Behavioral selection in structured populations

Matthias Borgstede[1]

Abstract

The *multilevel model of behavioral selection* (MLBS) by Borgstede and Eggert (Behav Process 186:104370. 10.1016/j.beproc.2021.104370, 2021) provides a formal framework that integrates reinforcement learning with natural selection using an extended Price equation. However, the MLBS is so far only formulated for homogeneous populations, thereby excluding all sources of variation between individuals. This limitation is of primary theoretical concern because any application of the MLBS to real data requires to account for variation between individuals. In this paper, I extend the MLBS to account for inter-individual variation by dividing the population into homogeneous sub-populations and including class-specific reproductive values as weighting factors for an individual's evolutionary fitness. The resulting formalism closes the gap between the theoretical underpinnings of behavioral selection and the application of the theory to empirical data, which naturally includes inter-individual variation. Furthermore, the extended MLBS is used to establish an explicit connection between the dynamics of learning and the maximization of individual fitness. These results expand the scope of the MLBS as a general theoretical framework for the quantitative analysis of learning and evolution.

Keywords Selection by consequences · Behavioral selection · Natural selection · Price equation · Covariance based law of effect · Multilevel model of behavioral selection

Introduction

Darwinian thinking has a long tradition in behavior analysis (e.g., Broadbent 1961; Campbell 1956; Gilbert 1970; Herrnstein 1964; Hull et al. 2001; Pringle 1951; Skinner 1966; Thorndike 1900). Although there seems to be a wide consensus that reinforcement learning can be described by analogy with natural selection, opinions about how exactly behavioral selection[1] is connected to natural selection diverge. For example, Skinner (1981) claims that natural selection and reinforcement learning are two instances of the same underlying causal principle: selection by consequences. A different view is articulated by McDowell (2013), who proposes that learning and evolution constitute a single, self-similar process. Others have taken the position that the analogy between natural selection and behavioral selection is

misleading because mechanisms of learning are themselves subject to natural selection (Burgos 2019).

The intricate relation between learning and evolution has also been acknowledged in evolutionary biology (e.g., Dunlap and Stephens 2016; McNamara and Houston 2009; Stephens 1991). Biological accounts of learning often focus on specific learning mechanisms as the target of natural selection (Aoki and Feldman 2014; Dridi and Lehmann 2014; Fawcett et al. 2013; Moore 2004). However, none of these approaches provides a general theoretical integration of learning and evolution by means of a formal model of selection. Such an integrated perspective would not only help to understand how evolution and learning interact, but also clarify under which conditions it is justified to assume that individual learning may lead to increases in individual

---

[1] In this paper, the term "behavioral selection" exclusively refers to learning processes on the level of the individual organism.

✉ Matthias Borgstede
matthias.borgstede@uni-bamberg.de

[1] University of Bamberg, Markusplatz 3, 96047 Bamberg, Germany

fitness (cf. Frankenhuis et al. 2019; Lewis et al. 2010; Singh et al. 2010).

The multilevel model of behavioral selection (MLBS) aims at clarifying the relation between learning and evolution from the perspective of selection theory (Borgstede and Eggert 2021). The MLBS is a direct extension of the quantitative theory of natural selection as described by the Price equation (Price 1970, 1972). Therefore, all conclusions drawn from the model can be interpreted as mathematical theorems of natural selection. Following the rationale of multilevel selection theory (cf. Gardner 2015), selection is described both, between and within individuals.

The MLBS describes learning as an interaction between an individual and its environment on a molar level (i.e., it treats behavior as extended over time), thereby shifting the focus from molecular mechanisms like associative learning to statistical measures of behavioral allocation like mean, variance and covariance (cf. Baum 1973; Rachlin 1978). One of the core features of the MLBS is the use of statistical fitness predictors ("fitness proxies") to define reinforcers, thereby establishing a functional relation between reinforcement learning and natural selection. In this view, a reinforcer is not a thing that is received by the individual, it is an event that co-varies with evolutionary fitness, given the conditions under which the organism has been shaped by natural selection. Such fitness predictors are phylogenetically important events (PIEs), as introduced by Baum (2012). A PIE that positively co-varies with evolutionary fitness will foster positive selection (i.e., reinforcement), whereas a PIE that negatively co-varies with fitness will result in negative selection (i.e., punishment). By replacing the concept of a reinforcer by the concept of PIEs, it is possible to analyze individual behavioral adaptations with regard to their evolutionary function, rather than the specific (molecular) mechanisms that realize this function.

The MLBS has inspired several theoretical papers that link the principle of behavioral selection to other analytical frameworks for behavior analysis, such as information seeking (Borgstede 2021) or matching theory (Borgstede and Luque 2021). Furthermore, first empirical applications of the framework have been proposed in Strand et al. (2022) and Borgstede and Anselme (2024).

