

# Selection and evolution of causally-covarying traits

Michael B. Morrissey

January 31, 2014

4 School of Biology, University of St Andrews

5

## contact

email: michael.morrissey@st-andrews.ac.uk

phone: +44 (0) 1334 463738

fax: +44 (0) 1334 46336

post: Dyers Brae House

School of Biology, University of St

St Andrews, Fife, UK, KY16 9TH

**6 Keywords:** natural selection, causation, selection gradients, path analysis, multivariate,

## 7 secondary theorem of selection

## 8 Abstract

9 When traits cause variation in fitness, the distribution of phenotype, weighted by fitness,  
10 necessarily changes. The degree to which traits cause fitness variation is therefore of central  
11 importance to evolutionary biology. Multivariate selection gradients are the main quan-  
12 tity used to describe components of trait-fitness covariation, but they quantify the direct  
13 effects of traits on (relative) fitness, which are not necessarily the total effects of traits on  
14 fitness. Despite considerable use in evolutionary ecology, path analytic characterizations of  
15 the total effects of traits on fitness have not been formally incorporated into quantitative  
16 genetic theory. By formally defining “extended” selection gradients, which are the total  
17 effects of traits on fitness, as opposed to the existing definition of selection gradients, a  
18 more intuitive scheme for characterizing selection is obtained. Extended selection gradients  
19 are distinct quantities, differing from the standard definition of selection gradients not only  
20 in the statistical means by which they may be assessed and the assumptions required for  
21 their estimation from observational data, but also in their fundamental biological mean-  
22 ing. Like direct selection gradients, extended selection gradients can be combined with  
23 genetic inference of multivariate phenotypic variation to provide quantitative prediction of  
24 microevolutionary trajectories.

## 25 Introduction

26 Natural selection is the phenomenon where effects of traits on fitness necessarily result in  
27 within-generation changes in the distribution of phenotype, weighted by fitness (Godfrey-  
28 Smith, 2007). When heritable traits are selected, and in the absence of antagonistic selec-  
29 tion of genetically correlated traits, the effect of a trait on fitness also results in changes  
30 in the distribution of breeding values. This change of the distribution of breeding values  
31 transmits within-generation phenotypic change to the next generation. This fundamen-  
32 tal evolutionary mechanism has led to a range of approaches and perspectives on how  
33 to explain phenotype-fitness relationships in terms of causal and correlative effects, and  
34 how to quantify the ultimate evolutionary consequences of selection (Endler, 1986; Lande  
35 and Arnold, 1983; Mitchell-Olds and Shaw, 1987; Price, 1970; Robertson, 1966; Schlüter,  
36 1988; Shaw and Geyer, 2010). The main partitioning of selection is the decomposition of  
37 a selection differential  $S$ , the covariance of a trait with relative fitness, into that resulting  
38 from direct effects, as represented by selection gradients  $\beta$ , and correlational effects (Walsh  
39 and Lynch, 2012), resulting from selection of phenotypically-correlated traits. Generally,  
40 selection gradients are characterized as describing the causal effects of a trait on fitness,  
41 i.e., representing “selection for” (Endler, 1986; Sober, 1984), rather than the total associ-  
42 ation of traits and fitness (“selection of”), and so are often the most central parameters in  
43 empirical and theoretical studies of natural selection.

44 Arnold (1983) provided the basis for a thought experiment that can be used to eluci-  
45 date the importance of the distinction between direct and total causal effects of traits on  
46 fitness. Consider two characters: an aspect of morphology, and an aspect of organismal  
47 performance, and also their relationships with fitness. Assume that morphology influences  
48 fitness via an effect on performance, which itself influences fitness, but that morphology  
49 does not affect fitness independently of performance. Figure 1 illustrates these relationships  
50 graphically. Arnold (1983) considered the problem of inference of selection and prediction  
51 of evolution of morphology, especially in light of the fact that it may be hard to simultane-

52 ously and meaningfully measure morphology and fitness on a sufficiently large number of  
53 individuals to make robust inferences in any given single study. Arnold (1983) emphasised  
54 how to use Wright's (1921; 1934) path rules to make inferences of the selection of mor-  
55 phology, given separate inferences about the effects of morphology on performance, and of  
56 performance on fitness. van Tienderen (2000) extended the approach, showing how demo-  
57 graphic principles can be used to evaluate performance (i.e., demographic rates, life history  
58 traits) - fitness relationships, and how to relate these to other traits, such as morphology.  
59 In the morphology-performance-fitness model, a selection gradient of morphology can be  
60 obtained as the product of the coefficients describing the morphology-performance and  
61 performance-fitness relationships. The product of this selection gradient and the genetic  
62 variance of the morphological trait yields a prediction of evolutionary change in perfor-  
63 mance. Multivariate evolutionary prediction follows in a standard manner in order to  
64 predict evolutionary trajectories of multiple aspects of morphology (Arnold, 1983; Lande,  
65 1979). However, simultaneous evaluation of selection and evolution of morphology and  
66 performance is not so straight forward<sup>1</sup>.

67 If morphology influences performance, three important consequences follow. First, the  
68 phenotypic covariance (partial covariance, formally, but these are equivalent in this simple  
69 case) of morphology and performance will be non-zero. Second, the genetic covariance  
70 will be non-zero; essentially, if morphology affects performance, breeding values for the  
71 morphological trait are consequentially a component of the breeding values for performance.  
72 Third, the effect of morphology on fitness will be non-zero, providing that performance  
73 indeed influences fitness. This illustrates two related and potentially non-intuitive features  
74 of selection gradients that necessitate care in their interpretation. First, selection gradients  
75 are not necessarily interpretable simply as 'effects' of a traits fitness. Rather, selection  
76 gradients describe the direct components of effects of traits on fitness. In the morphology-  
77 performance-fitness model, the selection gradient for morphology is zero, if morphology and

---

<sup>1</sup>van Tienderen (2000), page 676, suggests that causal relationships among the set of focal phenotypic traits can be accommodated, but does not provide guidance as to how.

78 performance are considered simultaneously, but the true value is non-zero if performance  
79 is not simultaneously considered. So, second, the selection gradient is partially a function  
80 of the (arbitrary) inclusion of traits that may mediate a focal character's ultimate effect  
81 on fitness.

82 The dependence of selection gradients on the choice (or constraints) of what traits are  
83 included in a study is *not* necessarily a case of selection being erroneously estimated,  
84 i.e., it is distinct from the “missing variable” problem (Hadfield, 2008; Morrissey et al.,  
85 2010; Rausher, 1992). A univariate analysis of selection, genetics, and predicted evolu-  
86 tion of morphology, where the genetic variance of morphology is multiplied by a selection  
87 gradient representing the total regression of relative fitness on morphology, would pro-  
88 vide a correct evolutionary prediction. Similarly, a bivariate analysis, where the genetic  
89 variance-covariance matrix of morphology and performance was post-multiplied by a vec-  
90 tor containing the partial regressions<sup>2</sup>, where the partial regression of relative fitness on  
91 morphology is zero, would yield a correct evolutionary prediction as well. The “missing  
92 variable problem” would occur if an unmeasured variable existed that caused covariance  
93 of morphology and/or performance with fitness, beyond the causal effects of the traits  
94 themselves (Morrissey et al., 2010).

95 Clearly, partitioning total selection, i.e. the selection differential,  $S$ , into direct and  
96 indirect selection neither results in full characterization of the different possible aspects of  
97 relationships among traits and fitness, nor does it match intuition. A selection coefficient  
98 describing the total effect, not simply the direct effect, of a trait on fitness will have  
99 substantial interpretive advantages. Definition of this third selection coefficient, effectively  
100 an “extended-sense” selection gradient  $\eta$ , allows the primary division of types of selection  
101 coefficients to be based on causation, rather than on direct versus indirect effects. As  
102 such, total selection is thought of as the result of causal effects of a focal trait on fitness,  
103 summarized by  $\eta$ , and indirect selection due to incidental correlations.  $\eta$ , the total causal

---

<sup>2</sup>Or more generally, the partial derivatives of relative fitness with respect to the two phenotypic traits, averaged over the distribution of the traits. This is an issue of what selection gradients mean, not an issue of the methodological means by which estimated selection gradients are obtained.

104 effect of a trait on fitness, can then be further considered in terms of its component direct  
105 ( $\beta$ ) and indirect but causal components. In addition to matching intuition about causation,  
106 selection and evolution, empirical evaluation of  $\eta$  for a given focal trait or set of focal traits  
107 (say, morphology) is invariant to whether or not other traits (i.e., performance, life history)  
108 that mediate the focal trait's or traits' ultimate effect(s) on fitness are simultaneously  
109 considered.

