



**HAL**  
open science

# Fitness Estimation for Ecological Studies: An Evaluation in Columbian Ground Squirrels

F. Stephen Dobson, Jan Murie, Vincent Viblanc

► **To cite this version:**

F. Stephen Dobson, Jan Murie, Vincent Viblanc. Fitness Estimation for Ecological Studies: An Evaluation in Columbian Ground Squirrels. *Frontiers in Ecology and Evolution*, 2020, 8, 10.3389/fevo.2020.00216 . hal-02903692

**HAL Id: hal-02903692**

**<https://hal.science/hal-02903692>**

Submitted on 30 Sep 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



23 **Abstract.** The topics of evolution by natural selection and ecological interactions are closely  
24 intertwined. Thus, measurements of evolutionary fitness are ubiquitous in the ecological  
25 literature. As an empirical problem, the components of fitness, reproduction and survival, may  
26 be analyzed to produce fitness estimates in several ways. These can be divided into annual  
27 estimates that are most appropriate for short-term (e.g., annual) and experimental studies, and  
28 lifetime fitness estimates that are most appropriate for evaluating functional organismal traits.  
29 The latter are appropriate for comparative studies of natural selection in species with lifetimes  
30 that extend over several years. These estimates may also be particularly useful for estimating  
31 the direct and indirect components of inclusive fitness, an important topic for the evolution of  
32 cooperation. We reviewed examples of some of these alternatives from our research on  
33 Columbian ground squirrels (*Urocitellus columbianus*). Using empirical data, we also test the  
34 degree of correspondence of annual fitness, lifetime reproductive success, and individual  
35 fitness measures that are based on matrix methods. We conclude that correspondence of  
36 different methods is not strong, though each method appears most appropriate for different  
37 types of traits and research questions about fitness differences among trait forms.

38

39

## INTRODUCTION

40

41

42

43

44

In his seminal book on evolutionary ecology, G. Evelyn Hutchinson described the interaction between ecology and evolution as “the ecological theater and the evolutionary play” (Hutchinson 1967). The idea is that ecological conditions provide the selective influences on phenotypic traits that evolve through the process of natural selection (Darwin 1859). In a population, phenotypic traits may change in frequency over time, and this change can be

45 inferred from fitness differences, patterns of reproduction and survival that are associated with  
46 different trait forms. This evolutionary principle, that trait adaptation by natural selection occurs  
47 in ecological environments, has been applied to many empirical studies, and is assumed by  
48 many more. Studies of natural selection on traits commonly measure changes in trait  
49 frequencies that occur over time, either from year to year or from generation to generation  
50 (e.g., studies reviewed by Endler 1986; Charmantier et al. 2014). Such changes are produced by  
51 variations in reproduction and survival, and are used as surrogate measures of evolutionary  
52 fitness for the trait forms that individuals carry.

53         Empirical studies often group individuals that express alternative forms of a trait into  
54 “trait groups” or use individuals to represent a continuum of trait values (termed “trait forms”  
55 for continuous traits). Although trait fitness is typically estimated through the reproduction and  
56 survival of individuals that carry specific trait forms, it is the change in frequency of traits, not  
57 individuals, that is the result of natural selection. In this sense, trait groups can be considered  
58 alternative adaptations that respond to environmental conditions. Good examples are studies  
59 of the influence of global changes in climate on the timing of reproductive events or other  
60 elements of a species’ lifecycle (e.g., Visser and Both 2005; Chamaillé-Jammes et al. 2006;  
61 Parmesan 2006; Lane et al. 2012; Tafani et al. 2013; Dobson et al. 2016; Radchuck et al. 2019).  
62 An important caveat is that individuals have many traits, some of which may be genetically  
63 correlated through linkage disequilibrium or pleiotropy (e.g., Duckworth and Kruuk 2009; Bize  
64 et al. 2017; Mullan et al. 2018). Thus, single-trait studies may not identify the traits that are the  
65 targets of selection, including important evolutionary tradeoffs, and the actual interactions with  
66 agents of selection in the environment (Falconer 1952; Lande and Arnold 1983; Roff and

67 Fairbairn 2012). In addition, the fitness of a trait form is relative to the success of alternative  
68 trait forms, as populations dynamically change. Hence, the fitness of traits is relative to what  
69 other individuals are doing at a given moment in time for a population that may be increasing  
70 or decreasing. Measuring fitness for alternative trait forms is thus complicated by the need to  
71 account for population demographics (Metcalf and Pavard 2007).

72         Nonetheless, studies of single or a few traits provide a good starting place for more in-  
73 depth examinations of suites of traits and suites of environmental factors. Early studies  
74 examined the influence of environmental events on morphological traits and differential  
75 survival (e.g., Bumpus 1899; Kettlewell 1961; Lande and Arnold 1983), but later studies  
76 examined both reproduction and survival (e.g., Pemberton et al. 1999). Measures that combine  
77 reproduction and survival (e.g., annual fitness; Qvarnström et al. 2006) have been applied to  
78 examine such traits expressed on an annual basis such as reproductive phenology (e.g., Lane et  
79 al. 2012; Dobson et al. 2017). Other measures examine traits expressed for a lifetime, like  
80 foraging patterns (using lifetime reproductive success, LRS; Clutton-Brock 1988; Altmann 1991)  
81 or age at first breeding (via individual fitness from matrix methods; McGraw and Caswell 1996).  
82 Lifetime reproductive success is not sensitive to the timing of successful reproductive events  
83 (that might occur at any point in a lifecycle), but individual fitness measures are sensitive to  
84 when such events occur (Brommer et al. 2002). The impact of this difference on studies of  
85 natural selection is currently unknown (Brommer et al. 2004; Reid et al. 2019).

86         Fitness is a population parameter, though it is often estimated for individuals. A variety  
87 of fitness estimates have been developed for individuals, using demographic data from  
88 populations (e.g., Qvarnström et al. 2006; Clutton-Brock 1988; McGraw and Caswell 1996;

89 Viblanc et al. 2010; Rubach et al., *submitted*). However, natural selection operates on trait  
90 groups in populations, or rather on suites of correlated traits. Thus, individuals are divided into  
91 groups, effectively sub-populations, that exhibit alternative expressions of the trait(s) of  
92 interest. When we compare these alternative trait groups (or forms, for continuous traits), a  
93 rough idea of the association of the trait form and fitness (reproduction or survival, but best is a  
94 combination of the two) can be obtained. This is a rough estimate of selection on the trait  
95 forms. One problem is that the trait of interest may be genetically correlated with other traits  
96 that constrain or magnify the change in trait forms from one generation to the next (e.g., Lande  
97 and Arnold 1983; Price and Langen 1992). Thus, identifying which trait forms are the actual  
98 “targets” of selection can be difficult. Another problem for inference about natural selection is  
99 that we may have little idea about the heritability of trait forms. However, given caution, it is  
100 instructive to understand which trait forms appear associated with fitness differences and  
101 which do not, remembering that selection and response to selection are two different things.