However, the formalism introduced by Borgstede and Eggert (2021) builds on the simplifying assumption of a homogeneous population. From a strictly formal position, such a restriction would rule out all inter-individual variation apart from the selected behavior. Borgstede and Eggert (2021) argue that the assumption of a homogeneous population is justified because the theory deals with laws of behavior on the most general level. Moreover, it is possible to justify individual selection estimates as proposed in Borgstede and Anselme (2024) by stating the MLBS for an abstract population of individuals that are conceived to be identical to the individual under study. Although formally consistent, it is difficult to give a theoretically sound interpretation to such fictitious populations. A more elegant way to connect the empirical methodology by Borgstede and Anselme (2024) to the theoretical underpinnings presented in Borgstede and Eggert (2021), would be to include inter-individual variation in the formal presentation of the MLBS. Dropping the assumption of a homogeneous population would further widen the scope of the theory considerably, especially with regard to its evolutionary foundations. Therefore, in this paper, I extend the MLBS to account for inter-individual variation. The aim is to close the theoretical gap between the formalization of behavioral selection in the MLBS and possible empirical applications of the theory that naturally require to account for individual differences. The results will then be used to investigate the relation between individual learning dynamics and the idea that individual learning maximizes evolutionary fitness. While the latter is routinely assumed in many models of behavioral ecology (cf. Davies et al. 2012), it has never been shown on a theoretical level that fitness maximization models may in fact be applied to individual learning.

In the remainder of the paper, I will first review the core concepts and formalism of the MLBS as presented in Borgstede and Eggert (2021). I will then proceed to incorporate variation between individuals by dividing the population into homogeneous sub-populations, so-called "classes." These classes formally account for individually different environments, constant individual traits, as well as regionally or culturally separated groups of individuals (Caswell 2001). Incorporating class-structure into the MLBS thus extends the scope of the model to arbitrary sources of inter-individual variation. The resulting generalized MLBS introduces the concept of *reproductive value* in the theory of behavioral selection. Reproductive value captures the relative contribution of a class of individuals to the future population and thus constitutes a weighting factor for the calculation of evolutionary fitness (Grafen 2006). The main result is that the principle of behavioral selection remains valid in inhomogeneous populations if individual fitness is evaluated in terms of expected gain in reproductive value. This result is further exploited to provide a formal justification for the assumption that individual learning tends to result in fitness maximization.

The multilevel model of behavioral selection

The multilevel model of behavioral selection describes behavioral selection as an integral component of natural selection. The general idea is that behavioral change within individuals can be described by means of natural selection, if (a) individuals are treated as their own offspring and (b)

evolutionary fitness is replaced by a statistical fitness predictor at the individual level. Evolutionary fitness is defined as the contribution of an individual to the future population (if individuals are treated as their own offspring, this corresponds to their survival probability).

Formally, the model relies on the most general description of natural selection as provided by the Price equation (Price 1970, 1972). The Price equation describes the average change of an evolving character from one generation (parents) to the next (offspring) by decomposing the change in mean character value $\Delta \bar{b}$ into one covariance term (selection) and one expectation term (non-selection)[2]:

$$\overline{w}\Delta\overline{b} = \text{Cov}(w_i, b_i) + \text{E}(w_i\Delta b_i) \qquad (1)$$

Here, b_i is the character value of parent i and w_i designates the contribution of parent i to the offspring generation (i.e., the relative frequency of individuals in the offspring generation that descend from parent i). It is important to note that the change in mean character value $\Delta \bar{b} = \text{E}(b_i\prime - b_i)$ is defined with respect to the parent generation, as well. This means that $b_i\prime$ refers to the mean character value of parent i's offspring (i.e., the index always refers to the parent). $\overline{w}$ is defined as the average contribution of parents to the offspring generation. With these definitions, the Price equation is a mathematical truth. In other words, for any two sets that can be related in the sense of "parents" and "offspring," the change in mean character value corresponds to the covariance between character value and individual fitness plus the expectation over the difference between a parents' and their offspring's character values. This mathematical insight describes the concept of change by a process of selection on the most general level. The Price equation is true for arbitrary characters and for arbitrary mechanisms of inheritance. Therefore, it seems plausible that it also provides a suitable mathematical background for the description of reinforcement learning (cf. Price 1995, written ca. 1971).

The potential of the Price equation as a formal account of reinforcement learning was further explored by Baum (2017). Baum's formal account of behavioral selection using the Price equation builds on the idea that *behaviors* are the elementary units of selection. Here, selection refers to the change in average behavioral allocation between one set of reinforcement trials (the "parent population") and a later set of reinforcement trials (the "offspring population"). While providing a mathematically sound description of behavioral

change for some special cases, Baum's account does not provide a complete account of how "parent behaviors" produce "offspring behaviors," hence it is difficult to make sense of the corresponding behavioral fitness equivalent (Borgstede and Eggert 2021).

