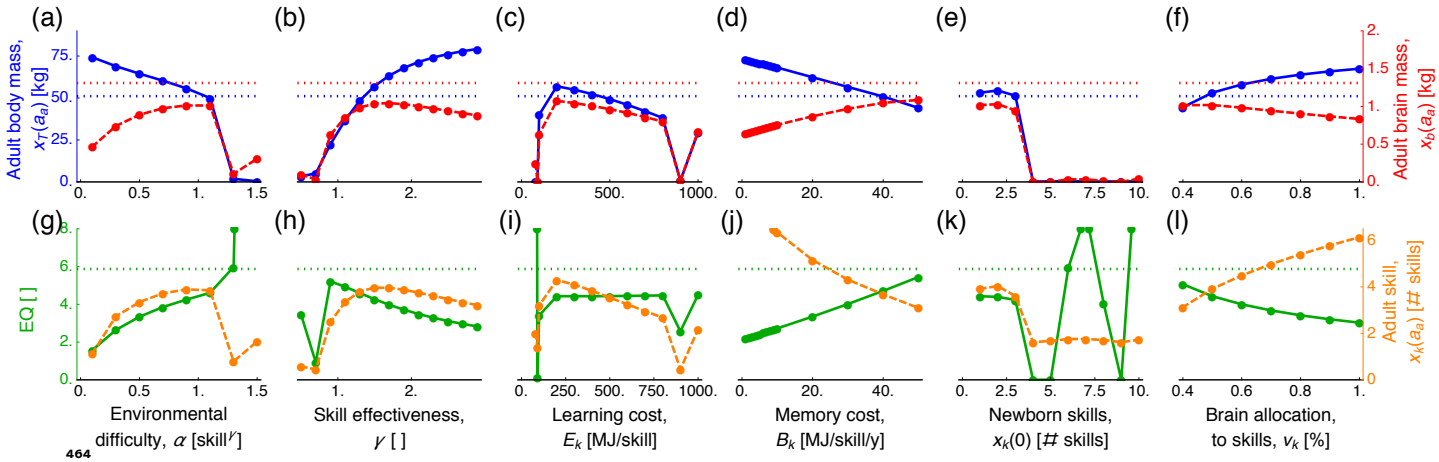


Figure S16: Predicted comparative patterns with power competence. See legend of Fig. 3. Jitter in EQ is due to increasing jittering in the controls when body and brain mass collapse.

References

Abramczuk, J. and Sawicki, W. (1974). Variation in dry mass and volume of nonfertilized oocytes and blastomeres of 1-, 2- and 4-celled mouse embryos. *J. Exp. Zool.*, **188**, 25–34.

Blaxter, K. (1989). *Energy Metabolism in Animals and Man.* Cambridge Univ. Press, Cambridge, UK.

Bryson, Jr., A.E. and Ho, Y.C. (1975). *Applied Optimal Control: Optimization, Estimation, and Control.* Taylor & Francis.

Dickey, R.P., Taylor, S.N., Lu, P.Y., Sartor, B.M., Rye, P.H. and Pyrzak, R. (2002). Effect of diagnosis, age, sperm quality, and number of preovulatory follicles on the outcome of multiple cycles of clomiphene citrate-intrauterine insemination. *Reproductive Endocrinology*, **78**, 1088–1095.

Diehl, M., Bock, H.G., Diedam, H. and Wieber, P.B. (2006). Fast direct multiple shooting algorithms for optimal robot control. In *Fast motions in biomechanics and robotics*, pages 65–93. Springer, Berlin.

Dorfman, R. (1969). An economic interpretation of optimal control theory. *Am. Econ. Rev.*, **59**, 817–831.

Eaton, S.B., Pike, M.C., Short, R.V., Lee, N.C., Trussell, J., Hatcher, R.A. et al (1994). Women's reproductive cancers in evolutionary context. *Q. Rev. Biol.*, **69**, 353–367.

Goyal, M.S., Hawrylycz, M., Miller, J.A., Snyder, A.Z. and Raichle, M.E. (2014). Aerobic glycolysis in the human brain is associated with development and neotenous gene expression. *Cell Metab.*, **19**, 49–57.

Gurven, M. and Kaplan, H. (2007). Longevity among hunter-gatherers: a cross-cultural examination. *Popul. Dev. Rev.*, **33**, 321–365.

Hou, C., Zuo, W., Moses, M.E., Woodruff, W.H., Brown, J.H. and West, G.B. (2008). Energy uptake and allocation during ontogeny. *Science*, **322**, 736–739.