102         Our purpose is to compare and contrast different methods of measuring fitness, applied  
103 during the course of our long-term research program on Columbian ground squirrels  
104 (*Urocitellus columbianus*). We provide examples from past studies using different fitness  
105 measures, all based on empirical studies of reproduction and survival during 28 years of  
106 research on ground squirrels. In these examples, we used information about reproduction,  
107 survival, and traits of individuals to evaluate fitness of alternative trait groups or trait forms.  
108 We used annual fitness estimates in some studies, and lifetime fitness in others. In some cases,  
109 experimental treatments provided the alternative trait groups, but we also applied continuous  
110 variables to estimate trait forms. In one study, we examined direct and indirect components of

111 inclusive fitness of cooperative behavior from helpful close kin. Herein, our goals were: first, to  
112 review our long-term studies of reproduction and survival in a comparison of annual and  
113 lifetime estimates of fitness, and lifetime reproductive success and individual fitness estimates  
114 from matrix models. We expected that annual and lifetime estimates of fitness might not  
115 correspond closely, since environmental conditions may vary among years (Lane et al. 2012;  
116 Dobson et al. 2016). Second, to test this expectation, by comparing annual and alternative  
117 lifetime fitness measures for correspondence. Different methods should correspond closely, if  
118 they give an accurate estimate of the fitness of individuals and their traits.

119

120

## METHODS

### 1. Field site and ground squirrels

122 Columbian ground squirrels (*Urocitellus columbianus*) were studied from 1992 to 2019  
123 in the Sheep River Provincial Park in Alberta, Canada (50°38'N, 114°39'W; elevation 1,550 m;  
124 see population #3 of Fig. 1 in Dobson 1994). The study population was about 1.8ha of meadow  
125 surrounded by pine-spruce forest on 3 sides and the gorge of the Sheep River on the final side.  
126 The meadowland contained grasses and forbs that provided forage for the ground squirrels,  
127 and was honeycombed with burrows that were dug out each year by these semi-fossorial  
128 rodents. Natural predators included coyotes (*Canis latrans*), red foxes (*Vulpes vulpes*), long-  
129 tailed weasels (*Mustela erminea*), golden eagles (*Aquila chrysaetos*), and northern goshawks  
130 (*Accipiter gentilis*). Columbian ground squirrels hibernate for 8-9 months each year and have a  
131 single breeding season during their short period of annual activity (Dobson and Murie 1987;  
132 Dobson et al. 1992). Our studies have focused primarily on female ground squirrels, because

133 they are matrilocal and many were monitored over their complete lifespans (only 3 of 338 adult  
134 females were immigrants, over 28 years). Female ground squirrels can live up to 9-10 years (an  
135 exceptional female lived to 14 years), and usually begin reproduction at 2 to 3 years of age  
136 (Broussard et al. 2008), though yearling females occasionally reproduce in our population  
137 (Rubach et al., submitted).

138           Each spring, we trapped all ground squirrels within 3 days of emergence from  
139 hibernation in live traps (13 X 13 X 40 cm; Tomahawk Live Trap, Hazelhurst, Wisconsin) baited  
140 with a small amount of peanut butter. Each individual was induced to enter a cloth bag for  
141 initial handling. We weighed each ground squirrel to the nearest 5g in the handling bag with a  
142 Pesola spring-slide balance (1000g; Schindellegi, Switzerland), fitted untagged individuals with a  
143 pair of numbered metal ear tags (Monel #1005-1; National Band and Tag, Newport, Kentucky),  
144 and measured their head width (zygomatic arch breadth) to the closest 0.1 mm with a pair of  
145 dial calipers. Each ground squirrel was examined for sexual condition (males abdominal or  
146 scrotal testes; females for appearance of the vulva, closed, open, open and swollen, and with or  
147 without a copulatory plug), presence and abundance of fleas, and wounds. We gave each  
148 ground squirrel a unique black mark on the pelage with black dye (human hair dye; Lady Clairol  
149 Hydrience, Proctor & Gamble Co., Cincinnati, Ohio) for later visual identification. Because  
150 unmarked individuals could be quickly identified and trapped, we were able to capture, mark,  
151 and examine every ground squirrel in the population that emerged from hibernation in every  
152 year of the study.

153           In the spring, female ground squirrels were watched daily for characteristic reproductive  
154 behavior (and regularly trapped to check the appearance of their vulva), to determine their



155 single annual day of estrus, during which they are typically receptive to mating males for a  
156 period of 5-6 hours. Subsequently, females typically went through 24 days of gestation and  
157 gave birth underground in single-entrance “natal burrows” where the young are nursed and  
158 protected by territorial mothers (Murie and Harris 1988) during about 27 days of lactation  
159 (Murie and Harris 1982). Juvenile ground squirrels first emerge above ground from their natal  
160 burrows near the time of weaning. Because mothers kept juveniles in single-entrance natal  
161 burrows, we were able to both identify the mother and capture emerging juveniles, usually on  
162 their first day above ground. Juveniles were trapped (Tomahawk or other cage traps), ear-  
163 tagged, examined to identify sex, examined for presence and number of fleas, and dye-marked  
164 with unique symbols. Males typically dispersed to other populations as yearlings, or remained  
165 residents on the study site but resided away from the area of their natal nest (and thus only  
166 occasionally interacting with female kin; dispersal pattern reviewed in Neuhaus 2006). In  
167 contrast, females were highly philopatric and interactive, very rarely dispersing to other  
168 populations (Arnaud et al. 2012). Thus, we were able to build lifetime records of reproduction  
169 and survival for females that could be used to estimate a variety of fitness measures.

170

## 171 **2. Fitness estimates**

### 172 ***2.1. Annual fitness***

173 Our fitness estimates were applied to females, because females were matrilocal (Arnaud  
174 et al. 2012) and paternity was not measured in all years of our study (Raveh et al. 2010; Balmer  
175 et al. 2019). Using the long-term data, we calculated annual fitness as the mean number of  
176 gene copies that each female had represented in the following year (Qvarnström et al. 2006). A

177 female had one copy if she survived to the next year, plus a half copy for each of her new  
178 offspring that also survived to the following year. For the estimate, these values were added  
179 together.

180

## 181 ***2.2. Lifetime reproductive success***

182 Lifetime reproductive success (LRS) was measured by simply adding up the number of  
183 offspring that each female produced over her lifetime. We did this for offspring produced near  
184 the time of weaning, when young first emerged from natal burrows. We also constructed a  
185 relative index of lifetime reproductive success of each mother compared to their peers, by  
186 regressing lifetime reproductive success for each female onto the lifetime reproductive success  
187 of her cohort. The residuals of this regression were used as a relative index ( $LRS_{rel}$ ).

188

## 189 ***2.3. Individual (lifetime) fitness***

190 We calculated individual fitness estimates ( $\lambda_{ind}$ ) for each female of the population that  
191 had lived-out her entire lifespan (thus, excluding those alive at the end of the study), following  
192 the matrix approach of McGraw and Caswell (1996). For each female, we constructed a matrix  
193 that had half her reproductive output (at either weaning of offspring, or offspring that survived  
194 their first hibernation season) specified on the top row of the matrix. This represented her  
195 annual contribution to reproductive production, the other half being from the male. The matrix  
196 had ones on the off diagonal, and zeros in all other unfilled elements (for an example, see  
197 Viblanc et al. 2010). The dominant right eigenvalue of this matrix (of value  $\lambda_{ind}$ ) was a growth

198 parameter for the matrix, and was taken as the lifetime individual fitness of the female and  
199 thus also her phenotype. The method was explained in detail by McGraw and Caswell (1996).