The MLBS takes a slightly different route, treating individuals as the elementary units of selection. Individuals behave, individuals reproduce and individuals die, whereas behaviors do nothing of the above. Therefore, fitness is defined at the individual level as the contribution of an individual to the future population (i.e., using the standard definition of evolutionary fitness as described above). Second, behavioral selection occurs within individuals. This means that the effect of natural selection as captured by the covariance term in the Price equation is negligible for behavioral selection. Therefore, the MLBS focuses on the expectation term, restricting the analysis to the survival part of evolutionary fitness. Third, because fitness is an individual characteristic, there is no variation in fitness within individuals (and hence no covariance between fitness and any target of behavioral selection). In order to enable selection, the individual is therefore assumed to adapt its behavior to the environment using statistical fitness proxies.

Adopting a molar view on behavior, the MLBS describes the change in mean behavioral allocation, where behavioral allocation is itself taken to be extended over time. For example, one may express the behavioral allocation of a foraging animal as the relative time spent at each food patch within a certain interval. These intervals constitute *behavioral episodes* and are defined by a uniform class of recurring contextual factors (i.e., by a certain structure of reinforcement contingencies). Behavioral change by means of selection is thus specified on two different levels, the level of mean population behavior (averaged over individuals) and the level of mean individual behavior (averaged over behavioral episodes) with behavioral episodes being nested within individuals.

Building on these concepts, it is now possible to describe the within individual change $w_i\Delta b_i$ within the expectation term of the Price equation by recursive expansion (i.e., by inserting the Price equation into itself). To indicate the different levels of selection, I introduce the index i for expected values and covariances at the population level and the index j for expected values and covariances at the within-individuals level:

$$\overline{w}\Delta\overline{b} = \text{Cov}_i(w_i, b_i) + \text{E}_i\big(\text{Cov}_j(w_{ij}, b_{ij}) + \text{E}_j(w_{ij}\Delta b_{ij})\big) \quad (2)$$

Consequently, the fitness-weighted change within individuals is given by:

$$w_i\Delta b_i = \text{Cov}_j(w_{ij}, b_{ij}) + \text{E}_j(w_{ij}\Delta b_{ij}) \qquad (3)$$

---

[2] Note that the notation used for the Price equation varies considerably between publications. If the evolving character is an allele frequency, it is often designated x, y or z. Sometimes authors differentiate between the phenotypic character and the part of it that is subject to evolutionary change, which is then referred to as either g or p. Since in this paper I am concerned with behavior (as opposed to genetic values), I prefer the letter b.

Equation 3 describes behavioral selection within individual i as the within individual covariance between behavioral allocations over episodes j (b_{ij}) and the corresponding evolutionary fitness w_{ij}. The second term refers to any changes in behavioral allocation within individual i apart from the selection component. In the MLBS, the varying fitness values w_{ij} are further replaced by the predicted values from a linear regression of the form $w_{ij} = \beta_0 + \beta_{wp}p_{ij} + \varepsilon$. Note that the regression coefficients refer to the population level here. Inserting this in the above equation (and defining $E_j(w_{ij}\Delta b_{ij}) = \delta$) yields:

$$w_i\Delta b_i = \text{Cov}_j\left(\beta_0 + \beta_{wp}p_{ij}, b_{ij}\right) + \delta \tag{4}$$
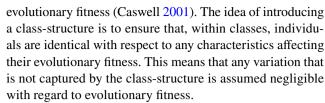
which can be rearranged to:

$$w_i\Delta b_i = \beta_{wp}\text{Cov}_j\left(p_{ij}, b_{ij}\right) + \delta \tag{5}$$

Equation 5 is called the *covariance based law of effect* (CLOE) since it describes the expected change in mean behavioral allocation due to behavioral selection in a quantitative way. The term β_{wp} refers to the regression slope of evolutionary fitness on p and is called the "reinforcing power" of a fitness proxy. Thus, mean change in behavioral allocation is proportional to the covariance between behavior and a statistical fitness predictor with the reinforcing power of the predictor as a scaling factor.

By linking the concept of behavioral selection to the theory of natural selection, the MLBS gives a quantitative account of reinforcement at the most general level. It gives a valid description of all processes of behavioral selection regardless of the specific mechanisms involved. However, in its current form, the model makes one limiting assumption: individuals are treated as if they were identical when conditioned on p. This means that strictly speaking, the CLOE is only valid if all fitness predictors apart from p are distributed randomly over the individuals. This is a considerable limitation of the scope of the theory and shall thus be addressed in the following section.

Introducing variation to the MLBS

In biological models of natural selection, inter-individual variation is usually captured by dividing the population into homogeneous sub-populations, so called *classes* (e.g., Batty et al. 2014; Grafen 2020; Taylor 1990). These classes may refer to different stages in the life cycle of the species (e.g., juvenile, adult, post-reproductive…), different sexes (male, female, hermaphrodite…), age (e.g., in seasonal species), spatial distribution (e.g., different food patches) or any other personal or environmental characteristic that is relevant for

evolutionary fitness (Caswell 2001). The idea of introducing a class-structure is to ensure that, within classes, individuals are identical with respect to any characteristics affecting their evolutionary fitness. This means that any variation that is not captured by the class-structure is assumed negligible with regard to evolutionary fitness.