110 Path analysis (Wright, 1934) of natural selection provides a means of simultaneously  
111 modelling how traits affect fitness and how phenotypic traits affect one another (Walsh  
112 and Lynch, 2013). As such, path analysis can provide insights, both quantitative and  
113 qualitative, into the mechanisms by which phenotypic traits cause fitness variation (Latta  
114 and McCain, 2009; Scheiner et al., 2000). The morphology-performance-fitness model is a  
115 simple path model. The procedure of obtaining the total effect of morphology on fitness as  
116 the product of the regressions of fitness on performance and performance on morphology  
117 is a simple application of Wright's path rules. The qualitative benefits of a path analytic  
118 perspective for making inferences about natural selection have been discussed from several  
119 perspectives (Arnold, 1983; Conner, 1996; Crespi and Brookstein, 1989; Kingsolver and  
120 Schemske, 1991; Latta and McCain, 2009; Scheiner et al., 2000; Shipley, 1997). While  
121 these authors have appreciated and clearly demonstrated the value of characterizing the  
122 causal effects of traits on one another and on fitness, the distinction between compound  
123 path-based selection coefficients and (traditional, direct) selection gradients has not been  
124 made clear. Consequently, some conclusions have been drawn based on the notion that  
125 path-based inferences of selection and traditional selection gradients represent different in-  
126 ferences (statistical, philosophical, or both) of the same biological quantity (Scheiner et al.,  
127 2002, 2000), but this is not the case. In addition to the previous lack of formal consider-  
128 ation of the mathematical and philosophical distinctions and commonalities between path  
129 coefficients and selection gradients in the traditional sense, the role of path coefficients in  
130 quantitative genetic theory has not yet been formally considered.

131 I show how path analysis-based extended selection gradients relate quantitatively to

132 genetic variation and evolutionary change by deriving an equation that relates extended se-  
 133 lection gradients to genetic variation in order to quantitatively predict evolutionary change.  
 134 I then provide two examples of the estimation and interpretation of extended selection gra-  
 135 dients in an evolutionary quantitative genetic context. In the first, I present a comparison  
 136 of extended and direct selection gradients of Soay sheep *Ovis aries* (Clutton-Brock and  
 137 Pemberton, 2004) neonatal traits. This provides a simple situation where the biological  
 138 meanings of the traits, and of their relationships with fitness, are fairly intuitive, allowing  
 139 illustration of the interpretive differences between  $\beta$  and  $\eta$ . I then show the incorpora-  
 140 tion of the path analytic approach into both the decomposition of phenotypic and genetic  
 141 (co)variances, and the simultaneous quantification of selection gradients, using data from  
 142 a laboratory rearing experiment based on a population of recombinant inbred lines derived  
 143 from contrasting ecotypes of the wild oat *Avena barbata* (Gardner and Latta, 2008; Latta  
 144 and McCain, 2009). The examples demonstrate (i) how  $\beta$  and  $\eta$  can differ qualitatively,  
 145 including how they can take different signs, and (ii) consequently how biological interpre-  
 146 tations that are typically sought regarding the selective meaning of trait variation must be  
 147 assessed via the extended view of selection gradients.

## 148 Multivariate evolutionary prediction using extended selection gra- 149 dients

150 Expected evolutionary change based on (path coefficient-based) estimates of extended se-  
 151 lection gradients can be obtained starting with the Lande equation (Lande, 1979),

$$\Delta \bar{z} = \mathbf{G}\beta, \quad (1)$$

152 where  $\Delta \bar{z}$  is the expected per-generation change in the vector of mean phenotype,  $\mathbf{G}$   
 153 is the matrix of additive genetic variances and covariances, and  $\beta$  is a vector of direct  
 154 selection gradients, i.e., the average partial derivatives of relative fitness integrated over

155 the distribution of the phenotype. In path analytic terms,  $\beta$  are the coefficients associated  
 156 with arrows directly between traits and relative fitness. To express the rest of the formula  
 157 in terms of path coefficients,  $\mathbf{G}$  needs to be related to causal effects of traits on one another  
 158 (path arrows among traits). Given a matrix of path coefficients  $\mathbf{b}$ , the total causal effects  
 159 of each trait on every other trait are

$$\Phi = (\mathbf{I} - \mathbf{b})^{-1}. \quad (2)$$

160 Following McArdle and McDonald (1984) and Gianola and Sorensen (2004),  $\mathbf{G}$  is deter-  
 161 mined in part by  $\Phi$  according to

$$\mathbf{G} = \Phi \mathbf{G}_\epsilon \Phi^T, \quad (3)$$

162 where  $\mathbf{G}_\epsilon$  represents the additive genetic component of sources of variance and covariance  
 163 among traits, beyond those attributable to causal relationships among traits. Diagonal  
 164 elements of  $\mathbf{G}_\epsilon$  represent the additive genetic components of exogenous inputs of variation  
 165 to a system of structural equations, often denoted  $U$  on path diagrams. Off-diagonal  
 166 elements of  $\mathbf{G}_\epsilon$ , if any, represent the additive genetic component of covariances that are  
 167 extrinsic to causal relations, often denoted with curved double-headed arrows on path  
 168 diagrams.

169 Substitution of equation 3 into equation 1 gives  $\Delta \bar{z} = \Phi \mathbf{G}_\epsilon \Phi^T \beta$ . Within this expres-  
 170 sion, the extended selection gradients,  $\eta$ , or total effects of each trait on relative fitness,  
 171 are  $\eta = \Phi^T \beta$ . In scalar form, this is  $\eta_x = \sum_{y=1}^n \Phi_{xy} \beta_y$ , defining the total effects on fitness  
 172 as the sum of the products of the effects of the traits on one another and on relative fitness.  
 173 So the evolution of the mean vector in terms of extended selection gradients is

$$\Delta \bar{z} = \Phi \mathbf{G}_\epsilon \eta. \quad (4)$$

174 It remains to consider how exogenous genetic variances and covariances are to be ob-  
 175 tained.  $\mathbf{G}_\epsilon$  and its components are not generally considered among the parameters of

176 interest in evolutionary quantitative genetics, but have specific evolutionary meaning and  
177 are obtainable through modifications of familiar mixed-model techniques (Henderson, 1973;  
178 Kruuk, 2004; Wilson et al., 2010). Standard structural equation modelling packages (e.g.,  
179 SEM; Fox 2006, LISREL; Joreskog and Van Thillo 1972) for implementing path analyses  
180 intrinsically estimate total exogenous variances and covariances, even if these are not typ-  
181 ically considered parameters of particular interest. The key to the decomposition of the  
182 total exogenous (co)variances into genetic and residual components is to view a path model  
183 as a system of mixed model equations. The twist, however, is that any trait that has a  
184 effect on any other trait is part of the response (i.e., its value is modelled), and also serves  
185 as a predictor of the observed values of other traits. If there are neither simultaneous (e.g.,  
186  $A \rightarrow B, B \rightarrow A$ ) relationships nor recursive loops (e.g.,  $A \rightarrow B, B \rightarrow C, C \rightarrow A$ ), the com-  
187 ponents of  $\mathbf{b}$  and  $\mathbf{G}_\epsilon$  can be estimated from separate mixed models describing parts of the  
188 path model. Path coefficients are simply continuous fixed effects, and exogenous variances  
189 are obtained as random effects, conditional on any fixed effects representing path coeffi-  
190 cients. If a path model involves exogenous covariances, then components of  $\mathbf{G}_\epsilon$  pertaining  
191 to these variables would be estimated using a multi-response mixed model. Decomposition  
192 of exogenous (co)variances into genetic, residual, and potentially other components can be  
193 implemented using standard mixed model techniques. For example, using general pedi-  
194 gree information, additive genetic exogenous variance components can be estimated using  
195 animal models (Henderson, 1973; Kruuk, 2004; Wilson et al., 2010) in which fixed effects  
196 are included to estimate path coefficients. All component mixed models must be simul-  
197 taneously evaluated in path models that contain recursive or simultaneous relationships  
198 (Gianola and Sorensen, 2004), but such features of path models do not generally appear  
199 in studies of natural selection.

## 200 Example 1: Selection of neonatal traits in Soay sheep

201 The purpose of this example is (*i*) to consider the differences between estimates of  $\beta$  and  
 202  $\eta$  in the context of covariances among traits in a real dataset, and (*ii*) to consider the  
 203 ways in which interpretations about natural selection can be made given estimates of  $\beta$   
 204 and  $\eta$ . Here, I consider the relations among birth date, twin status, birth weight, weight  
 205 in August, and their selection via relationships with relative fitness in the first year of life  
 206 of female Soay sheep lambs on St Kilda, Outer Hebrides, Scotland, during the period of  
 207 1985 to 2009. The fitness component is overwinter survival. In total, the analysis was  
 208 conducted on complete records of all traits and overwinter survival for 1284 individuals.  
 209 More detail about the study system is available in Clutton-Brock and Pemberton (2004).

210 Covariances among birth date, twin status, birth weight, August weight, and relative  
 211 fitness (overwinter survival scored as 0 and 1, divided by year-specific mean survival) are  
 212 given in table 1. A plausible model relating the traits to one another and to relative fitness  
 213 is

$$w_i = \mu_w + b_{w,awt} \cdot awt_i + b_{w,bwt} \cdot bwt_i + b_{w,twn} \cdot twn_i + b_{w,bdy} \cdot bdy_i + e_i(w) \quad (5a)$$

$$awt_i = \mu_{awt} + b_{awt,bwt} \cdot bwt_i + b_{awt,twn} \cdot twn_i + b_{awt,bdy} \cdot bdy_i + e_i(awt) \quad (5b)$$

$$bwt_i = \mu_{bwt} + b_{bwt,twn} \cdot twn_i + b_{bwt,bdy} \cdot bdy_i + e_i(bwt) \quad (5c)$$

214 where  $w$  represents relative fitness,  $awt$  represents weight in August (kg),  $bwt$  represents  
 215 birth weight (kg),  $twn$  represents twin status (scored as zero or one), and  $bdy$  represents  
 216 birth date (day of the year).  $i$  indexes individuals,  $e_i$  terms are residual errors of the  
 217 bracketed quantities, and  $\mu$  are intercepts. I evaluated the three multiple regressions in  
 218 equation 5 using MCMCglmm (Hadfield, 2010). This allowed statistical uncertainty in  
 219 both direct and extended selection gradient estimates to be evaluated by integration over  
 220 the joint posterior distributions of the solutions to equation 5.