200 The individual fitness measure of McGraw and Caswell (1996) required additional  
201 attention. Any estimate of fitness of an individual is relative to others in the population. Over a  
202 lifetime that may extend several years, rodent populations are well known to fluctuate (e.g.,  
203 Boonstra and Krebs 2012; Fauteux et al. 2016; Brommer et al. 2017; Bonnet and Postma 2018).  
204 For example, our population initially grew for about 10 years, then declined by close to 50%  
205 over two years, and then had a relative stable and gently increasing period for about 15 years  
206 (Figure 1). During times of population increase, a female with an estimated  $\lambda_{\text{ind}}$  of 1.0 does  
207 poorly compared to others in the population, since population growth (and the estimated  
208 fitness of an average female in the population) would have a  $\lambda$  value greater than 1.0. A female  
209 would be doing quite well, however, if the population was decreasing (population  $\lambda < 1.0$ ).  
210 Thus, it may be necessary to adjust  $\lambda_{\text{ind}}$  for changes in population size or population growth, if  
211 one wishes to compare the traits among females that experience different populations changes  
212 over their lifetimes.

213 We adjusted our lifetime individual fitness measure for changes in the population in two  
214 ways. First, we regressed individual fitness ( $\lambda_{\text{ind}}$ ) on a measure of population change in which  
215 comparative population values were produced by specifying proportional year-to-year  
216 population changes into a matrix in lieu of reproductive fertility values, but still with values of  
217 one in the off diagonal (to create  $\lambda_{\Delta N}$ , after Viblanc et al. 2010; Dobson et al. 2012). These  
218 matrices were constructed for the same years as each female's lifetime, and gave an indication  
219 of how the population, on average and over all females, was changing over time. Then  $\lambda_{\text{ind}}$

220 values were regressed on  $\lambda_{\Delta N}$  values. The residual values of this analysis adjust for changes in  
221 population size during a female's lifetime, and we added 1.0 to these to make interpretation  
222 easier (producing  $\lambda_{rel\Delta N}$ ). Second, we calculated actual population growth using a Leslie matrix  
223 (Leslie 1945) for the cohort of each female, during that female's lifetime, and then similarly  
224 regressed  $\lambda_{ind}$  on  $\lambda_{Leslie}$  and added 1.0 to the residuals of the regression (producing  $\lambda_{relL}$ ; after  
225 Rubach et al., submitted).  $\lambda_{rel\Delta N}$  and  $\lambda_{relL}$  were used as estimates of a female's fitness relative to  
226 her competitors, compared to the population at large and compared to her cohort, during her  
227 lifetime.

#### 228 ***2.4. Direct and indirect components of inclusive fitness:***

229 We estimated the direct and indirect components of inclusive fitness in two related  
230 ways. Inclusive fitness was relative to a particular trait, the presence of close kin ("genial  
231 neighbors") that demonstrably improve reproductive success and  $\lambda_{ind}$  relative to changes in  
232 population size ( $\lambda_{\Delta N}$ ), to produce  $\lambda_{rel\Delta}$  (after Viblanc et al. 2010; Dobson et al. 2012). First, we  
233 computed  $\lambda_{rel\Delta N}$  for mothers with and without co-breeding close kin and compared these  
234 values. Next, we estimated inclusive fitness from a direct component (the mean relative  
235 individual fitness of mothers without kin present) and an indirect component (averaged:  
236 relative individual fitness for relatives present minus the mean fitness of mothers without kin  
237 present, times the degree of relatedness). These estimates were calculated based on the  
238 number of weaned offspring and included as kin only those relatives that appear to be  
239 recognized as close kin in the field (King 1989). "Uterine kin" (viz., mothers, daughters, and  
240 littermate sisters) recognize one another via social familiarization in the natal nest-burrow  
241 (Hare and Murie 1996; Viblanc 2010). Other females were classed as distant and non-kin, and

242 used for comparison to close kin. Further details of inclusive fitness calculations can be found  
243 in Dobson et al. (2012).

244

245

## RESULTS AND DISCUSSION

246

### 1. Fitness estimates in Columbian ground squirrels: review from a long-term study

247

#### 1.1. Annual fitness estimates

248

Raveh et al. (2015) conducted an experiment on reproductive Columbian ground

249

squirrel females that involved the removal of fleas (*Oropsylla* spp.) using a spot-on pet

250

insecticide. The experiment artificially created the trait groups “with natural flea loads” and

251

“with no fleas,” and thus tested for a fitness cost to parasitism. The results indicated no

252

significant difference between treated and untreated mothers using the Qvarnström et al.

253

(2006) annual fitness measure (respectively;  $1.13 \pm 0.13$ ,  $n = 28$ ; and  $1.06 \pm 0.14$ ,  $n = 26$ ; mixed

254

model,  $\chi^2 = 0.06$ ,  $df = 1$ ,  $P = 0.80$ ). In addition, they found no significant difference in the

255

annual fitness measure for mothers and young that occupied nests that were infested with

256

fleas during lactation from nests that were not infested (respectively;  $1.25 \pm 0.16$ ,  $n = 20$ ; and

257

$1.08 \pm 0.14$ ,  $n = 26$ ; mixed model,  $\chi^2 = 1.29$ ,  $df = 1$ ,  $P = 0.26$ ). The conclusion of the study was

258

that the removal of ectoparasites was not a significant influence on annual fitness, nor on

259

several other measures related to maternal fitness. The Qvarnström et al. (2006) method was

260

an appropriate measure for the experimental contrast and natural comparison of the

261

ectoparasite “treatments,” and two trait groups (viz., with and without parasites) were used in

262

both comparisons. This method tracks the number of gene copies that were passed on in the

263 population from one year to the next: one complete copy if the mother survives, and a half  
264 copy of her genes passes on through her male and female offspring.

265 Viblanc et al. (2016) applied annual fitness to a comparative network study of  
266 aggression by adult female Columbian ground squirrels. There were no significant differences  
267 in fitness among reproductive females according to the aggression that they received from  
268 other mothers. But aggression directed towards non-close kin was 2.3 times greater than  
269 aggression directed towards reproductive close kin (randomized network analysis,  $P < 0.001$ ;  
270 Fig. 3 in Viblanc et al. 2016). Mothers that most commonly committed aggression towards  
271 other females had significantly greater annual fitness (randomized network analysis,  $P = 0.004$ ;  
272 see Fig. 4 in Viblanc et al. 2016). The trait forms were different levels of committed aggression  
273 (chases and fights). Here, the Qvarnström et al. (2006) method was an appropriate measure of  
274 network analyses from a single year of behavioral data, but in this case the trait form was a  
275 continuous axis.

276 Annual fitness can also be applied to longer-term studies, such as those examining  
277 influences of changing climate. In Columbian ground squirrels, variations in spring and summer  
278 climates have significant influences on annual fitness (Lane et al. 2012; Dobson et al. 2016).  
279 Over a 20-year period, when adult females emerged from hibernation earlier their annual  
280 fitness was greater (Figure 2). For this study, the trait forms were different dates of emergence  
281 from hibernation, a variable that was repeated for most individuals among years. Emergence  
282 from hibernation was influenced by snowmelt, and lower fitness ensued when spring melt-off  
283 of snowpack was later (Figure 3). Additionally, dry hot conditions during summer also had a  
284 strong influence on fitness (Figure 4). These seasonal climatic influences on the fitness of adult

285 females produced strong influences on population size as well, so that a year of especially late  
286 spring melt-off of snow and hot and dry conditions in summer were associated with a nearly  
287 50% decline in the population (Dobson et al. 2016). In these studies, annual fitness was an  
288 appropriate index for comparisons, in part because annual events were studied. The  
289 Qvarnström et al. (2006) method has been used in a similar manner in several comparative and  
290 experimental field studies (e.g., Arnaud et al. 2013; Lane et al. 2015; Hoogland and Brown 2016;  
291 Lane et al. 2019).