There are different versions of the Price equation for class-structured populations with various notations, depending on the type of class-structure and the aim of the underlying model (e.g., Grafen 2015; Lion 2018; Taylor 1990). Nevertheless, these different formulations all follow the same general rationale. First, natural selection is formulated by means of the Price equation for each class separately. And second, the overall population change of the evolving character is calculated as a weighted mean of the contributions from each class. Formally, for the Price equation to make sense, the weighting factors are arbitrary. However, in the context of biological evolution, adequate weighting factors have been shown to be unique only up to a positive constant and correspond to the *reproductive values* of the corresponding classes (Batty et al. 2014; Grafen 2006, 2015; Taylor 1990). Reproductive values were introduced by Fisher (1930) and refer to the relative genetic contribution of an individual to the future population. As such, reproductive values depend on class-specific fertility rates and reproduction rates (i.e., they are demographic parameters). Given a demographic model of the population dynamics, reproductive values can be calculated using standard methods from matrix algebra (Caswell 2001, 2010). For reasons of mathematical convenience, in the following I assume reproductive values to be scaled such that the sum of all class reproductive values v_x is one. The weighted average thus reduces to a weighted sum over classes.

The Price equation for classes can now be stated as:

$$\overline{w}\sum_x v_x\Delta\overline{b}_x = \sum_x v_x\text{Cov}_x\left(w_i, b_i\right) + \sum_x v_x E_x\left(w_i\Delta b_i\right) \tag{6}$$

Note that covariances and expectation values are taken not over all individuals, but conditional on the respective classes x. Moreover, the "average change" in population value $\Delta\overline{b}$ in the class-structured case is also actually a weighted average of the conditional mean changes for each class ($\Delta\overline{b}_x$) with the class reproductive values (v_x) as weighting factors. Finally, individual fitness w_i is defined as a weighted sum over the number of offspring n_y in each class y multiplied by the corresponding offspring reproductive values v_y:

$$w_i = \sum_y v_y n_{iy} \tag{7}$$

Like in the simple case without classes, it is straightforward to expand the expectation term by applying the Price

equation recursively to describe mean character change from individual i to the descendants of individual i, i.e., $w_i \Delta b_i$. Since the MLBS focuses only on the survival part of natural selection, this corresponds to within individuals change in behavioral allocation b_i from one set of behavioral episodes to the next. Furthermore, as long as individuals only change classes between these time steps, all behavioral episodes j within individuals i correspond to the same class x. Therefore, the class subscript can be omitted and the equation for behavioral selection remains unchanged[3]:

$$w_i \Delta b_i = \mathrm{Cov}_j\left(w_{ij}, b_{ij}\right) + \mathrm{E}_j\left(w_{ij} \Delta b_{ij}\right) \quad (8)$$

Like in the simple MLBS without class-structure, behavioral selection is expressed with respect to a statistical fitness predictor p, rather than fitness itself. However, since individual fitness is now defined using offspring reproductive values v_y as weighting factors, it is necessary to specify the effect of p with respect to each possible subsequent class y. This means that instead of one simple linear regression of fitness on p, we now use a linear regression of each fitness component on p. These fitness components correspond to the offspring classes, i.e., the possible classes y to which the individual may transition from the current class x. Formally, this is accomplished by a class specific linear regression for each number n_y of descendants in offspring class y of the form: $n_y = \beta_{xy0} + \beta_{xyp}p + \varepsilon$. Like in the simple case, these regressions are defined on the level of the population. The evolutionary fitness of individual i can thus be written as:

$$w_i = \sum_y v_y\left(\beta_{xy0} + \beta_{xyp}p + \varepsilon\right) = \sum_y v_y \beta_{xy0} + v_y \beta_{xyp}p + v_y \varepsilon \quad (9)$$

Substituting this for the predicted fitness in the behavioral selection part of the Price equation yields:

$$w_i \Delta b_i = \mathrm{Cov}_j\left(\sum_y v_y \beta_{xy0} + v_y \beta_{xyp}p_{ij}, b_{ij}\right) + \\ \mathrm{E}_j\left(\sum_y v_y \beta_{xy0} + v_y \beta_{xyp}p_{ij} \Delta b_{ij}\right) \quad (10)$$

Since we are concerned with the selection part only, we simplify by treating the expectation term as a residual δ. Within-individuals change in behavioral allocation can now be rearranged to:

$$w_i \Delta b_i = \sum_y v_y \beta_{xyp} \mathrm{Cov}_j\left(p_{ij}, b_{ij}\right) + \delta \quad (11)$$

This is the covariance based law of effect with class-structure. Its general form resembles the simple version derived in Borgstede and Eggert (2021): behavioral selection is proportional to the covariance between behavioral allocation and an arbitrary fitness predictor. However, in the class-structured case, the reinforcing power of p is now defined as the reproductive value weighted sum over the effects of p on all fitness components y (i.e., $\sum_y v_y \beta_{xyp}$).