221 The estimates of  $\beta$  from equation 5a, the fixed components of which are essentially

222 Lande and Arnold's (1983) multiple regression analysis for directional (direct) selection  
223 gradients, and estimates of  $\eta$  obtained by applying path rules to coefficients obtained from  
224 equations 5a-c, are given in table 2. Figure 2 shows the relationships described in equation  
225 5 as a path diagram, with representation of path strengths (variance-standardized, i.e.  
226 partial correlations, except for regressions of relative fitness on the traits, which are the  
227 partial regressions of  $w$  on the variance-standardized traits) as thickness of arrows.

228 August weight has a substantial direct effect on fitness, while the other traits have  
229 smaller direct effects (table 2, figure 2). However, in this model, twin status and birth  
230 weight have effects as well on fitness as well, but they are largely indirect. Birth weight  
231 has a positive effect on fitness via its effect on August weight, jointly with the fact that  
232 August weight affects fitness. Similarly, while twin status has little or no direct effect  
233 on fitness, it does have negative effects on both birth weight and August weight, and  
234 consequently a negative total, if mostly indirect, effect on fitness.

235 The interpretation of extended selection gradients is well-illustrated by this example.  
236 Twin status and birth weight have very small direct influence on fitness, and therefore  
237 small  $\beta$ . Insofar as it is reasonable to assume that these traits may have causal effects  
238 on August weight and fitness, it is very worth quantifying the total effect of this trait  
239 on fitness if we are trying to understand the adaptive significance of variation in birth  
240 weight.  $\eta$  most closely reflects the concept of "selection for" (Endler, 1986; Sober, 1984)  
241 birth weight, as it reflects the selective significance of birth weight in a way that existing  
242 selection coefficients do not.

243 These results do not necessarily represent a comprehensive study of selection of lamb  
244 traits via variation in first year overwinter survival in female Soay sheep. For example, the  
245 study population experiences substantial environmental variation with respect to popula-  
246 tion density, food availability and weather (Clutton-Brock and Pemberton, 2004; Coulson  
247 et al., 2001), and the relationships among traits and between traits and fitness may vary in  
248 important ways with environmental conditions (Catchpole et al., 2000). I present this ex-  
249 ample as a simple calculation of  $\eta$  given a path diagram, and of the interpretive differences

250 between  $\eta$  and  $\beta$ .

251 **Example 2: Path model-based simultaneous inference of selection  
252 and genetics in wild oats**

253 The purpose of this section is to illustrate the simultaneous estimation of extended selection  
254 gradients and genetic and residual exogenous variances for the purpose of quantitative  
255 path model-based microevolutionary prediction. The focal dataset in this section is from  
256 Gardner and Latta's (2008) experiment in which recombinant inbred lines of wild oat *Avena*  
257 *barbata*, derived from contrasting ecotypes, were grown in the lab to evaluate relationships  
258 among a number of phenological, vegetative, and reproductive traits. The experimental  
259 design using inbred lines greatly simplifies the statistical inference of genetic parameters.  
260 However, the (co)variances among individuals attributable to line must be interpreted as  
261 broad-sense genetic parameters, i.e., representing total genetic effects, not only additive  
262 genetic effects. Extension to analysis of classical breeding designs (Lynch and Walsh, 1998)  
263 and general pedigrees (Gianola and Sorensen, 2004; Henderson, 1973; Kruuk, 2004; Wilson  
264 et al., 2010) is relatively straight forward (see section 2), once the basic principle is clear.

265 Closely following Latta and McCain (2009), I adopted the path model structure in figure  
266 3 as an *a priori* set of causal assumptions about covariances among the phenological, veg-  
267 etative, and reproductive traits, and relative fitness, based on the number of reproductive

268 spikes. The set of mixed models characterizing this causal scheme is

$$w_i = \mu_7 + b_{7,4} \cdot mass_i + b_{7,5} \cdot rpt_i + b_{7,6} \cdot mrt_i + g_{\epsilon, line_i}(w) + e_{\epsilon,i}(w), \quad (6a)$$

$$mrt_i = \mu_6 + b_{6,4} \cdot mass_i + b_{6,3} \cdot dtf_i + g_{\epsilon, line_i}(mrt) + e_{\epsilon,i}(mrt), \quad (6b)$$

$$rpt_i = \mu_5 + b_{5,4} \cdot mass_i + b_{5,3} \cdot dtf_i + g_{\epsilon, line_i}(rpt) + e_{\epsilon,i}(rpt), \quad (6c)$$

$$mass_i = \mu_4 + b_{4,2} \cdot m60_i + g_{\epsilon, line_i}(mass) + e_{\epsilon,i}(mass), \quad (6d)$$

$$dtf_i = \mu_3 + b_{3,2} \cdot m60_i + g_{\epsilon, line_i}(dtf) + e_{\epsilon,i}(dtf), \quad (6e)$$

$$m60_i = \mu_2 + b_{2,1} \cdot dgerm_i + g_{\epsilon, line_i}(m60) + e_{\epsilon,i}(m60), \quad (6f)$$

$$dgerm_i = \mu_1 + g_{\epsilon, line_i}(dgerm) + e_{\epsilon,i}(dgerm), \quad (6g)$$

269 where the traits are (numerical indices for model term subscripts in brackets): (1) days  
 270 to germination, *dgerm*, (2) mass on day 60, *m60*, in grams, (3) days to first flower, *dtf*,  
 271 (4) final total mass, *mass*, in grams, (5) number of reproductive tillers, *rpt*, (6) combined  
 272 mass of reproductive tillers, *mrt*, in grams, and (7) relative fitness, *w*. I obtained relative  
 273 fitness by dividing the number of seed spikes (each spike contains two seeds) by the mean  
 274 number of spikes. For consistency with Latta and McCain (2009), I standardized each  
 275 trait observation by subtracting block averages (the experimental rearing was conducted  
 276 in three blocks) prior to the mixed model analyses.  $\mu$  are intercepts, and  $b_{jk}$  values are  
 277 partial regression coefficients, where  $j$  indexes response variables and  $k$  indexes predictors.  
 278  $g_{\epsilon, line_i}(k)$  are the trait ( $k$ )-specific exogenous genetic values of each line, and are assumed  
 279 to be drawn from normal distributions with estimated variances  $g_{\epsilon, line_i}(k) \sim N(0, \sigma_{\epsilon, g, k}^2)$   
 280 where  $N(0, \sigma^2)$  represents a normal probability distribution with mean 0 and variance  
 281  $\sigma^2$ .  $g$ , for genetic value, replaces  $a$ , for breeding value, above, in the typical notation  
 282 of the genetic effects, simply because the estimated parameters, given the inbred line-  
 283 based experiment, are broad-sense (exogenous) genetic values rather than additive genetic  
 284 effects. Similarly,  $e_{\epsilon,i}(k)$  are residuals, drawn from normal distributions with trait-specific  
 285 estimated variances, i.e.,  $e_{\epsilon,i}(k) \sim N(0, \sigma_{\epsilon, e, k}^2)$ . As for equation 5, I evaluated each multiple

286 regression mixed model specified by equation 6 separately using MCMCGLMM (Hadfield,  
287 2010).

288 Conditional on the structure of the path model defined by equation 6, the estimates of  
289  $\mathbf{b}$  and  $\Phi$  are given in table 3. The genetic and residual variance-covariance matrices (and  
290 ultimately the phenotypic variance-covariance matrix, their sum), obtained using equation  
291 3, and equivalently,  $\mathbf{R} = \Phi \mathbf{R}_\epsilon \Phi^T$  are given in supplemental table S2. These estimated  
292 variance-covariance matrices generally match previously-reported genetic parameters from  
293 this experiment (Gardner and Latta, 2008), as well as a mixed model-based estimate of  
294 the genetic variances and covariances made without any assumptions (i.e., without the  
295 path model; supplemental table S3), using a multi-response mixed model-based analysis  
296 to estimate the covariance matrix associated with line and the residual covariance matrix,  
297 using MCMCGLMM. However, the path analysis-based estimates of the matrices generally  
298 contained estimates of individual covariance components that are smaller in magnitude  
299 than the unconstrained estimates, for variances and covariances involving mass, number of  
300 reproductive tillers, mass of reproductive tillers, and relative fitness.

301 I obtained extended selection gradient estimates by application of equation 2 to the  
302 estimate of  $\mathbf{b}$  from equation 6, and obtained credible intervals by integrating this analysis  
303 over the posterior distribution of the solution to equation 6. Path model-based inference  
304 of direct and extended selection gradients revealed negative total effects of the two pheno-  
305 logical traits on fitness, and positive total effects of the vegetative and reproductive traits.  
306 Trivially, the path-based estimate of the direct effects of number and mass of reproductive  
307 tillers on fitness were also positive, because  $\boldsymbol{\eta}$  and  $\boldsymbol{\beta}$  are identical for these traits, given  
308 the path model (figure 3, equation 6). The direct effect of mass on fitness is negative.