292

### 293 ***1.2. Lifetime reproductive success and individual fitness estimates***

294 Interest in the individuals that carry traits often leads to comparisons of the lifetime  
295 reproductive success of individuals with different phenotypic values of traits. But natural  
296 selection applies to traits and combinations of traits, rather than individuals per se. The  
297 frequencies of trait groups change over time when natural selection occurs. Thus, rather than  
298 the number of offspring or even grand-offspring for measuring changes due to natural selection  
299 (Brommer et al. 2004; Reid et al. 2019), the most useful measure is the growth rate of trait  
300 forms or associations of trait forms among generations. The most commonly used measure of  
301 increase in trait forms is lifetime reproductive success (e.g., Clutton-Brock 1988; Grafen 1988;  
302 Merilä and Sheldon 2000; Jensen et al. 2004; Descamps et al. 2006; McLoughlin et al. 2007).

303 The growth of trait groups can also be estimated by calculation of individual fitness over  
304 the lifetime of individuals in the population (McGraw and Caswell 1996). This method uses  
305 matrix algebra (classically used for estimating population growth) to estimate the increase in  
306 trait forms from changes in the frequency of phenotypic traits. The logic of doing this follows a

307 long history of measuring fitness from the intrinsic growth rate of individuals that carry  
308 different trait forms (Stearns 1992; Roff 2002). A major advantage of the individual fitness  
309 approach is that it is sensitive to the timing of reproduction, so that offspring produced early in  
310 a mother's life contribute more to fitness than those produced later (Brommer et al. 2002).

311

### 312 *1.2.1. Age at first reproduction*

313 Demographic theory suggests that offspring produced early in life, if they carry a  
314 particular trait form, may themselves begin to reproduce earlier, thus contributing to further  
315 spread of the trait form over time (Stearns 1992). Individual fitness was devised for an  
316 examination of age at reproductive maturity in Sparrowhawks (*Accipiter nisus*) and blue tits  
317 (*Parus caeruleus*) (McGraw and Caswell 1996). For Columbian ground squirrels, a similar  
318 exploration of individual fitness revealed greater individual fitness values for individuals that  
319 begin to reproduce successfully at an earlier age (Rubach et al., submitted). This was done with  
320 a calculation similar to that of McGraw and Caswell (1996). The number of weaned offspring  
321 was used to estimate reproductive success. Regression of  $\lambda_{ind}$  (as in McGraw and Caswell 1996)  
322 on  $\lambda_{Leslie}$  (cohort growth) showed that an adjustment for changes in population growth was  
323 needed ( $R^2 = 0.52$ ,  $F = 52.6$ ,  $P < 0.0001$ ,  $n = 148$ ). Thus, Rubach et al. (submitted) used the  
324 residuals of this regression ( $\lambda_{rel}$ ) to estimate relative individual fitness.  $\lambda_{rel}$  differed  
325 significantly among females that first reproduced at ages 1, 2, and 3-5 years old (Figure 5).  
326 Lifetime reproductive success, however, showed no significant difference among the trait  
327 groups (i.e., females that first reproduce at ages of 1, 2, and 3 and above). Statistically  
328 comparing lifetime reproductive success for mothers to that of their peers (i.e., their cohort



329 lifetime reproductive success, to produce  $LRS_{rel}$ ) produced little difference among females that  
330 first reproduced at different ages. Additionally, an earlier study that used lifetime reproductive  
331 success as a fitness measure found no significant difference between females that first  
332 reproduced at ages 2 and 3 (Neuhaus et al. 2004).

333         Examples that examine age at maturity have presented the strongest case for using  
334 individual fitness estimates, since the age at which reproduction begins has a strong influence  
335 on population growth (Cole 1954; Oli and Dobson 2003). Examples of fitness differences that  
336 appear to favor early breeding include blue tits and sparrowhawks (McGraw and Caswell 1996),  
337 Ural owls (Brommer et al. 1998), wood ducks (*Aix sponsa*) (Oli et al. 2002), and yellow-bellied  
338 marmots (*Marmota flaviventris*) (Oli and Armitage 2003). Oli and Armitage (2008), found that  
339 female marmots that delayed breeding suffered a loss of inclusive fitness, even when direct  
340 fitness was augmented by indirect fitness benefits from reproduction of close kin. The attempt  
341 of Rubach et al. (submitted) to compare the lifetime reproductive success of individual females  
342 to that of their cohorts ( $LRS_{rel}$ ) yielded no beneficial insights to simple use of lifetime  
343 reproductive success. Thus, based on present evidence, no measure of fitness, whether annual  
344 or over the lifespan, is certain to provide an accurate description of natural selection or  
345 evolutionary response to natural selection. For long-term studies that examine complete  
346 lifespans, a cautious approach might be to apply a relative measure based on individual fitness  
347 and either on changes in population size or on population growth rate.

348         Relative individual fitness (e.g.,  $\lambda_{relL}$  or  $\lambda_{rel\Delta N}$ ) takes changes in population size into  
349 account, and gives an estimate of the growth rate of the different trait groups. By asking how  
350 fitness differed among the trait groups, Rubach et al. (submitted) assumed that females that

351 reproduce at different ages have a trait that can be passed on to future generations: this is the  
352 assumption of genetic variation and heritability. While this idea might be challenged, the whole  
353 point of looking at a fitness measure is the search for an evolutionary advantage. If a trait  
354 undergoes selection, but exhibits no response to selection (viz., due to limited heritability,  
355 antagonistic pleiotropy, or genetic correlations with other traits) the results are perhaps less  
356 interesting. Additionally, when the number of offspring that survived until their first possible  
357 reproductive season was used to estimate reproductive success, the advantage for earlier  
358 reproduction by mothers was not quite significant, though it still had a small to medium effect  
359 size. Production of the next generation is meaningful for natural selection in terms of offspring  
360 that themselves survive to reproduce in the next generation (e.g., Boyce and Perrins 1985).  
361 Naturally, the different trait groups might vary due to environmental factors or random  
362 variations in resource acquisition over time, so it is important to remember that offspring may  
363 not express the trait forms of their parents, particularly for phenotypically plastic traits, where  
364 expression might be influenced by the environment.

365

### 366 *1.2.2. Kin selection and inclusive fitness*

367 Individual fitness measures are most appropriate for traits that are expressed once  
368 during an individual's lifetime, as is the case with many developmental traits. Age at maturity is  
369 one such, but many traits of temperate species are expressed on an annual basis (e.g., litter  
370 size, phenology of reproduction, seasonal cycles in activity or body mass, etc.). Nonetheless,  
371 some evolutionary characteristics may vary during an individual's lifetime, but the primary  
372 interest is in the cumulative effect of the social or ecological environment on fitness. Inclusive

373 fitness is an example of an advantageous phenomenon (the presence of cooperative and  
374 reproductive close kin that augment maternal fitness) that can accumulate over a lifespan.  
375 Here the trait is usually some sort of behavioral cooperation with close genetic relatives, so that  
376 kin selection is a possible influence on the behaviors (Hamilton 1964). In this case, inclusive  
377 fitness (an individual's total fitness) has two components: fitness accrued by an individual in  
378 the absence of help from genetic relatives (the direct component) and fitness accrued from the  
379 help that the individual gives to genetic relatives (the indirect component).