Irie, T. and Iwasa, Y. (2005). Optimal growth pattern of defensive organs: the diversity of shell growth among mollusks. *Am. Nat.*, **165**, 238–249.

Iwasa, Y. and Roughgarden, J. (1984). Shoot/root balance of plants: optimal growth of a system with many vegetative organs. *Theor. Popul. Biol.*, **25**, 78–105.

Kamien, M.I. and Schwartz, N.L. (2012). *Dynamic Optimization*. Dover, Mineola, NY, 2nd edition.

Kelley, H.J., Kopp, R.E. and Moyer, H.G. (1967). Singular extremals. In G. Leitmann, editor, *Topics in Optimization*, pages 63–101. Academic Press, New York.

King, D. and Roughgarden, J. (1982). Graded allocation between vegetative and reproductive growth for annual plants in growing seasons for random length. *Theor. Popul. Biol.*, **22**, 1–16.

Kleiber, M. (1932). Body size and metabolism. *Hilgardia*, **6**, 315–351.

Kleiber, M. (1961). *The Fire of Life*. Wiley.

Kozłowski, J. (1992). Optimal allocation of resources to growth and reproduction: implications for age and size at maturity. *Trends Ecol. Evol.*, **7**, 15–19.

Kuzawa, C.W., Chugani, H.T., Grossman, L.I., Lipovich, L., Muzik, O., Hof, P.R. et al (2014). Metabolic costs and evolutionary implications of human brain development. *Proc. Nat. Acad. Sci. USA*, **111**, 13010–13015.

Lande, R. (1982). A quantitative genetic theory of life history evolution. *Ecology*, **63**, 607–615.

Magnusson, C., Hillensjö, T., Hamberger, L. and Nilsson, L. (1986). Oxygen consumption by human oocytes and blastocysts grown in vitro. *Hum. Reprod.*, **1**, 183–184.

McGee, E.A. and Hsueh, A.J.W. (2000). Initial and cyclic recruitment of ovarian follicles. *Endocrine Reviews*, **21**, 200–214.

Moses, M.E., Hou, C., Woodruff, W.H., West, G.B., Nekola, J.C., Zuo, W. et al (2008). Revisiting a model of ontogenetic growth: estimating model parameters from theory and data. *Am. Nat.*, **171**, 632–645.

Mylius, S.D. and Diekmann, O. (1995). On evolutionarily stable life histories, optimization and the need to be specific about density dependence. *Oikos*, **74**, 218–224.

O'Herlihy, C., De Crespigny, L.C., Lopata, A., Johnston, I., Hoult, I. and Robinson, H. (1980). Preovulatory follicular size: a comparison of ultrasound and laparoscopic measurements. *Fertil. Steril.*, **34**, 24–26.

Patterson, M.A. and Rao, A.V. (2014). GPOPS-II: A MATLAB software for solving multiple-phase optimal control problems using hp-adaptive gaussian quadrature collocation methods and sparse nonlinear programming. *ACM Trans. Math. Softw.*, **41**, 1–37.

Perrin, N. (1992). Optimal resource allocation and the marginal value of organs. *Am. Nat.*, **139**, 1344–1369.

517 Peters, R.H. (1983). *The Ecological Implications of Body Size*. Cambridge Univ. Press, Cambridge, UK.

518 Schmidt-Nielsen, K. (1984). *Scaling*. Cambridge Univ. Press.

519 Schniter, E., Gurven, M., Kaplan, H.S., Wilcox, N.T. and Hooper, P.L. (2015). Skill ontogeny among Tsimane
520 forager-horticulturalists. *Am. J. Phys. Anthrop.*, **158**, 3–18.

521 Sears, K.E., Kerkhoff, A.J., Messerman, A. and Itagaki, H. (2012). Ontogenetic scaling of metabolism, growth,
522 and assimilation: Testing metabolic scaling theory with *Manduca sexta* larvae. *Physiol. Biochem. Zool.*, **85**,
523 159–173.

524 Sydsæter, K., Hammond, P., Seierstad, A. and Strom, A. (2008). *Further Mathematics for Economic Analysis*.
525 Prentice Hall, 2nd edition.

526 West, G.B., Brown, J.H. and Enquist, B.J. (2001). A general model for ontogenetic growth. *Nature*, **413**, 628–631.

527 Ziółko, M. and Kozłowski, J. (1983). Evolution of body size: an optimization model. *Math. Biosci.*, **64**, 127–143.