309 To compare the path-based estimate of  $\boldsymbol{\beta}$  with unconstrained estimates, I estimated  $\boldsymbol{\beta}$   
310 by multiple regression of spike number on the other six traits. For this I fitted a model  
311 directly analogous to equation 6a, but including partial regressions of relative fitness on all  
312 other traits, and without the estimate of the among-line variance of relative fitness. For  
313 the traits with non-zero  $\beta$  as defined by the path model, the path-based and unconstrained

314 estimates of  $\beta$  are similar. Unconstrained inference of  $\beta$  suggests a negative direct effect  
315 of days to first flower on fitness (table 3).

316 As in the Soay sheep example, differences between  $\beta$  and  $\eta$  in the wild oats illustrate  
317 important ways in which formalization the path-analysis perspective into evolutionary  
318 quantitative genetic inferences yields insight into selective mechanisms. For mass at day 60  
319 and days to first flower,  $\eta$ , suggests much more substantial selection than does  $\beta$ . Selection  
320 of mass is particularly interesting, as the two types of selection gradients have different  
321 estimated signs. Except insofar as individuals with greater mass may have greater mass of  
322 reproductive organs, total vegetative mass is (trivially) not itself a component of fecundity.  
323 Since fecundity variation is the only source of fitness variation in this greenhouse-based  
324 experimental system, the portion of the effect of mass on fitness that is independent of  
325 effects acting via fecundity is unlikely to be positive. Because non-reproductive structures  
326 must be maintained, they must be costly in-and-of themselves, and so the direct selection  
327 gradient of mass is negative. However, individuals with greater total mass also have greater  
328 reproductive capacity, and so the extended selection gradient of mass, i.e., the total causal  
329 effect of mass on fitness, is positive.

330 To compare evolutionary predictions based on extended selection gradients with alter-  
331 native approaches to evolutionary prediction, I made predictions of microevolution based  
332 on application of the Lande equation  $\Delta\bar{z} = \mathbf{G}\beta$  (Lande, 1979), and on the secondary  
333 theorem of selection, whereby expected evolutionary change is the genetic covariance of  
334 each trait with relative fitness  $\Delta\bar{z} = \sigma_a^2(z, w)$  (Morrissey et al., 2010; Robertson, 1966). I  
335 estimated  $\mathbf{G}$  (broad-sense genetic variances and covariances) as the among-line covariance  
336 matrix using a multi-response mixed model treating the six traits other than fitness (for the  
337 Lande equation) or all seven traits (for the secondary theorem of selection) as dependent  
338 variables.

339 All three systems of evolutionary prediction yield qualitatively similar results (figure  
340 4). Based on all three systems of prediction, little evolution of days to germination and  
341 mass at day 60 is expected, days to first flower is expected to advance, a modest increase

342 in total mass is expected, and finally, substantial evolution of greater number and mass of  
343 reproductive tillers is expected. In general, the predictions based on the Lande equation  
344 and the secondary theorem of selection are greater in magnitude than those based on  
345 the path analysis of extended selection gradients (figure 4). The smaller predictions of  
346 evolutionary change based on the path model seem to be due to lower path model-based (co)  
347 variance estimates (supplemental tables S2 and S3), rather than any substantial differences  
348 in trait-fitness relationships (table 3).

## 349 Discussion

350 Extended selection gradients provide a means of quantitatively summarizing selection that  
351 reflects the concept of “selection for” (Endler, 1986; Sober, 1984), i.e., they reflect the  
352 total dependence of relative fitness on variation in a trait. The example analyses of Soay  
353 sheep and wild oat data illustrate scenarios where total and direct effects of traits on  
354 fitness differ in important ways. The inferred effect of sheep birth mass on fitness might  
355 be relegated to indirect selection of a mere correlated trait, if only  $\beta$  was considered.  
356 Similarly, the positive covariance of oat plant mass and fitness might also be relegated to a  
357 case of indirect selection where the positive relationship is an indirect result of selection for  
358 reproductive traits. Such conclusions would represent, at best, incomplete interpretations  
359 of the selective consequences of variation in Soay sheep birth weight and wild oat plant  
360 mass.

361 Direct integration of an hypothesis about the mechanism of selection into the statistical  
362 mechanics of the estimation of genetic and phenotypic variances and covariances has several  
363 potential benefits, but also necessitates careful interpretation and explicit consideration  
364 of the associated assumptions. First, two potential misconceptions must be addressed.  
365 Extension of a causal model of phenotypic covariance among traits to the genetic level does  
366 not require any additional assumptions beyond those that are involved in application of  
367 path analysis at the phenotypic level. If trait *A* causes variation in trait *B*, then the *partial*

368 genetic and phenotypic regressions of  $B$  on  $A$  are the same (see Hadfield 2008; Morrissey  
369 et al. 2010; Queller 1992; Robertson 1966 for further discussion of the manifestation of  
370 causation as equivalent genotypic and phenotypic partial regressions). Note that this is  
371 only true for the partial regressions - the action of other traits or of environmental variation  
372 might make the total genetic and phenotypic regressions different, and failure to account  
373 for all the contributors to covariances among traits may result in erroneous estimation of  
374 any focal partial regression parameters, just as in any selection analysis (Hadfield, 2008;  
375 Morrissey et al., 2010; Rausher, 1992; Robertson, 1966).

376 Second, equivalence of phenotypic and genetic partial regressions does not imply equiv-  
377 alence, or even common signs, of phenotypic and genetic covariances and correlations. The  
378 magnitudes and signs of phenotypic and genetic correlations are determined jointly by  
379 the partial regressions and the relative magnitudes of the genetic and non-genetic compo-  
380 nents of the exogenous (co)variances of traits. Consider, for example, a situation in which  
381 a trade-off occurs between two heritable traits (perhaps a trade-off between life history  
382 traits). This could be manifested as a negative partial regression of one trait on the other.  
383 However, if the values of the two traits are both partially determined by a third trait  
384 (perhaps resource availability or acquisition rate; this generates a model very similar to de  
385 Jong and van Noordwijk's (1992) model of resource acquisition and allocation), then they  
386 may covary positively despite the inherent trade-off. In a situation where the third trait  
387 is highly variable but not heritable, it could cause a positive overall phenotypic covariance  
388 between the first two traits while they could covary negatively at the genetic level, even  
389 though the phenotypic and genetic partial regressions among all the traits are equal.

390 Path-analytic estimates of genetic variance-covariance matrices will generally be (sta-  
391 tistically) more precise than unconstrained estimates of genetic parameters. Consequently,  
392 evolutionary predictions based on  $\eta$  will be estimable with less sampling variance (i.e.,  
393 smaller standard errors). This effect may generally be dramatic, because path-based es-  
394 timation of  $\mathbf{G}$  uses information about the partial regressions of traits on one another,  
395 obtained from phenotypic data in conjunction with an *a priori* causal model of trait co-

396 variance. The extent to which the statistical precision of path-based estimation of  $\mathbf{G}$  is  
397 justified depends on the validity of the path model. Essentially, statistical uncertainty is  
398 traded against the validity of assumptions. Under the assumption that the wild oat path  
399 model (figure 3, equation 6) represents a valid causal explanation of the covariances among  
400 the traits, the standard deviations of the posterior distributions (interpretable as similar  
401 to standard errors) of the elements of  $\mathbf{G}$  (supplemental table S2a) are about half of what  
402 they are based on unconstrained estimation of  $\mathbf{G}$  (supplemental table S3), and uncertainty  
403 in evolutionary predictions based on  $\eta$  is correspondingly smaller as well (figure 4).

404 Incorporation of path analysis into evolutionary quantitative genetic theory generates  
405 a new system of evolutionary prediction that is statistically and philosophically distinct  
406 from the breeder's and Lande equations (Lande, 1979; Lush, 1937), and from the secondary  
407 theorem of selection (Robertson, 1966). Path analysis-based evolutionary prediction relies  
408 most heavily on *a priori* assumptions of the causal nature of phenotype-fitness covariance.  
409 Evolutionary prediction based on the breeder's equation assumes that all traits directly  
410 responsible for multivariate phenotype-fitness covariances are identified, meaningfully mea-  
411 sured, and adequately modelled, but makes no assumptions about the causal structure of  
412 phenotypic and genetic relationships among traits. Finally, evolutionary prediction based  
413 on the secondary theorem of selection (Etterson and Shaw, 2001; Morrissey et al., 2010,  
414 2012; Price, 1970; Robertson, 1966, 1968) does not require that all, or indeed any, causal  
415 sources of trait-fitness covariance are identified, nor does it make any assumptions about  
416 the causal structure of phenotypic or genetic covariation among traits or between traits  
417 and fitness.

418 The three systems for evolutionary prediction (ordered as above, i.e., path, breed-  
419 ers/Lande, and secondary theorem) vary in three more practical aspects: (1) This order  
420 represents decreasing statistical precision of evolutionary predictions when all the assump-  
421 tions of each system are met. (2) This order represents decreasing risk of erroneous pre-  
422 dictions when the assumptions are not met. And (3), this order represents decreasing  
423 level of insight into the mechanisms of natural selection. In fact, the secondary theorem

of selection provides a prediction of evolutionary change, but yields almost no insight into natural selection: genetic covariances of a trait and fitness may be due to selection of those traits, selection of other genetically correlated traits (measured or not), or may be due to drift, population structure, or variation in accumulated mutation (in the last two cases the covariation of traits with fitness may nonetheless be reasonably characterized as selection). Robertson's theorem could be considered a primary quantitative genetic theorem of evolution, neither necessarily nor specifically of selection.