380           Our examination of possible kinship advantages in fitness terms began with an  
381 examination of whether there was a difference in the direct fitness component between female  
382 Columbian ground squirrels that reproduced in the presence of close kin and those that had no  
383 co-reproductive close kin with which to cooperate (Viblanc et al. 2010). The form that  
384 cooperation took was lowered aggression (viz. greater tolerance) during co-reproduction of  
385 adult females and their mother, littermate sisters, and daughters (King 1989; Viblanc et al.  
386 2016). The numbers of co-breeding close kin females per year (during the reproductive  
387 lifespan) were the trait forms. We found a significant association of the number of co-breeding  
388 close female kin and number of offspring at weaning, though at a medium effect size  
389 (estimated from path analysis of  $\lambda_{rel\Delta N}$ ;  $\rho = 0.29$ ,  $P = 0.01$ ,  $n = 70$ ). In turn, litter size was highly  
390 significantly associated with relative individual fitness ( $\rho = 0.79$ ,  $P < 0.0001$ ,  $n = 70$ ), resulting in  
391 a small-to-medium indirect effect of number of close kin on fitness (indirect path coefficient,  $\rho$   
392 = 0.23, significant when both direct coefficients are significant, Cohen 1988). Viblanc et al.'s  
393 (2010) use of relative individual fitness ( $\lambda_{rel\Delta N}$ ) was necessary because individual fitness ( $\lambda_{ind}$ )

394 was significantly associated with changes in population density (estimated from  $\lambda_{\Delta N}$ ;  $R^2 = 0.33$ ,  $F$   
395 = 33.387,  $n = 70$ ,  $P = 0.001$ ,  $n = 70$ ).

396 To extend our study of kin effects, we estimated inclusive fitness, to include the indirect  
397 component in addition to the direct component previously calculated (Dobson et al. 2012). Our  
398 goal was to evaluate the possible importance of an indirect component to inclusive fitness in  
399 Columbian ground squirrels. Mothers were classified as having co-breeding close kin versus not  
400 having close kin during their reproductive lifespans, so the alternative trait groups were the  
401 presence or absence of potentially helpful close kin. The number of weaned offspring was used  
402 to estimate female reproductive success, and the analyses used an adjustment for changes in  
403 the population during a female's lifetime (in this case,  $\lambda_{ind}$  was regressed on  $\lambda_{\Delta N}$  and 1.0 was  
404 added to the residuals, to produce  $\lambda_{rel\Delta N}$  for each female). Relative indirect fitness accounted  
405 for over 40% of a mother's inclusive fitness, a substantial and significant amount ( $0.43 \pm 0.08$   
406 SE,  $t = 5.71$ , d.f. = 28,  $P < 0.0001$ ). In addition, as the number of close co-breeding close  
407 relatives increased, a mother's inclusive fitness increased significantly (Figure 6). The analyses  
408 were greatly facilitated by use of the relative individual fitness approach.

409

## 410 **2. Empirical comparison of annual fitness, lifetime reproductive success, and individual** 411 **fitness**

412 The use of annual and lifetime fitness estimates made us question how closely these  
413 estimates correspond. To compare annual and lifetime measures, we used samples of females  
414 when they were different prime breeding ages, namely 3, 4, 5, and 6 years old (respectively;  
415 101, 80, 59, 32 mothers). Of course, these values for individual females were not independent

416 and were from different years over the 28-year study period. The number of offspring at  
417 weaning was used to estimate reproductive success for these subsequent analyses. We  
418 compared the annual fitness values for the females in each age group separately to the relative  
419 lifetime individual fitness estimates ( $\lambda_{rel\Delta N}$ ) and to the lifetime reproductive success of these  
420 same females with correlations (Figure 7; e.g., a single datum would be an annual value for a 3-  
421 year-old female and her lifetime fitness, the latter estimated by relative individual fitness  $\lambda_{rel\Delta N}$   
422 or estimated by lifetime reproductive success). The values of the correlations were used as  
423 indications of effect size (Cohen 1988) for the similarities of annual and lifetime values, with  
424 small ( $r = 0.10$ ), medium ( $r = 0.30$ ), and large effects ( $r = 0.50$ ). Similarity of estimates of the  
425 association of annual and individual fitness (unadjusted) were consistently between medium  
426 and large, averaging around 40% ( $r = 0.380$ ). Similarity of annual fitness and lifetime  
427 reproductive success also averaged around 40% ( $r = 0.423$ ), but were much more variable. The  
428 annual estimates of fitness were meant to reveal the influence of an experiment or annual  
429 comparisons. Since year-to-year variations in the environment occurred and these variations  
430 might well average out during an extended lifetime (9-12 years for 6 of the females in our  
431 sample of 101 mothers), a modest effect size might have been expected.

432         When relative individual fitness ( $\lambda_{rel\Delta N}$ ) was compared to lifetime reproductive success,  
433 the correlation was significant but fairly moderate (Figure 8;  $r = 0.345$ ,  $n = 132$  mothers,  $t =$   
434  $4.186$ ,  $P < 0.0001$ ). When the cohort Leslie matrix was used to estimate relative individual  
435 fitness ( $\lambda_{relL}$ ), however, the correlation with lifetime reproductive success was very low and  
436 insignificant ( $r = 0.016$ ,  $t = 0.184$ ,  $P = 0.85$ ). The two estimates of relative individual fitness  
437 ( $\lambda_{rel\Delta N}$  and  $\lambda_{relL}$ ) were strongly associated ( $r = 0.817$ ,  $t = 16.15$ ,  $P < 0.0001$ ). The lack of strong

438 association of the relative indices of individual fitness and lifetime reproductive success indicate  
439 that the widespread use of lifetime reproductive success as a fitness measure might have to be  
440 reconsidered. On the other hand, both individual fitness and lifetime reproductive success  
441 were equally moderate predictors of the number of maternal gene copies passed on to future  
442 generations in collared flycatchers and Ural owls (respectively, *Ficedula albicollis*, *Strix*  
443 *uralensis*; Brommer et al. 2004), and in a study that included both male and female gene copies  
444 in song sparrows (*Melospiza melodia*; Reid et al. 2019). These studies were conducted,  
445 however, without obvious evidence of selection favoring particular traits.

446

447

#### CONCLUDING REMARKS

448 In conclusion, when conducting experiments or looking at annual events, Qvarnström et  
449 al.'s (2006) annual fitness method seems to be a good tool that takes both reproduction and  
450 survival into account. Annual events were illustrated by the experimental ectoparasite (flea)  
451 removal treatment study (Ravel 2015) and response to annual climatic conditions (Lane et al.  
452 2012; Dobson et al. 2016). However, this method may be less appropriate for traits that occur  
453 as part of the ontogenetic sequence of events during the lifespan, or for judging longer-term  
454 success for conditions that have a cumulative influence on fitness. These latter cases were  
455 illustrated by the study of age at maturity (Rubach et al., submitted), and the studies of kin  
456 selection (Viblanco et al. 2010) and inclusive fitness (Dobson et al. 2012). We summarize when  
457 we think these different estimates might be most appropriate in Table 1. Between the two  
458 ways that individual fitness ( $\lambda_{ind}$ ) might be adjusted for changes in population dynamics ( $\lambda_{rel\Delta N}$   
459 and  $\lambda_{relL}$ ), we prefer  $\lambda_{rel\Delta N}$ . This measure is based on changes in population size over time, and

460 includes all competitors (in the present case, all females) in the population. By contrast,  $\lambda_{\text{rel}}$   
461 compares a reproductive female only to others in her cohort, a subset of the population. Yet  
462 over their lifetime, individual in numerous species are exposed not only to individuals from the  
463 same cohort, but to overlapping generations of multiple cohorts. Thus, a fair comparison should  
464 contrast a particular individual against all individuals of the population over her lifetime. In any  
465 case, the choice of empirical fitness measures should be carefully considered, the one most  
466 appropriate to the research question, and amenable for the species under study (e.g., long vs.  
467 short-lived).