Learning and fitness maximization

Several authors have approached individual learning from a maximization perspective, where stable state behavior is analyzed in terms of maximal reinforcer value (cf. Rachlin et al. 1981, 1976; Rachlin and Burkhard 1978). Reinforcer value in this context is closely connected to the concept of subjective utility from behavioral economics (Herrnstein et al. 1993; Loewenstein et al. 2009). Similarly, behavioral ecology routinely applies fitness maximization as an explanatory mode for behavioral adaptations (cf. Caswell 1982; Davies et al. 2012). Although the latter is concerned with evolutionary adaptations (i.e., the possible outcomes of natural selection), many models in behavioral ecology actually refer to the outcome of individual level adaptations. For example, the ideal free distribution (IFD) describes how individuals in a population are distributed over two different food patches when the outcome is negatively related to the number of individuals that already are at the patch. Taking consumed food as a proxy of evolutionary fitness, individual fitness will be maximal for each individual if and only if the ratio of individuals at the two patches matches the ratio of food resources at the two sites (Fretwell 1972; Fretwell and Lucas 1969). Although the formal model is meant to explain evolutionary adaptations, the actual distribution of individuals over food patches is hardly ever the result of natural selection. The model correctly predicts that natural selection should produce an IDF for plants or other organisms whose spatial position is fixed over the lifetime of a single individual. However, many applications of the IFD involve moving animals, whose spatial position with regard to food patch distribution is often not the result of natural selection, but rather of individual learning (Houston 2008; Kraft et al. 2002). To apply an evolutionary model of fitness optimization to individual behavioral adaptations (i.e., learning), presupposes that learning can be explained by the same principles as evolution. Many results from behavioral ecology seem to imply that this might indeed be the case (cf. Davies et al. 2012; Stephens and Krebs 1986). However, unless evolutionary theory and learning theory are described within a unified model, the assumption that individual learning leads to the maximization of evolutionary fitness is unjustified.

---

[3] An analogous result is presented by Gardner (2015) who argues that social groups may be considered as units of selection if the groups are homogenous with respect to their respective classes.
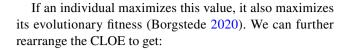
The MLBS may provide such a unified theoretical perspective and thus close the theoretical gap between individual learning and fitness maximization. In the following, I will establish an explicit link between the class-structured version of the CLOE developed above and the principle of fitness maximization by exploiting the formal equivalence of evolutionary fitness and reinforcer value established in Borgstede (2020). The paper argues that if there is a maximand of reinforcement learning, it must be defined such that maximizing reinforcement coincides with maximizing evolutionary fitness. The link between reinforcement and evolutionary fitness is then established under the assumption that maximization occurs on both levels.

However, the theoretical framework in Borgstede (2020) does not include the dynamics of learning (as described by the CLOE), but only the possible endpoints of individual learning given that learning maximizes reinforcer value. Inter-individual variation is modeled by an arbitrary class-structure, with evolutionary fitness being the reproductive value weighted sum of an individual's offspring. Like in this paper, individuals are formally treated as their own offspring, thereby bridging the different time scales of natural selection and behavioral selection. The main difference is that the notation is slightly different in Borgstede (2020), using partial derivatives to define the key concepts. These partial derivatives correspond to the partial regression coefficients used to derive the CLOE. Therefore, the marginal change in fitness components y per unit change in behavior (designated $\pi_{xy}(b_x)$ in Borgstede 2020) corresponds to the partial regression coefficients β_{xyp} introduced above. Thus, the definition of "reinforcing power" in the maximization paper is identical to the one derived above and corresponds to the reproductive value weighted sum of fitness effects (i.e., $\sum_y v_y \beta_{xyp}$). Marginal reinforcer value of a behavior b in class x is further designated $r(b_x)$ and defined as the product of the reinforcing power of p and the marginal change in p per unit change in behavioral allocation. Like before, this marginal change (designated p_x in Borgstede 2020) can be identified with the slope of a linear regression of the form $p_{ij} = \beta_{x0} + \beta_{xpb} b_{ij} + \varepsilon$. To retrieve the slope β_{xpb} from the CLOE, one can write the covariance between behavior and p as:

$$\mathrm{Cov}_j(p_{ij}, b_{ij}) = \beta_{xpb} \mathrm{Var}_j(b_{ij}) \tag{12}$$

Following the definition in Borgstede (2020), the marginal reinforcer value of mean behavioral allocation of an individual in class x can be re-stated as:

$$r(b_x) = \beta_{xpb} \sum_y v_y \beta_{xyp} \tag{13}$$

If an individual maximizes this value, it also maximizes its evolutionary fitness (Borgstede 2020). We can further rearrange the CLOE to get:

$$w_i \Delta b_i = \beta_{xpb} \sum_y v_y \beta_{xyp} \mathrm{Var}_j(b_{ij}) + \delta \tag{14}$$