Issues pertaining to fundamental meaning, as opposed to the inference, of causal mechanisms of selection must be kept distinct. First, understanding the mechanistic, i.e., causal, basis of natural selection can bring an understanding of natural selection that statistical quantification of trait fitness relationships cannot provide alone. Whether one is interested in direct selection gradients (direct causal effects), extended selection gradients (total causal effects), or selection differentials (covariance arising from selective processes), each of these parameters is in some way a reflection of a causal process (Godfrey-Smith, 2007; Sober, 1984). Inference of selection gradients relies on the existence of a correct causal model of the mechanism underlying trait-fitness covariance. For direct selection gradients, this model of direct effects of traits on fitness is implicit in the concept of partial derivatives of relative fitness with respect to phenotype, which in practice is normally assessed by multiple regression. Failure to include traits that covary with focal traits, and that cause fitness variation amounts to applying an incorrect model of direct effects of traits on fitness. This type of 'wrong model' problem, arising from missing traits, is well discussed (Holland, 2008; Kruuk et al., 2003; Morrissey et al., 2010; Rausher, 1992; Robertson, 1966). Inference of extended selection gradients similarly requires that all factors that ultimately cause focal traits to covary with fitness be measured, and additionally, requires that a valid scheme by which to relate their causal effect on one another and on fitness is available. It must be kept in mind that the fit of observational data to a causal model of any kind (whether it be a multiple regression model, or a more complex causal hypothesis) provides only the weakest kind of inference about the validity of the model. Wright (1934) describes

452 this very well:

453 “In considering the reliability of path coefficients there are two questions which  
454 must be kept distinct. First is the adequacy of the qualitative scheme to which  
455 the path coefficients apply and the second is the reliability of the coefficients, if  
456 one accepts the scheme as representing a valid point of view. The setting up of  
457 a qualitative scheme depends primarily on information outside of the numerical  
458 data and the judgement as to its validity must rest primarily on this outside  
459 information. One may determine from standard errors whether the observed  
460 correlations are compatible with the scheme and thus whether it is a possible  
461 one, but not whether it correctly represents the causal relation.”

462 The current work (*i*) highlights why the causal structure of trait-fitness relationships mat-  
463 ters for making inferences about natural selection, and (*ii*) derives the statistical quan-  
464 titative genetic mechanics that relate causal schemes to selection, genetic variation, and  
465 evolutionary change. The current work does not provide any recipe for determining the  
466 causal structure of trait-fitness relationships. Inferences of  $\eta$  in any given application will  
467 vary with different assumed causal structures, but this does not mean that  $\eta$  is in any way  
468 arbitrary: there will be a correct causal structure that yields correct inference of extended  
469 selection gradients. As Wright points out (quote above), observational data such as that  
470 typically used for quantitative genetic inference of selection provides only the weakest kind  
471 of test of the adequacy of causal hypotheses. However, the necessity of understanding the  
472 causal structure of trait-fitness covariance could indeed benefit from a range of different  
473 kinds of information about causal relationships. Logical decisions based on chronology,  
474 natural history, existing theory, and experiment could all in principle be brought to bear.  
475 For example, in the sheep example, I considered all relationships among traits plausible;  
476 some effects may be small, but rather than exclude them *a priori*, I allowed them to be  
477 estimated as small values. As such the sheep analysis can be seen as a contrivance to  
478 exploit the least restrictive possible path model, guided only by a linear view of time and

479 causality.

480 It is unlikely that relationships among measured variables in any study system will ever  
481 completely reflect all causes of covariance. With careful consideration of the biology of  
482 any given study system, it is plausible that relationships among measured variables could  
483 often reflect the major causes of covariance, but in general, unmeasured traits and aspects  
484 of the environment will generally also cause covariance among measured quantities. The  
485 consequences of this simple and realistic view of empirical data have profound implications  
486 for what can be achieved using the many existing procedures in the path analysis liter-  
487 ature for assessing the fit of different models to the same dataset. In particular, in the  
488 presence of modest effects of unmeasured variables, essentially correct causal structures  
489 (among measured quantities) may appear to be preferred when modest amounts of data  
490 are available, but with increasing data, there will be a tendency for indices of statistical fit  
491 to lead to preference of more complicated models, i.e., models that contain effects that do  
492 not exist, but reflect spurious associations due to unmeasured quantities. This principle,  
493 where data-driven analytical decisions, especially in frequentist analytical frameworks, will  
494 generally result in preference for overly complex and wrong models, applies to statistical  
495 modelling in general, not just to path analysis.

496 This is not to say that assessing fit is irrelevant. Rather, what one does with information  
497 about fit is what matters. Under the assumption that each wild oat plant is independent  
498 (it is not, as each belongs to an inbred line), the residual mean squared error of approxi-  
499 mation (Steiger, 1990) is 0.109 (90% CI: 0.085-0.133), which by most arbitrary thresholds  
500 indicates a marginal fit, and Bentler's (1990) comparative fit index is 0.966, which is quite  
501 good (indices of fit from SEM Fox 2006, based on fitting the model in equation 6, but  
502 without accounting for inbred line). The  $\chi^2$  value arising from the difference between the  
503 covariance structure implied by the fitted path model, and the observed covariances is 80.0,  
504 which on 11 degrees of freedom (the covariance matrix of the seven variables has 28 unique  
505 elements, minus the number of free parameters, which include 7 exogenous variances and  
506 10 partial regression coefficients) indicates that a more complex model could provide highly

507 statistically significantly better fit. Note that assessment of fit in these ways is not relevant  
508 to the Soay sheep example, as it is based on a saturated model. Imperfect fit *may* indicate  
509 that there are paths that should be added to the model, or it may indicate the presence  
510 of some unmeasured variable. If a path model is well considered, the latter will often be  
511 the case. A statistical solution will sometimes be available via fitting latent variables. La-  
512 tent variables are not directly considered here, but their use is common in path analysis,  
513 and the quantitative genetic principles pertaining to systems of causally covarying traits  
514 should be relatively easily extended to models that include latent variables. More usefully,  
515 imperfect fit could be used to inform future data collection, or could motivate experiments.  
516 In the wild oat example, a saturated path model (detailed in the supplemental material),  
517 ordered chronologically, yields the extended selection gradients (SD of posterior distribu-  
518 tion): dgerm, 0.021 (0.026); m60, 0.351 (0.076); dtf, -0.0232 (0.001); mass, 0.033 (0.005);  
519 rpt, 0.033 (0.002); and mrt, 0.086 (0.006). These inferences of selection based on a satu-  
520 rated model, which by definition fits the data perfectly, represent only small quantitative  
521 differences from those based on the original model (table 3b and figure 3). The main differ-  
522 ence is the slightly more negative selection of days to first flower (see also table 3c, which  
523 shows a potential direct component of the effect of dtf on fitness, over and above the effects  
524 included in the path model). The addition of such a direct effect to the path model may  
525 be justified on the (data-driven, post-hoc) argument that advanced phenology gives more  
526 time for optimal allocation of resources to different aspects of reproduction. However, for  
527 the present illustrative purposes, I have deferred to the expert opinion that contributed to  
528 the original publication of the *Avena* path model (Latta and McCain, 2009). Inferences  
529 of extended selection gradients and associated evolutionary predictions based on another  
530 alternative (highly *post hoc*) path model are presented in the supplemental material, and  
531 generate very similar results.

532 Experimental data may in principle be more powerful for testing causal hypotheses  
533 (Fisher, 1935), though experimentalists know that specific causal inference from any kind  
534 of data can be difficult! Manipulations of traits, or of the selective context in which traits

535 are expressed, are under-used approaches to characterizing mechanistic basis of natural se-  
536 lection. The concept of extended selection gradients may greatly facilitate the experimental  
537 verification of observational inferences about natural selection, especially for approaches  
538 based on trait manipulation. Developmental associations among traits make experimental  
539 verification of  $\beta$  notoriously difficult. The basic experiment to verify or quantify a direct  
540 selection gradient requires that a trait be manipulated independently of other traits, to  
541 test whether relative fitness changes by  $\beta \cdot \Delta z_{experimental}$ . However, developmental asso-  
542 ciations of traits - which may themselves be part of the causal structure of selection -  
543 generally make independent manipulation of traits difficult if not impossible or irrelevant.  
544 In contrast, experimental verification of extended selection gradients is not in principle  
545 opposed to the existence of developmental relationships among traits. Importantly, exper-  
546 imentation should be seen not only as a means of qualitatively verifying causal hypotheses,  
547 but also of quantitatively parameterizing mechanistic models. The statistical mechanics  
548 presented here for relating extended selection gradients to genetic variation and evolution  
549 are equally applicable using inferences from observational or experimental data, separately  
550 or in combination.