468         Other methods for estimating fitness are not directly applicable to the problem of  
469 comparing trait forms in empirical research, but may hold promise for future improvements.  
470 Perhaps the most attractive is the use of offspring from a pedigree, perhaps even including  
471 both male and female relatives (Brommer et al. 2004; Reid et al. 2019). An alternative to our  
472 use of changes in population size to adjust for environmental variation might involve fitness  
473 measures that take demographic and environmental stochasticity into account (e.g., Benton  
474 and Grant 2000; Engen et al. 2009; Sæther and Engen 2015). These measures have not yet  
475 been applied to alternative trait forms. Finally, methods that examine the comparative  
476 sensitivity of population growth to reproduction and survival (e.g., Dobson and Oli 2001; Oli and  
477 Dobson 2003; Coulson et al. 2006) might be used to answer similar questions about individual  
478 fitness.

479

480

## ACKNOWLEDGMENTS

481 Our field research on Columbian ground squirrels has been assisted by many volunteers, under-  
482 graduate and graduate students, and technicians. We express our warmest thanks to all of  
483 them. We thank the Biogeosciences Institute of the University of Calgary provided housing and  
484 laboratory space in the field, especially Direct E. Johnson, Station Managers J. Mappin-  
485 Buchanan and A. Cunnings, and K. Ruckstuhl (faculty organizer and researcher at the R.B.  
486 Miller Research Station). Alberta Ministry of Environment and Paris, Division of Parks provided  
487 permits for research in Sheep River Provincial Park. Fish and Wildlife Division provided permits  
488 for research, as well as capture, marking, and release of wild Columbian ground squirrels. All  
489 field work occurred under several approved protocols by the Institutional Animal Care and Use  
490 Committee of Auburn University, and co-acceptance of these permits by the University of  
491 Calgary. Discussions with colleagues J.E. Lane, P. Neuhaus, and S. Raveh have greatly furthered  
492 our understanding of many aspects of ground squirrel biology. Two anonymous reviewers  
493 provided helpful suggestions on the manuscript. We thank the Institute of Advanced Studies of  
494 the University of Strasbourg for their financial support through an USIAS fellowship for FSD, and  
495 the Région Grand Est and the Eurométropole de Strasbourg for the award of a Gutenberg  
496 Excellence Chair to FSD during the writing of this project. The long-term research was funded  
497 through successive collaborative grants, including a Natural Sciences and Engineering Research  
498 Council of Canada grant to JOM, a National Science Foundation grant (DEB-0089473) to FSD, a  
499 post-doctoral research grant from the AXA Research Fund to VAV, a Fyssen Research grant to  
500 VAV, and a CNRS Projet International de Coopération Scientifique grant (PICS-07143) to VAV.

501

502



503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523

**LITERATURE CITED**

Altmann, S.A. (1991). Diets of yearling female primates (*Papio cynocephalus*) predict lifetime fitness. *Proc. Natl. Acad. Sci. USA* 88, 420-423.

Arnaud, C.M., Dobson, F.S. and Murie, J.O. (2012). Philopatry and within colony movements in Columbian ground squirrels. *Mol. Ecol.* 21, 493-504.

Arnaud, C.M., Becker, P.H., Dobson, F.S., and Charmantier, A. (2013). Canalization of phenology in common terns: genetic and phenotypic variations in spring arrival date. *Behav. Ecol.* 24, 683-690.

Balmer, A., Zinner, B., Gorrell, J.C., Coltman, D.W., Raveh, S., and Dobson, F.S. (2019). Alternative reproductive tactics and lifetime reproductive success in a polygynandrous mammal. *Behav. Ecol.* 30, 474-482.

Benton, T.G., and Grant, A. (2000). Evolutionary fitness in ecology: comparing measures of fitness in stochastic, density-dependent environments. *Evolutionary Ecology Research* 2, 769-789.

Bize, P., Daniel, G., Viblanc, V. A., Martin, J. G., and Doligez, B. (2017). Negative phenotypic and genetic correlation between natal dispersal propensity and nest-defence behaviour in a wild bird. *Biol. Letters* 13, 20170236.

Bonnet, T., and Postma, E. (2018). Fluctuating selection and its (elusive) evolutionary consequences in a wild rodent population. *J. Evol. Biol.* 31, 572-586.

Boonstra, R., and Krebs, C.J. (2012). Population dynamics of red-backed voles (*Myodes*) in North America. *Oecologia* 168, 601-620.

524 Brommer, J.E., Pietiäinen, H., and Kolunen, H. 1998. The effect of age at first breeding on Ural  
525 owl lifetime reproductive success and fitness under cyclic food conditions. *J. Anim. Ecol.*  
526 67, 359–369

527 Brommer, J.E., L. Gustafsson, H. Pietiäinen, and J. Merilä. 2004. Single-generation estimates of  
528 individual fitness as proxies for long-term genetic contribution. *Amer. Nat.* 164, 505-  
529 517.

530 Brommer, J. E., Merilä, J., and Kokko, H. (2002). Reproductive timing and individual fitness. *Ecol.*  
531 *Letters* 5, 802-810.

532 Brommer, J.E., Wistbacka, R., and Selonen, V. (2017). Immigration ensures population survival  
533 in the Siberian flying squirrel. *Ecology and Evolution* 7, 1858-1868.

534 Broussard, D.R., Dobson, F.S., and Murie, J.O. (2008). Previous experience and reproductive  
535 investment of female Columbian ground squirrels. *J. Mammal.* 89, 145-152.

536 Bumpus, H.C. (1899). The elimination of the unfit as illustrated by the introduced sparrow,  
537 *Passer domesticus*. *Biol. Lectures, Woods Hole Marine Biol. Station* 6, 209-226.

538 Chamailé-Jammes, S., Massot, M., Aragon, P., and Clobert, J. (2006). Global warming and  
539 positive fitness response in mountain populations of common lizards *Lacerta vivipara*.  
540 *Global Change Biology* 12, 392-402.

541 Charmantier, A., Garant, D., and Kruuk, L.E.B. (2014). *Quantitative Genetics in the Wild*. Oxford  
542 University Press, New York, New York.

543 Clutton-Brock, T.H. (1988). Reproductive Success. In: *Reproductive Success* (ed. Clutton-Brock,  
544 T.H.). University of Chicago Press, Chicago, pp. 472–485.

545 Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences, 2nd ed.* Lawrence  
546 Erlbaum Associates, Hillsdale, New Jersey.

547 Coulson, T., Benton, T.G., Lundberg, P., Dall, S.R.X., Kendall, B.E., and Gaillard, J.-M. (2006).  
548 Estimating individual contributions to population growth: evolutionary fitness in  
549 ecological time. *Proc. Royal Soc. Lond. B. Biol. Sci.* 273, 547-555.

550 Descamps, S., Boutin, S., Berteaux, D., and Gaillard, J.-M. (2006). Red squirrels trade a long life  
551 for an early reproduction. *Proc. Royal Soc. Biol. Sci.* 273, 2369-2374.

552 Dobson, F.S. (1994). Measures of gene flow in the Columbian ground squirrel. *Oecologia* 100,  
553 190-195.

554 Dobson, F.S., Badry, M.J., and Geddes, C. (1992). Seasonal activity in the Columbian ground  
555 squirrel. *Can. J. Zool.* 70, 1364-1368.

556 Dobson, F.S., Becker, P.H., Arnaud, C.M., Bouwhuis, S., and Charmantier, A. (2017). Plasticity  
557 results in delayed breeding in a long-distance migrant seabird. *Ecology and Evolution* 7,  
558 3100-3109.