which is equivalent to:

$$w_i \Delta b_i = r(b_i) \mathrm{Var}_j(b_{ij}) + \delta \tag{15}$$

Thus, in the MLBS, behavioral selection equals the product of marginal reinforcer value and (within-individuals) behavioral variance. This means that, if there is no variation in behavior, there is no behavioral selection. Furthermore, behavioral selection only occurs if changes in behavior are associated with changes in reinforcer value, and thus, evolutionary fitness. Consequently, if a behavior is optimal with regard to evolutionary fitness (such as the distribution of individuals over food patches according to the IFD), the MLBS predicts that this behavior will also be selected by individual learning as described by the CLOE. This result integrates individual learning dynamics with the concept of fitness maximization via the abstract principle of behavioral selection.

Conclusion

This paper presents an extension of the multilevel model of behavioral selection (MLBS) to account for inter-individual variation. Variation between individuals is modeled by dividing the population into homogeneous classes. These classes can represent any source of variation between individuals, be it internal or external. Within this framework, a generalized version of the covariance based law of effect (CLOE) was derived. It turned out that the CLOE remains valid even if there is variation between individuals. The only difference to the simple version without classes is that the reinforcing power of a fitness predictor is now defined as a reproductive value weighted sum of fitness effects. This general result was then exploited to close a theoretical gap between evolutionary models of fitness maximization and individual learning dynamics. By linking the theory of behavioral selection to the concept of reproductive value, fitness maximization turns out to be a natural result of individual behavioral adaptations when learning is understood as a selection process as described in the MLBS.

The main implications of the extended MLBS for class-structured populations are thus that (a) applications of behavioral selection models to individuals from an inhomogeneous population are valid and (b) applications of evolutionary maximization models to the outcomes of individual learning are valid. Both results have a high intuitive appeal,

which is probably why they have hardly been questioned in the past. Although the formal justification of these two assumptions has no practical implications, it is of great theoretical value to know that the current modus operandi is indeed theoretically sound.

The formalism presented here is not the first to introduce inter-individual variation into a multilevel Price equation. Gardner (2015) incorporates class-structure into a multilevel Price equation that bears some similarity to the one derived here. Although concerned with group-level and individual-level selection (rather than between-individuals and within-individuals selection as the MLBS), the two approaches address a similar problem. Nevertheless, the MLBS partitioning substantially differs from Gardner's derivation both technically and conceptually. On a technical level, Gardner uses parental reproductive value as the target of selection, thereby allocating class transitions in the offspring generation. Conceptually, this implies that Gardner's partitioning exclusively focuses on selection due to reproduction. However, since the MLBS is concerned with learning (i.e., selection within individuals), the relevant aspect of selection is not reproduction, but survival (i.e., selection due to class transitions). Therefore, although Gardner's formalism is perfectly adequate as a basis of a genetical theory of selection, it does not account for selection that may occur within individuals. In contrast to Gardner's approach, the MLBS, treats parental fitness as the target of selection, which is defined such that offspring are weighted according to the corresponding *offspring* reproductive values. Because a surviving individual may formally be treated as its own offspring, the MLBS accounts for fitness effects due to class transitions, which is a necessary condition for within-individuals selection.

The MLBS aims to provide a general explanatory framework for individual learning in an evolutionary context. It is not, however, intended to be a specific learning model for, say, associative or non-associative processes in a given species. As argued in Borgstede and Luque (2021), the MLBS is best understood as a conceptual framework for the quantitative analysis of behavior from an evolutionary perspective. Within this framework, the CLOE provides the fundamental principle of behavioral selection theory. Like the simple Price equation, given the definitions of the model primitives, the CLOE states a mathematical truth. Any empirical application would require to further specify the constraining conditions and auxiliary laws that are needed to describe a specific learning scenario (cf. Borgstede and Eggert 2023a,b). Given such specializations of the fundamental theoretical principles, the MLBS can indeed provide explanations for various well-known empirical phenomena,

such as conditions of undermatching and overmatching in operant choice (Davison and McCarthy 2016) or the blocking of uninformative stimuli in classical conditioning (Kamin 1969), as shown in Borgstede and Eggert(2021) and Borgstede and Luque (2021). Further theoretical developments may extend the MLBS such that it could also account for adaptive behavior that is based on imitation or instruction. Such work may hopefully shed light on the connection between learning, evolution and culture.

Behavioral selection theory is only at the beginning of its development as a meta-theory of adaptive behavior. Instead of describing how associations are established within an individual by singular events, behavioral selection describes behavior on a larger time scale and is concerned with the long-term interactions between learning individuals and their environment. The MLBS thus provides a general formal framework to derive quantitative laws on a molar level, thereby linking the dynamics of learning to the theory of evolution by natural selection.