551 Perhaps the most important conceptual contribution in Arnold's (1983) paper is the  
552 demonstration of how to link theoretical and empirical perspectives on relationships among  
553 traits and relationships among traits and fitness in a quantitative framework. To date,  
554 applications of path analysis in studies of natural selection have relied almost entirely  
555 on observational data. In some cases, this includes complete life history data, which  
556 entirely determines fitness. Analyses are then conducted treating fitness as a (statistically)  
557 independently observed variable, when in fact it is derived entirely from other observed life  
558 history variables; van Tienderen (2000)'s methods provide the mathematical machinery to  
559 combine evolutionary demographic theory with path analytic approaches, but the method  
560 has been surprisingly little used (but see Coulson et al. 2003). The generalization of  
561 evolutionary demographic theory of quantitative traits provided by integral projection  
562 models (Coulson et al., 2010; Ellner and Rees, 2006) should provide a general means of

563 integrating demographic perspectives on fitness variation into path analysis and empirical  
564 studies of selection. Integration of path analytic approaches to characterizing natural  
565 selection into integral projection models will provide the analytical tools to model the  
566 consequences of non-linear causal effects of traits on one another and on fitness<sup>3</sup>, and to  
567 rigorously model non-normal distributions of traits. Also, Rice (2002, 2004) provides a  
568 complimentary set of theoretical principles by which a more comprehensive quantitative  
569 genetic theory of the selection, genetics, and evolution of non-normal and non-linearly  
570 causally-covarying traits could be developed. By these approaches, more theoretically and  
571 statistically-sound inference of causal relationships, and corresponding path coefficients  
572 and extended and direct selection gradients, among traits and fitness could be obtained  
573 directly from life history theory. In this context, life history and demographic theory  
574 can also be exploited to provide robust inference of path coefficients when traits interact  
575 multiplicatively.

## 576 Summary

577 Given *a priori* assumptions about causal relationships among traits and between traits and  
578 fitness, path analysis can provide inference of the total causal effects of traits on fitness.  
579 Formalization of such characterizations of selection as extended selection gradients, and  
580 consideration of how these coefficients relate to quantitative genetic variation and evolu-  
581 tionary change, provides the basis for incorporation of path analysis into the theoretical  
582 and empirical evolutionary quantitative genetics tool box. In particular, extended selection  
583 gradients may prove to be particularly useful for comparisons of selection across studies.  
584 While traditional, direct selection gradients provide entirely valid evolutionary predictions  
585 when used with their associated statistical quantitative genetic machinery (Lande, 1979),  
586 their biological interpretation is hindered by the fact that they do not describe the total

---

<sup>3</sup>The direct application of path rules to squared deviations of trait values from population means (e.g., as advocated by Scheiner et al. 2000) does not generally yield quantitatively or qualitatively correct inference of non-linear selection. It is not clear whether or not general analytical expressions for path-based inference of non-linear selection will be tractable, except in very simple restrictive cases. Outside of an integral project model framework, path-based inference of compound non-linear selection gradients could be obtained by numerical techniques.

587 causal effects of traits on fitness, and that their (correct) values vary arbitrarily as func-  
588 tions of what traits are studied. This statistical, rather than biological, definition can lead  
589 to trivialization of the mechanism of selection. In particular, evolution of traits that *cause*  
590 fitness variation indirectly, and traits that are incidentally correlated with selected traits,  
591 are both treated as cases of evolution due to genetic correlations in microevolutionary stud-  
592 ies based only on direct selection gradients. Empirical extended selection gradient-based  
593 inferences of microevolutionary processes rely heavily on *a priori* assumptions about cau-  
594 sation, or in other words, on additional information about the mechanism of selection, but  
595 perhaps only slightly more so than the use of direct selection gradients (Morrissey et al.,  
596 2010). The validity of such assumptions cannot be comprehensively assessed with obser-  
597 vational data (Wright, 1934) alone, such as that with which path-based studies of natural  
598 selection are typically parameterized. However, *a priori* biological knowledge can be used  
599 to construct plausible causal schemes. Furthermore, the clarification provided herein of  
600 how hypotheses about the organismal biology underlying trait-fitness relationships relate  
601 to selection gradients in a formal quantitative genetic sense should motivate and facili-  
602 tate further use of experimental approaches to understanding selective mechanisms. Path  
603 model-based thinking about natural selection should provide the means for formally link-  
604 ing observational, theoretical, and experimental inferences (Arnold, 1983), and this will  
605 greatly complement application of the statistical quantitative genetic principles pertaining  
606 to extended selection gradients.

## 607 Acknowledgements

608 The Soay sheep data were provided by Josephine Pemberton and Loeske Kruuk, and  
609 were collected primarily by Jill Pilkington and Andrew MacColl with the help of many  
610 volunteers. The collection of the Soay sheep data is supported by the National Trust for  
611 Scotland and QinetQ, with funding from NERC, the Royal Society, and the Leverhulme  
612 Trust. The phenotypic data on the wild oat lines were collected primarily by Kyle Gardner.

613 The data were provided by Bob Latta, who also developed the recombinant inbred lines  
614 with funding from NSERC, and contributed invaluable comments on an earlier draft of this  
615 manuscript. Bill Hill, Bruce Walsh, Jarrod Hadfield, Graeme Ruxton, Charles Goodnight,  
616 Kerry Johnson, Tom Meagher, Steve Frank, Sam Scheiner, Tim Coulson, Mike Ritchie,  
617 and Diane Byers provided valuable discussions and comments.

## 618 References

- 619 Arnold, S. J. 1983. Morphology, performance, and fitness. *American Zoologist* 23:347–361.
- 620 Bentler, P. M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin*  
621 107:238–46.
- 622 Catchpole, E. A., J. T. Morgan, T. N. Coulson, S. Freemand, and S. D. Albon. 2000.  
623 Factors influencing soay sheep survival. *Applied Statistics* 49:453–472.
- 624 Clutton-Brock, T. H. and J. M. Pemberton, 2004. Soay sheep dynamics and selection in  
625 an island population. Cambridge University Press, Cambridge.
- 626 Conner, J. K. 1996. Understanding natural selection: an approach integrating selection  
627 gradients, multiplicative fitness components, and path analysis. *Ethology Ecology Evo-*  
628 *lution* 8:387–397.
- 629 Coulson, T., E. A. Catchpole, S. D. Albon, B. J. T. Morgan, J. M. Pemberton, T. H.  
630 Clutton-Brock, M. J. Crawley, and B. T. Grenfell. 2001. Age, sex, density, winter  
631 weather, and population crashes in Soay sheep. *Science* 292:1528–1531.
- 632 Coulson, T., L. E. B. Kruuk, G. Tavecchia, J. Pemberton, and T. Clutton-Brock. 2003.  
633 Estimating the form of selection on neonatal traits in red deer using elasticity path  
634 analysis. *Evolution* 57:2879–2892.
- 635 Coulson, T., S. Tuljiapurkar, and D. Z. Childs. 2010. Using evolutionary demography to

- 636 link life history theory, quantitative genetics, and population ecology. *Journal of Animal*  
637 *Ecology* 76:1226–1240.
- 638 Crespi, B. J. and F. L. Brookstein. 1989. A path-analytic model for the measurement of  
639 selection on morphology. *Evolution* 43:18–28.
- 640 Ellner, S. and M. Rees. 2006. Integral projection models for species with complex life  
641 cycles. *The American Naturalist* 167:410–428.
- 642 Endler, J. A., 1986. *Natural selection in the wild*. Princeton University Press.
- 643 Etterson, J. R. and R. G. Shaw. 2001. Constraint to adaptive evolution in response to  
644 global warming. *Science* 294:151–154.
- 645 Fisher, R. A., 1935. *The Design of Experiments* (9th ed. 1971). Macmillan, New York.
- 646 Fox, J. 2006. Structural equation modeling with the sem package in R. *Structural Equation*  
647 *Modeling* 13:465–486.
- 648 Gardner, K. M. and R. G. Latta. 2008. Heritable variation and genetic correlation of quan-  
649 titative traits within and between ecotypes of *Avena barbata*. *Journal of Evolutionary*  
650 *Biology* 21:737–748.
- 651 Gianola, D. and D. Sorensen. 2004. Quantitative genetic models for describing simultaneous  
652 and recursive relationships between phenotypes. *Genetics* 168:1407–1424.
- 653 Godfrey-Smith, P. 2007. Conditions for evolution by natural selection. *The Journal of*  
654 *Philosophy* 104:489–516.
- 655 Hadfield, J. 2008. Estimating evolutionary parameters when viability selection is operating.  
656 *Proceedings of the Royal Society of London* 275:723–734.
- 657 ———. 2010. MCMC methods for multi-response generalized linear mixed models: The  
658 *MCMCglmm* R package. *Journal of Statistical Software* 33:1–22.