559 Dobson, F.S., Lane, J.E., Low, M., and Murie, J.O. (2016). Fitness implications of seasonal climate  
560 variation in Columbian ground squirrels. *Ecology and Evolution* 6, 5614-5622.

561 Dobson, F.S., and Murie, J.O. (1987). Interpretation of intraspecific life history patterns:  
562 evidence from Columbian ground squirrels. *Amer. Nat.* 129, 382-397.

563 Dobson, F.S., and Oli, M.K. (2001). The demographic basis of population regulation in  
564 Columbian ground squirrels. *Amer. Nat.* 158, 236-247.

565 Dobson, F.S., Viblanc, V.A., Arnaud, C.A., and Murie, J.O. (2012). Kin selection in Columbian  
566 ground squirrels: direct and indirect fitness benefits. *Mol. Ecol.* 21, 524-531.

567 Duckworth, R.A., and Kruuk, L.E.B. (2009). Evolution of genetic integration between dispersal  
568 and colonization ability in a bird. *Evolution* 63, 968–977.

569 Endler, J. (1986). *Natural Selection in the Wild*. Princeton University Press, Princeton, New  
570 Jersey.

571 Engen, S., Lande, R., Sæther, B.-E., and Dobson, F.S. (2009). Reproductive value and the  
572 stochastic demography of age-structured populations. *Amer. Nat.* 174, 795-804.

573 Falconer, D.S. (1952). The problem of environment and selection. *Amer. Nat.* 86, 293-298.

574 Fauteux, D., Gauthier, G., and Berteaux, D. (2016). Top-down limitation of lemmings revealed  
575 by reduction of predators. *Ecology* 97, 3231-3241.

576 Grafen, A. (1988). On uses of data on lifetime reproductive success. In: *Reproductive Success*,  
577 pp. 454-471, T.H Clutton-Brock (ed.) University of Chicago Press, Chicago, Illinois.

578 Hare, J.F., and Murie, J.O. (1996). Ground squirrel sociality and the quest for the ‘holy grail’:  
579 does kinship influence behavioral discrimination by juvenile Columbian ground  
580 squirrels? *Behav. Ecol.* 7, 76-81.

581 Hoogland, J.L., and Brown, C.R. (2018). Prairie dogs increase fitness by killing interspecific  
582 competitors. *Proc. Royal Soc. Lond. B. Biol. Sci.* 283, 20160144.

583 Hutchinson, G.E. (1965). *The Ecological Theater and the Evolutionary Play*. Princeton University  
584 Press, Princeton, New Jersey.

585 Jensen, H., Sæther, B.-E. Ringsby, T.H., Tufto, J., Griffith, S.C., and Ellegren, H. (2004). Lifetime  
586 reproductive success in relation to morphology in the house sparrow *Passer domesticus*.  
587 *J. Anim. Ecol.* 73, 599-611.

588 Kettlewell, H.B.D. (1961). The phenomenon of industrial melanism in the Lepidoptera. *Ann.*  
589 *Rev. Ent.* 6, 245-262.

590 King, W.J. (1989). Differential behavior of adult female Columbian ground squirrels. *Anim.*  
591 *Behav.* 38, 354-356.

592 Lande, R., and Arnold, S.J. (1983). The measurement of selection on correlated  
593 characters. *Evolution* 37, 1210-1226.

594 Lane, J.E., Kruuk, L.E.B., Charmantier, A., Murie, J.O., and Dobson, F.S. (2012). Delayed  
595 phenology and reduced fitness associated with climate change in a wild hibernator.  
596 *Nature* 489, 554-557.

597 Lane, J.E., Czenze, Z.J., Findlay-Robinson, R., and Bayne, E. (2019). Phenotypic plasticity and  
598 local adaptation in a wild hibernator evaluated through reciprocal translocation. *Amer.*  
599 *Nat.* 194, 516-528.

600 Lane, J.E., McAdam, A.G., Charmantier, A., Humphries, M.M., Coltman, D.W., Fletcher, Q.,  
601 Gorrell, J.C., and Boutin, S. (2015). Post-weaning parental care increases fitness but is  
602 not heritable in North American red squirrels. *J. Evol. Biol.* 28, 1203-1212.

603 Leslie, P.H. (1945). On the use of matrices in certain population mathematics. *Biometrika* 33,  
604 183-212.

605 McLoughlin, P. D., Gaillard, J.-M. Boyce, M.S., Bonenfant, C., Messier, F., Duncan, P., Delorme,  
606 D., Säid, S., and Klein, F. (2007). Lifetime reproductive success and composition of the  
607 home range in a large herbivore. *Ecology* 88, 3192-3201.

608 Merilä, J., and Sheldon, B.C. (2000). Lifetime reproductive success and heritability in nature.  
609 *Amer. Nat.* 155, 301-310.

610 McGraw, J.B, and Caswell, H. (1996). Estimation of individual fitness from life-history data.  
611 *Amer. Nat.* 147, 47-64.

612 Metcalf, C.J.E., and Pavard, S. (2007). Why evolutionary biologists should be  
613 demographers. *Trends in Ecology & Evolution* 22, 205-212.

614 Mullon, C., Keller, L., and Lehman, L. (2018). Social polymorphism is favoured by co-evolution of  
615 dispersal with social behaviour. *Nature Ecology & Evolution* 2, 132.

616 Murie, J.O., and Harris, M.A. (1982). Annual variation of spring emergence and breeding in  
617 Columbian ground squirrels (*Spermophilus columbianus*). *J. Mammal.* 63:431-439.

618 Murie, J.O., and Harris, M.A. (1988). Social interactions and dominance between female and  
619 male Columbian ground squirrels. *Can. J. Zool.* 66, 1414-1420.

620 Neuhaus, P. (2006). Causes and consequences of sex-biased dispersal in Columbian ground  
621 squirrel, *Spermophilus columbianus*. *Behaviour* 143, 1013-1031.

622 Neuhaus, P., Broussard, D.R., Murie, J.O., and Dobson, F.S. (2004). Age of primiparity and  
623 implications of early reproduction on life history in female Columbian ground squirrels.  
624 *J. Anim. Ecol.* 73, 36–43.

625 Oli, M.K., and Armitage, K.B. (2003). Sociality and individual fitness in yellow-bellied marmots:  
626 insights from a long-term study (1962–2001). *Oecologia* 136, 543–550.

627 Oli, M.K., and Armitage, K.B. (2008). Indirect fitness benefits do not compensate for the loss of  
628 direct fitness in yellow-bellied marmots. *J. Mammal.* 89, 874-881.

629 Oli, M.K., and Dobson, F.S. (2003). The relative importance of life-history variables to  
630 population growth rate in mammals: Cole’s prediction revisited. *Amer. Nat.* 161, 422-  
631 440.

632 Oli M.K., Hepp, G.R., and Kennamer, R.A. (2002). Fitness consequences of delayed maturity in  
633 female wood ducks. *Evol. Ecol. Res.* 4, 563–576.

634 Price, T., and Langen, T. (1992). Evolution of correlated characters. *Trends in Ecology &*  
635 *Evolution* 7, 307-310.

636 Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annu.*  
637 *Rev. Ecol. Evol. Syst.* 37, 637-669.

638 Qvarnström, A., Brommer, J.E., and Gustafsson, L. (2006). Testing the genetics underlying the  
639 co-evolution of mate choice and ornament in the wild. *Nature* 441, 84–86.