Declarations

References

Aoki K, Feldman MW (2014) Evolution of learning strategies in temporally and spatially variable environments: a review of theory. Theor Popul Biol 91:3–19. https://doi.org/10.1016/j.tpb.2013.10.004

Batty CJK, Crewe P, Grafen A, Gratwick R (2014) Foundations of a mathematical theory of darwinism. J Math Biol 69(2):295–334. https://doi.org/10.1007/s00285-013-0706-2

Baum WM (1973) The correlation-based law of effect. J Exp Anal Behav 20(1):137–153. https://doi.org/10.1901/jeab.1973.20-137

Baum WM (2012) Rethinking reinforcement: allocation, induction, and contingency. J Exp Anal Behav 97(1):101–124. https://doi.org/10.1901/jeab.2012.97-101

Baum WM (2017) Selection by consequences, behavioral evolution, and the price equation. J Exp Anal Behav 107(3):321–342. https://doi.org/10.1002/jeab.256

Borgstede M (2020) An evolutionary model of reinforcer value. Behav Process. https://doi.org/10.1016/j.beproc.2020.104109

Borgstede M (2021) Why do individuals seek information? A selectionist perspective. Front Psychol. https://doi.org/10.3389/fpsyg.2021.684544

Borgstede M, Anselme P (2024) Model-based estimates for operant selection. BioRxiv. https://doi.org/10.1101/2022.07.22.501082

Borgstede M, Eggert F (2021) The formal foundation of an evolutionary theory of reinforcement. Behav Process 186:104370. https://doi.org/10.1016/j.beproc.2021.104370

Borgstede M, Eggert F (2023a) Meaningful measurement requires substantive formal theory. Theory Psychol. https://doi.org/10.1177/09593543221139811

Borgstede M, Eggert F (2023b) Squaring the circle: from latent variables to theory-based measurement. Theory Psychol. https://doi.org/10.1177/09593543221127985

Borgstede M, Luque VJ (2021) The covariance based law of effect: A fundamental principle of behavior. Behav Philos 49:63–81

Broadbent DE (1961) Behaviour. Methuen, London

Burgos JE (2019) Selection by reinforcement: a critical reappraisal. Behav Proc 161:149–160. https://doi.org/10.1016/j.beproc.2018.01.019

Campbell DT (1956) Adaptive behavior from random response. Behav Sci 1(2):105–110. https://doi.org/10.1002/bs.3830010204

Caswell H (1982) Optimal life histories and the maximization of reproductive value: a general theorem for complex life cycles. Ecology 63(5):1218–1222. https://doi.org/10.2307/1938846

Caswell H (2001) Matrix Population Models. Construction, analysis, and interpretation, 2nd edn. Sinauer Associates, Sunderland

Caswell H (2010) Reproductive value, the stable stage distribution, and the sensitivity of the population growth rate to changes in vital rates. Demogr Res 23:531–548. https://doi.org/10.4054/DemRes.2010.23.19

Davies NB, Krebs JR, West SA (2012) An introduction to behavioural ecology, 4th edn. Wiley-Blackwell, Hoboken

Davison M, McCarthy D (2016) The matching law. Routledge. https://doi.org/10.4324/9781315638911

Dridi S, Lehmann L (2014) On learning dynamics underlying the evolution of learning rules. Theor Popul Biol 91:20–36. https://doi.org/10.1016/j.tpb.2013.09.003

Dunlap AS et al (2016) Reliability, uncertainty, and costs in the evolution of animal learning. Curr Opin Behav Sci 12:73–79. https://doi.org/10.1016/j.cobeha.2016.09.010

Fawcett TW, Hamblin S, Giraldeau L-A (2013) Exposing the behavioral gambit: the evolution of learning and decision rules. Behav Ecol 24(1):2–11. https://doi.org/10.1093/beheco/ars085

Fisher RA (1930) The genetical theory of natural selection. Clarendon Press, Oxford

Frankenhuis WE, Panchanathan K, Barto AG (2019) Enriching behavioral ecology with reinforcement learning methods. Behav Proc 161:94–100. https://doi.org/10.1016/j.beproc.2018.01.008

Fretwell SD (1972) Populations in a seasonal environment. Monographs in population biology, vol 5. Princeton University Press, Princeton

Fretwell SD, Lucas HL (1969) On territorial behavior and other factors influencing habitat distribution in birds. Acta Biotheor 19(1):16–36. https://doi.org/10.1007/BF01601953

Gardner A (2015) The genetical theory of multilevel selection. J Evol Biol 28:305–319. https://doi.org/10.1111/jeb.12566

Gilbert RM (1970) Psychology and biology. Can Psychol Psychol Can 11(3):221–238. https://doi.org/10.1037/h0082574

Grafen A (2006) A theory of Fisher's reproductive value. J Math Biol 53(1):15–60. https://doi.org/10.1007/s00285-006-0376-4

Grafen A (2015) Biological fitness and the price equation in class-structured populations. J Theor Biol 373:62–72. https://doi.org/10.1016/j.jtbi.2015.02.014

Grafen A (2020) The Price equation and reproductive value. Philos Trans R Soc Lond Ser B Biol Sci 375(1797):20190356. https://doi.org/10.1098/rstb.2019.0356