- 659 Henderson, C. R., 1973. Proceedings of the Animal Breeding and Genetics Symposium in  
660 Honour of Dr. Jay L. Lush, chap. Sire evaluation and genetic trends. Published by the  
661 American Society of Animal Science and the American Dairy Science Association.
- 662 de Jong, G. and A. J. van Noordwijk. 1992. Acquisition and allocation of resources: Genetic  
663 (co) variances, selection, and life histories. *The American Naturalist* 139:749–770.
- 664 Joreskog, K. G. and M. Van Thillo, 1972. LISTREL: a general computer program for  
665 estimating linear structural equation systems involving multiple indicators of unmeasured  
666 variables.
- 667 Kingsolver, J. G. and D. W. Schemske. 1991. Analyzing selection with path analysis.  
668 *Trends in Ecology and Evolution* 6:276–280.
- 669 Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the  
670 ‘animal model’. *Philosophical Transactions of the Royal Society of London B* 359:873–  
671 890.
- 672 Kruuk, L. E. B., J. Merilä, and B. C. Sheldon. 2003. When environmental variation  
673 short-circuits natural selection. *Trends in Ecology and Evolution* 18:207–209.
- 674 Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to  
675 brain:body size allometry. *Evolution* 33:402–416.
- 676 Lande, R. and S. J. Arnold. 1983. The measurement of selection on correlated characters.  
677 *Evolution* 37:1210–1226.
- 678 Latta, R. G. and C. McCain. 2009. Path analysis of natural selection via survival and  
679 fecundity across contrasting environments in *Avena barbata*. *Journal of Evolutionary  
680 Biology* 22:2458–2469.
- 681 Lush, J. L., 1937. Animal breeding plans. Iowa State College Press, Ames, Iowa.
- 682 Lynch, M. and B. Walsh, 1998. Genetics and analysis of quantitative traits. Sinauer,  
683 Sunderland, MA.

- 684 McArdle, J. J. and R. P. McDonald. 1984. Some algebraic properties of the reticular action  
685 model. *British Journal of Mathematical and Statistical Psychology* 37:234–251.
- 686 Mitchell-Olds, T. and R. G. Shaw. 1987. Regression analysis of natural selection: statistical  
687 inference and biological interpretation. *Evolution* 41:1149–1161.
- 688 Morrissey, M. B., L. E. B. Kruuk, and A. J. Wilson. 2010. The danger of applying the  
689 breeder's equation in observational studies of natural populations. *Journal of Evolutionary  
690 Biology* 23:2277–2288.
- 691 Morrissey, M. B., D. J. Parker, P. Korsten, J. M. Pemberton, L. E. B. Kruuk, and A. J. Wil-  
692 son. 2012. The prediction of adaptive evolution: empirical application of the secondary  
693 theorem of selection and comparison to the breeder's equation. *Evolution* 66:2399–2410.
- 694 Price, G. R. 1970. Selection and covariance. *Nature* 227:520–521.
- 695 Queller, D. C. 1992. Quantitative genetics, inclusive fitness, and group selection. *The  
696 American Naturalist* 139:540–558.
- 697 Rausher, M. D. 1992. The measurement of selection on quantitative traits: biases due to  
698 environmental covariances between traits and fitness. *Evolution* 46:616–626.
- 699 Rice, S. H. 2002. A general population genetic theory for the evolution of developmental  
700 interactions. *Proceedings of the National Academy of Sciences* 99:15518–15523.
- 701 ———. 2004. Developmental associations between traits: covariance and beyond. *Genetics*  
702 166:513–526.
- 703 Robertson, A. 1966. A mathematical model of the culling process in dairy cattle. *Animal  
704 Production* 8:95–108.
- 705 ———, 1968. Population Biology and Evolution, chap. The spectrum of genetic variation,  
706 Pp. 5–16. Syracuse University Press, New York.
- 707 Scheiner, S. M., K. Donohue, L. A. Dorn, S. J. Mazer, and L. M. Wolfe. 2002. Reducing  
708 environmental bias when measuring natural selection. *Evolution* 56:2156–2167.

- 709 Scheiner, S. M., R. J. Mitchell, and H. S. Callahan. 2000. Using path analysis to measure  
710 natural selection. *Journal of Evolutionary Biology* 13:423–433.
- 711 Schlüter, D. 1988. Estimating the form of natural selection on a quantitative trait. *Evo-  
712 lution* 42:849–861.
- 713 Shaw, R. G. and C. J. Geyer. 2010. Inferring fitness landscapes. *Evolution* 64:2510–2520.
- 714 Shipley, B. 1997. Exploratory path analysis with applications in ecology and evolution.  
715 *The American Naturalist* 146:1113–1138.
- 716 Sober, E., 1984. *The Nature of Selection*. University of Chicago Press, Chicago.
- 717 Steiger, J. H. 1990. Structural model evaluation and modification. *Multivariate Behavioral  
718 Research* 25:935–943.
- 719 van Tienderen, P. H. 2000. Elasticities and the link between demographic and evolutionary  
720 dynamics. *Ecology* 81:666–679.
- 721 Walsh, B. and M. Lynch, 2012. Evolution and selection of quantitative traits, chap.  
722 Short-term changes in the mean: 1. The breeders' equation. Advance copy: ni-  
723 tro.biosci.arizona.edu/zbook/NewVolume\_2/newvol2.html.
- 724 ———, 2013. Evolution and selection of quantitative traits,  
725 chap. 27. Measuring multivariate selection. Advance copy: ni-  
726 tro.biosci.arizona.edu/zbook/NewVolume\_2/newvol2.html.
- 727 Wilson, A. J., D. Reale, M. N. Clements, M. B. Morrissey, C. A. W. E. Postma, L. E. B.  
728 Kruuk, and D. H. Nussey. 2010. An ecologist's guide to the animal model. *Journal of  
729 Animal Ecology* 79:13–26.
- 730 Wright, S. 1921. Correlation and causation. *Journal of Agricultural Research* 20:557–585.
- 731 ———. 1934. The method of path coefficients. *Annals of Mathematical Statistics* 5:161–  
732 215.

Table 1: Variances (diagonal), covariances (below diagonal) and correlations (above diagonal) of lamb traits and relative fitness (based on first year over-winter survival) in female Soay sheep. Traits are Julian birth date  $b\text{dy}$ , twin status  $t\text{wn}$ , birth weight  $b\text{wt}$  (kg), weight in August  $a\text{wt}$  (kg), and relative fitness is denoted  $w$  (first year survival scored as 0 and 1, divided by annual mean survival). Values are the modes of the posterior distribution of the (co)variances or correlations and bracketed values are standard deviations of the posterior distribution, interpretable similarly to standard errors.

	$b\text{dy}$	$t\text{wn}$	$b\text{wt}$	$a\text{wt}$	$w$
$b\text{dy}$	61.19 (2.41)	-0.040 (0.028)	0.143 (0.028)	-0.084 (0.028)	0.040 (0.028)
$t\text{wn}$	-0.134 (0.092)	0.176 (0.007)	-0.398 (0.023)	-0.243 (0.027)	-0.136 (0.026)
$b\text{wt}$	0.629 (0.127)	-0.094 (0.007)	0.3191 (0.013)	0.390 (0.023)	0.169 (0.027)
$a\text{wt}$	-2.54 (0.871)	-0.399 (0.049)	0.877 (0.067)	15.561 (0.608)	0.282 (0.026)
$w$	0.208 (0.154)	-0.038 (0.008)	0.067 (0.011)	0.779 (0.081)	0.492 (0.020)

Table 2: Standardized (a) path coefficients, and (b) compound path coefficients, i.e.,  $\eta_{ij}$  based on the fitted path model relating sheep neonatal and lamb traits to relative fitness during the first year of life. The bottom row of (a) are equivalent to direct selection gradients, and the bottom row of (b) are path model-based extended selection gradients. Units are: birth day, days; twin status  $\in [0, 1]$ , birth and August weights, kg. Values are the modes of the posterior distribution estimates and bracketed values are standard deviations of the posterior distribution, interpretable similarly to standard errors.

(a) path coefficients ( $b$ , bottom row are path-based $\beta$ )				
	<i>birth day</i>	<i>twin status</i>	<i>birth weight</i>	<i>August weight</i>
<i>birth weight</i>	0.126 (0.026)	-0.396 (0.026)		
<i>August weight</i>	-0.139 (0.025)	-0.103 (0.028)	0.363 (0.027)	
$w$	0.036 (0.019)	-0.028 (0.021)	0.033 (0.021)	0.174 (0.021)

(b) compound path coefficients ( $\Phi$ , bottom row are $\eta$ )				
	<i>birth day</i>	<i>twin status</i>	<i>birth weight</i>	<i>August weight</i>
<i>birth weight</i>	0.126 (0.026)	-0.396 (0.026)		
<i>August weight</i>	-0.100 (0.027)	-0.254 (0.028)	0.363 (0.027)	
$w$	0.022 (0.019)	-0.092 (0.020)	0.096 (0.021)	0.174 (0.021)

Table 3: Unstandardized (a) path coefficients, (b) compound path coefficients, i.e.,  $\eta_{ij}$  based on the fitted path model, and (c) unstandardized, unconstrained direct selection differentials, of phenological, vegetative, and reproductive traits in a greenhouse experiment with a population of recombinant inbred lines of wild oat *Avena barbata* derived from contrasting ecotypes. Traits are days to germination  $dgerm$ , mass at day 60  $m60$  (g), days to first flower  $dtf$ , total final mass  $mass$  (g), number of reproductive tillers  $rpt$ , mass of reproductive tillers  $mrt$  (g), and relative fitness  $w$ . The bottom row of (a) are path model-based direct selection gradients, and the bottom row of (b) are path model-based extended selection gradients. The unconstrained direct selection gradients in (c) are obtained by the multiple regression of relative fitness on all six traits. Values are the modes of the posterior distribution estimates and bracketed values are standard deviations of the posterior distribution, interpretable similarly to standard errors.