640 Raveh, S., Heg, D., Dobson, F.S., Coltman, D.W., Gorrell, J.C., Balmer, A., and Neuhaus, P. (2010).  
641 Mating order and reproductive success in male Columbian ground squirrels (*Urocitellus*  
642 *columbianus*). *Behav. Ecol.* 21, 537-547.

643 Raveh, S., Neuhaus, P., and Dobson, F.S. (2015). Ectoparasites and fitness of female Columbian  
644 ground squirrels. *Philos. Trans. Royal Soc. Lond.* 370, 20140113.

645 Reid, J.M., Neitlisbach, P., Wolak, M.E., Keller, L.F., and Arcese, P. (2019). Individuals' expected  
646 genetic contributions to future generations, reproductive value, and short-term metrics  
647 of fitness in free-living song sparrows (*Melospiza melodia*). *Evolution Letters* 3, 271-285.

648 Roff, D.A. (2002). *Life History Evolution*. Sinauer Associates, Inc., Sunderland, Massachusetts.

649 Roff, D.A., and Fairbairn, D.J. (2012). The evolution of trade-offs under directional and  
650 correlational selection. *Evolution* 66, 2461-2474.

651 Rubach, K.K., Dobson, F.S., Zinner, B., Murie, J.O., and Viblanc, V.A. (*submitted*) Comparing  
652 fitness measures and the influence of age of first reproduction in Columbian ground  
653 squirrels. *J. Mammal.*

654 Sæther, B.-E., and Engen, S. (2015). The concept of fitness in fluctuating environments. *Trends*  
655 *in Ecology & Evolution* 30, 273-281.

656 Stearns, S.C. (1992). *The Evolution of Life Histories*. Oxford University Press, Oxford, U.K.

657 Tafani, M., Cohas, A., Bonenfant, C., Gaillard, J.M., and Allainé, D. (2013). Decreasing litter size  
658 of marmots over time: a life history response to climate change? *Ecology* 94, 580-586.

659 Viblanc, V.A., Arnaud, C.M., Dobson, F.S., and Murie, J.O. (2010). Kin selection in Columbian  
660 ground squirrels (*Urocitellus columbianus*): littermate kin provide individual fitness  
661 benefits. *Proc. Royal Soc. Lond. B. Biol. Sci.* 277, 989-994.

662 Viblanc, V.A., Pasquaretta, C., Sueur, C., Boonstra, R., and Dobson, F.S. (2016). Aggression in  
663 Columbian ground squirrels: relationships with age, kinship, energy allocation, and  
664 fitness. *Behav. Ecol.* 27, 1716-1725.

665 Visser, M.E., and Both, C. (2005). Shifts in phenology due to global climate change: the need for  
666 a yardstick. *Proc. Royal Soc. Lond. B. Biol. Sci.* 272, 2561-2569.

667



668 Figure 1. Spring population size of females in a colony of female Columbian ground squirrels ( $\geq$   
669 1 year old) in Sheep River Provincial Park, Alberta Canada from 1992-2019. (used with  
670 permission, from Rubach et al., submitted)

671

672

673 Figure 2. Influence of mean emergence date from hibernation on mean annual fitness of adult  
674 female Columbian ground squirrels over a period of 20 years. The influence of standardized  
675 emergence date (standardized to a mean of zero and unit variance) on mean relative annual  
676 fitness (individual values divided by the population mean). From a mean-centered selection  
677 analysis, the between-individual selection gradient = -0.14 (highest posterior density interval =  
678 -0.27 to -0.03). (used with permission, from Lane et al. 2012)

679

680

681 Figure 3. Regression of annual fitness on Julian date of snowmelt. Points represent annual  
682 averages, between 1992 and 2012. The trendline reflects the significant negative regression  
683 slope ( $\beta = 0.008 \pm 0.003$  SE,  $R^2 = 0.23$ ,  $F = 5.8$ , d.f. = 1,19,  $P = 0.03$ ). (used with permission, from  
684 Dobson et al. 2016)

685

686

687 Figure 4. Regression of annual fitness on (A) early summer temperature and (B) early summer  
688 rainfall after the young of the year were weaned (late June and July), when all ground squirrels  
689 were fattening for subsequent emergence into hibernation. Points represent daily period

690 averages for 21 different years (1992-2012). Trendlines reflect significant negative and positive  
691 slopes, respectively ( $\beta = 0.12 \pm 0.04$  SE,  $R^2 = 0.18$ , d.f. = 1,19,  $F = 4.1$ ,  $P = 0.05$ ; and  $\beta = 0.12 \pm$   
692  $0.04$  SE,  $R^2 = 0.30$ , d.f. = 1,19,  $F = 8.1$ ,  $P = 0.01$ ). (used with permission, from Dobson et al. 2016)

693

694

695 Figure 5. Fitness estimate for females that first reproduced at different ages, from mean  
696 adjusted lambda (measured as  $\lambda_{relL}$ ) and from mean lifetime reproductive success, both with  
697 standard errors, and reproductive success estimated from number of weaned offspring.

698 Differences among ages were significant for  $\lambda_{relL}$ , but not for lifetime reproductive success  
699 (respectively;  $R^2 = 0.309$ , d.f. = 1,146,  $F = 65.3$ ,  $P < 0.0001$ ;  $R^2 = 0.001$ , d.f. = 1,146,  $F = 0.20$ ,  $P =$

700  $0.65$ ). Samples from 28 years (1992 – 2019) were included. (used with permission, from

701 Rubach et al., submitted)

702

703

704 Figure 6. Estimated inclusive fitness and number of co-surviving close kin (mother-daughter and  
705 littermate sister dyads) that were both of reproductive age (2 years old and older) and  
706 actively reproducing at the same time ( $r = 0.425$ ,  $n = 35$ ,  $P = 0.005$ ; data from 1992-2008).

707 (used with permission, from Dobson et al. 2012)

708

709

710 Figure 7. Comparison of annual fitness values for each reproducing female, and her lifetime

711 adjusted individual fitness (measured as  $\lambda_{rel\Delta}$ ) and lifetime reproductive success (both lifetime

712 estimates calculated from number of offspring at weaning). For annual fitness, all mothers 3-  
713 years old (n = 101), 4-years old (n = 80), 5-years old (n = 59), and 6-years old (n = 32) were  
714 included.

715

716

717 Figure 8. Associations of lifetime reproductive success (LRS) and adjusted individual fitness (A)  
718 based on  $\lambda_{rel\Delta N}$ , and LRS and adjusted individual fitness (B) based on  $\lambda_{relL}$ . The former (A)  
719 exhibits significant association, but the latter (B) does not (see text). Trendlines are presented  
720 for visual inspection.

721

722 Author Contribution Statement

723

724 FSD and VAV designed the study; FSD wrote the manuscript; FSD, VAV, and JOM collected the  
725 data and revised drafts of the manuscript

726

727

728

729 Impact Statement

730

731 Evolutionary biology studies how natural selection operates on traits and combinations of traits  
732 by comparing differences in fitness for individuals that exhibit different trait forms. Thus, how  
733 fitness is measured a key issue for every evolutionary study. We used examples from our past  
734 research on Columbian ground squirrels to compare and contrast methods of measuring  
735 fitness. These include short-term, usually annual measurements, and alternative measures of  
736 lifetime reproduction. In particular, lifetime reproductive success and "individual fitness"  
737 measures (based on matrix methods) have provided conflicting results in past studies. In direct  
738 comparisons using our long-term, 28-year data set, alternative methods exhibited moderate  
739 but disappointing associations with one another. Methods to estimate fitness must be carefully  
740 chosen and considered with caution.