Herrnstein RJ (1964) Will. Proc Am Philos Soc 108(6):455–458

Herrnstein RJ, Loewenstein GF, Prelec D, Vaughan W (1993) Utility maximization and melioration: Internalities in individual choice. J Behav Decis Mak 6(3):149–185. https://doi.org/10.1002/bdm.3960060302

Houston AI (2008) Matching and ideal free distributions. Oikos 117(7):978–983. https://doi.org/10.1111/j.0030-1299.2008.16041.x

Hull DL, Langman RE, Glenn SS (2001) A general account of selection: biology, immunology, and behavior. Behav Brain Sci 24(3):511–528. https://doi.org/10.1017/S0140525X01004162

Kamin LJ (1969) Predictability, surprise, attention and conditioning. In: Campbell BA, Church RM (eds) Punishment and aversive behavior. New York, pp 279–296

Kraft JR, Baum WM, Burge MJ (2002) Group choice and individual choices: modeling human social behavior with the Ideal Free Distribution. Behav Proc 57(2–3):227–240. https://doi.org/10.1016/S0376-6357(02)00016-5

Lewis RL, Singh S, Barto AG (2010) Where do rewards come from? In: Proceedings of the international symposium on AI-inspired biology at the AISB 2010 convention

Lion S (2018) Class Structure, demography, and selection: reproductive-value weighting in nonequilibrium. Polymor Popul Am Nat 191(5):620–637. https://doi.org/10.1086/696976

Loewenstein Y, Prelec D, Seung HS (2009) Operant matching as a Nash equilibrium of an intertemporal game. Neural Comput 21(10):2755–2773. https://doi.org/10.1162/neco.2009.09-08-854

McDowell JJ (2013) A quantitative evolutionary theory of adaptive behavior dynamics. Psychol Rev 120(4):731–750. https://doi.org/10.1037/a0034244

McNamara JM, Houston AI (2009) Integrating function and mechanism. Trends Ecol Evol 24(12):670–675. https://doi.org/10.1016/j.tree.2009.05.011

Menzies P, Price H (1993) Causation as a secondary quality. Br J Philos Sci 44(2):187–203. https://doi.org/10.1093/bjps/44.2.187

Moore BR (2004) The evolution of learning. Biol Rev Camb Philos Soc 79(2):301–335. https://doi.org/10.1017/S1464793103006225

Price GR (1970) Selection and Covariance. Nature 227(5257):520–521. https://doi.org/10.1038/227520a0

Price GR (1972) Extension of covariance selection mathematics. Ann Hum Genet 35(4):485–490. https://doi.org/10.1111/j.1469-1809.1957.tb01874.x

Price, GR (1995, written ca. 1971). The nature of selection (Written circa 1971, published posthumously). J Theor Biol 175(3):389–396. https://doi.org/10.1006/jtbi.1995.0149

Pringle J (1951) On the parallel between learning and evolution. Behaviour 3(1):174–214. https://doi.org/10.1163/156853951X00269

Rachlin H (1978) A molar theory of reinforcement schedules. J Exp Anal Behav 30(3):345–360. https://doi.org/10.1901/jeab.1978.30-345

Rachlin H, Burkhard B (1978) The temporal triangle: response substitution in instrumental conditioning. Psychol Rev 85:22–47

Rachlin H, Green L, Kagel JH, Battalio RC (1976) Economic demand theory and psychological studies of choice. Psychol Learn Motiv 10:129–154

Rachlin H, Battalio RC, Kagel JH, Green L (1981) Maximization theory in behavioral psychology. Behav Brain Sci 4:371–417

Singh S, Lewis L, Barto AG (2010) Intrinsically motivated reinforcement learning: an evolutionary perspective. IEEE Trans Auton Ment Dev 2(2):70–82

Skinner BF (1966) The phylogeny and ontogeny of behavior. Contingencies of reinforcement throw light on contingencies of survival in the evolution of behavior. Science 153(3741):1205–1213. https://doi.org/10.1126/science.153.3741.1205

Skinner BF (1981) Selection by consequences. Science 213(4507):501–504

Stephens DW (1991) Change, regularity, and value in the evolution of animal learning. Behav Ecol 2(1):77–89. https://doi.org/10.1093/beheco/2.1.77

Stephens DW, Krebs JR (1986) Foraging theory. Monographs in behavior and ecology. Princeton University Press, Princeton

Strand PS, Robinson MJF, Fiedler KR, Learn R, Anselme P (2021) Quantifying the instrumental and noninstrumental underpinnings of Pavlovian responding with the Price equation. Psychon Bull Rev. https://doi.org/10.3758/s13423-021-02047-z

Taylor PD (1990) Allele-frequency change in a class-structured population. Am Nat 135(1):95–106

Thorndike EL (1900) The associative processes in animals. In: Biological lectures from the marine biological laboratory of woods holl, vol 1899, pp 69–91