(a) path coefficients ( $b$ , bottom row are path-based $\beta$ )						
	$dgerm$	$m60$	$dtf$	$mass$	$rpt$	$mrt$
$m60$	-0.004 (0.014)					
$dtf$		-12.707 (2.174)				
$mass$		3.966 (0.517)				
$rpt$			-0.265 (0.018)	0.303 (0.075)		
$mrt$			-0.125 (0.006)	0.401 (0.023)		
$w$				-0.014 (0.004)	0.029 (0.002)	0.098 (0.005)
(b) compound path coefficients ( $\Phi$ , bottom row are $\eta$ )						
	$dgerm$	$m60$	$dtf$	$mass$	$rpt$	$mrt$
$m60$	-0.004 (0.014)					
$dtf$	0.058 (0.188)	-12.707 (2.174)				
$mass$	-0.019 (0.056)	3.966 (0.517)				
$rpt$	-0.022 (0.067)	4.791 (0.683)	-0.265 (0.018)	0.303 (0.075)		
$mrt$	-0.014 (0.045)	3.239 (0.353)	-0.125 (0.006)	0.401 (0.023)		
$w$	-0.002 (0.006)	0.412 (0.049)	-0.020 (0.001)	0.034 (0.005)	0.029 (0.002)	0.098 (0.005)
(c) unconstrained direct selection gradients ( $\beta$ )						
	$dgerm$	$m60$	$dtf$	$mass$	$rpt$	$mrt$
$w$	0.014 (0.013)	0.032 (0.043)	-0.0033 (0.0011)	-0.012 (0.0045)	0.027 (0.002)	0.086 (0.0064)

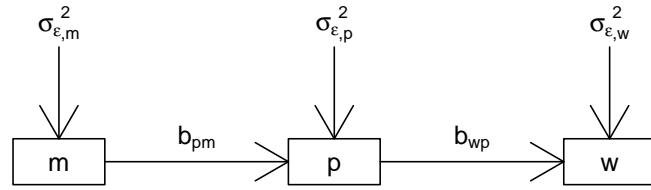


Figure 1: A hypothetical relationship between a morphological trait ( $m$ ), and aspect of organismal performance ( $p$ ), and fitness ( $w$ ). Arrows indicate hypothesized causal relationships. Path coefficients,  $b_{pm}$  and  $b_{wp}$  are the regression coefficients of performance on morphology, and relative fitness on performance, respectively.  $\sigma_\epsilon^2$  are exogenous variances, i.e., variation in endogenous variables beyond that attributable to causal effects in the path model.

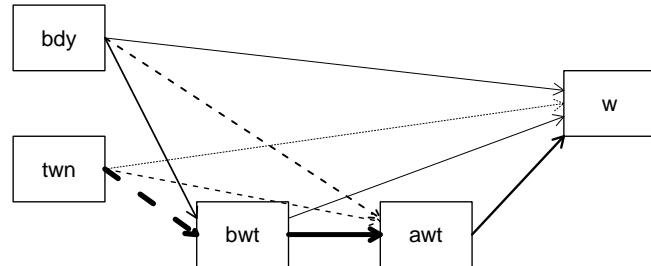


Figure 2: Parameterized path diagram representing relationships among lamb traits, and among lamb traits and relative fitness ( $w$ , based on first year over-winter survival), in female Soay sheep. Traits are Julian birth date  $b_{dy}$ , twin status  $t_{wn}$ , birth weight  $b_{wt}$ , weight in August  $a_{wt}$ . Path coefficients among traits are standardized, i.e., they represent partial correlations, and path coefficients between traits and relative fitness represent unit variance-standardized partial regression coefficients. The thickness of arrows represents the strength of the corresponding path coefficients. Solid arrows represent positive relationships and dashed arrows represent negative relationships. Exogenous inputs of variance are omitted for clarity.

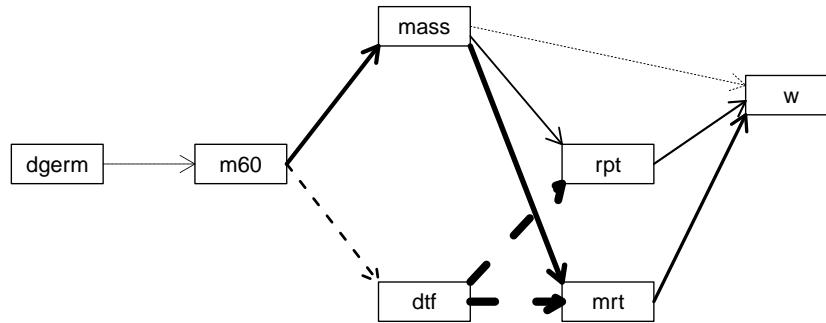


Figure 3: Parameterized path diagram representing relationships among phenological, vegetative, and reproductive traits in a population of recombinant inbred lines derived from contrasting ecotypes and raised under greenhouse conditions. The traits are days to germination  $dgerm$ , mass at day 60  $m60$ , days to first flower  $dtf$ , final total mass  $mass$ , mass of reproductive tillers  $mrt$ , number of reproductive tillers  $rpt$ , and relative fitness  $w$ , as assessed via variation in fecundity. Path coefficients among traits are standardized, i.e., they represent partial correlations, and path coefficients between traits and relative fitness represent unit variance-standardized partial regression coefficients. The thickness of arrows represents the strength of the corresponding path coefficients. Solid arrows represent positive relationships and dashed arrows represent negative relationships. Exogenous inputs of variance are omitted for clarity.

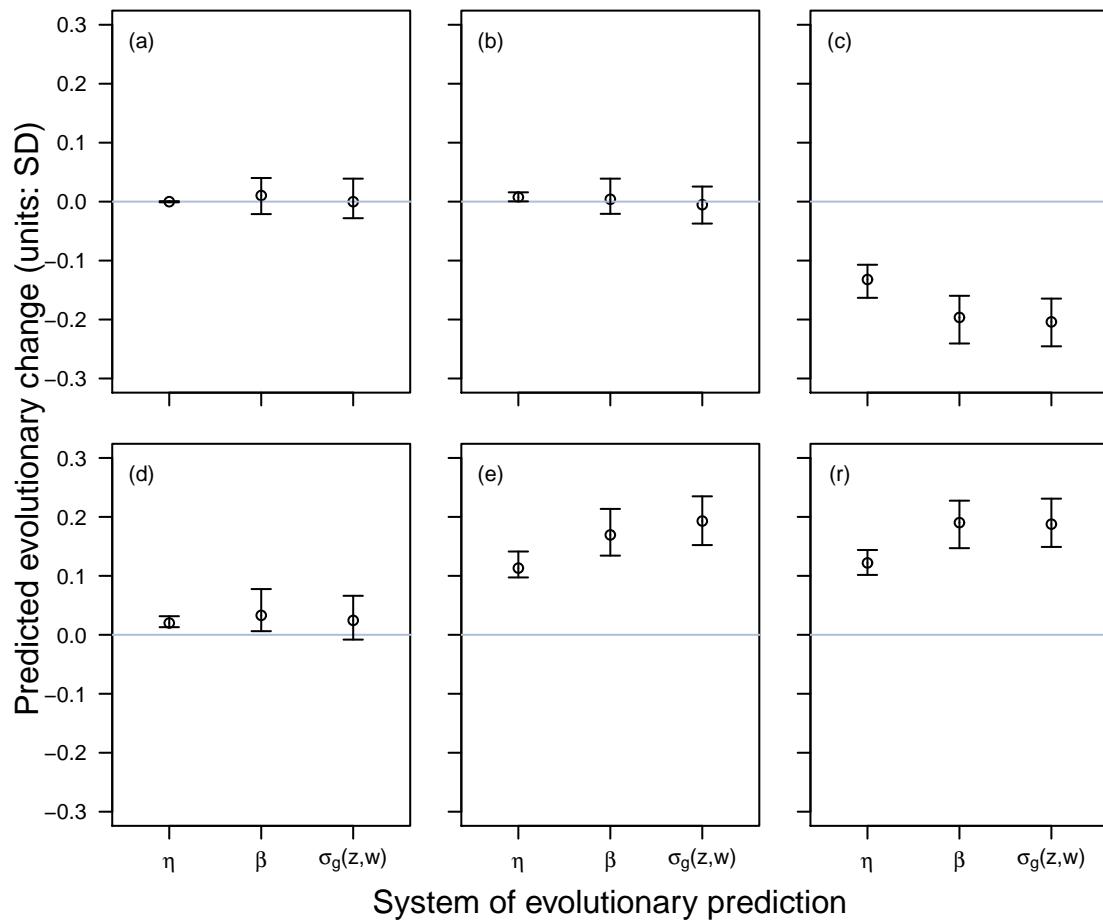


Figure 4: Evolutionary prediction for vegetative and reproductive traits from a laboratory experiment on a population of recombinant inbred lines of wild oat *Avena barbata* using extended selection gradient-based evolutionary prediction, the breeder's equation (specifically, Lande's formulation based on direct selection gradients,  $\beta$ ), and the secondary theorem of selection, i.e., the genetic covariance of each trait with relative fitness,  $\sigma_g(z, w)$ . Traits are (a) days to germination, (b) mass at day 60, (c) days to first flower, (d) final total mass, (e) number of reproductive tillers, and (f) total mass of reproductive tillers. Points are mean values of the posterior distribution of the evolutionary prediction based on each predictive framework (path-based extended selection gradients:  $\eta$ , multiple regression-based application of the Lande equation:  $\beta$ , and application of the secondary theorem of selection:  $\sigma_g(z, w)$ ), and the error bars denote 95% credible intervals.