Phenotypic selection in natural populations: what have we learned in 40 years?

Erik I. Svensson

Department of Biology, Lund University, Lund, Sweden

Corresponding author: Department of Biology, Sölvegatan 37, Lund University, 223 62 Lund, Sweden. Email: erik.svensson@biol.lu.se

Abstract

In 1983, Russell Lande and Stevan Arnold published “The measurement of selection on correlated characters,” which became a highly influential citation classic in evolutionary biology. This paper stimulated a cottage industry of field studies of natural and sexual selection in nature and resulted in several large-scale meta-analyses, statistical developments, and method papers. The statistical tools they suggested contributed to a breakdown of the traditional dichotomy between ecological and evolutionary time scales and stimulated later developments such as “eco-evolutionary dynamics.” However, regression-based selection analyses also became criticized from philosophical, methodological, and statistical viewpoints and stimulated some still ongoing debates about causality in evolutionary biology. Here I return to this landmark paper by Lande and Arnold, analyze the controversies and debates it gave rise to and discuss the past, present, and future of selection analyses in natural populations. A remaining legacy of Lande & Arnold, 1983 is that studies of selection and inheritance can fruitfully be decoupled and be studied separately, since selection acts on phenotypes regardless of their genetic basis, and hence selection and evolutionary responses to selection are distinct processes.

Keywords: agents of selection, environmental covariance, fitness, multivariate selection, phenotypic traits, quantitative genetics

Natural selection is not evolution (Fisher, 1930).

Natural selection acts on phenotypes, regardless of their genetic basis, and produces immediate phenotypic effects within a generation that can be measured without recourse to principles of heredity or evolution. In contrast, evolutionary response to selection, the genetic change that occurs from one generation to the next, does depend on genetic variation. (Lande & Arnold, 1983).

In 1983, Russell Lande and Stevan Arnold published “The measurement of selection on correlated characters” in Evolution, which became one of the most cited papers in this journal (Lande & Arnold, 1983). The first author was the young evolutionary biologist Lande, who had made a name by himself from his work beginning with a key paper in 1976 (Lande, 1976). Lande took statistical tools from the plant and animal breeding literature and merged them with the paleontologist George Gaylord Simpson’s version of the adaptive landscape for phenotypic characters, thereby giving birth to a new discipline: evolutionary quantitative genetics (Arnold et al., 2001; Lande, 1976, 1977, 1979, 1980a, 1980b; Svensson & Calsbeek, 2012). The second author was Arnold, a herpetologist and field biologist interested in animal behavior (Arnold, 1983). In their paper, Lande and Arnold introduced regression analysis as a novel tool to estimate selection simultaneously on multiple phenotypic characters in the field, provided that individual fitness data could be connected to individual trait variation.

Central to the new approach proposed by Lande and Arnold was that selection and inheritance were empirically separable and distinct processes. Thus, estimating selection on a character does not require any information about the genetic basis of the trait. Indeed selection can operate on traits without any heritable basis at all, although then there will of course be no be evolutionary change. This insight was not entirely new; the first sentence in the mathematical population geneticist Fisher’s classical book “The Genetical Theory of Natural Selection” (Fisher, 1930) states something similar (see quotation above); selection is a within-generation process whereby some phenotypes are more successful than others, whereas evolution by natural selection is the transmission of such selection to the next generation, which requires that phenotypes are at least partly heritable (Lewontin, 1970). This key point made by Lande’s former doctoral advisor Richard Lewontin was re-emphasized by Lande and Arnold, but was not entirely new. This message had been made explicit already in 1948 in a pioneering paper in one of the first issues of the new journal Evolution by Michael Lerner and Everett Dempster. They made an explicit analogy between plant and animal breeding and evolution, and they suggested that insights from the former literature could be used to study selection in natural populations (Lerner & Dempster, 1948). However, Lerner and Dempster’s paper seemed to have been largely forgotten in 1983 (and was interestingly not cited by Lande and Arnold), so Lande and Arnold re-introduced their idéa to introduce and methods from the breeding literature to the evolutionary biology community. The Lande and Arnold paper also proposed solutions for how to estimate selection on several characters simultaneously, when characters were correlated with each other, as well as suggestions of how to estimate nonlinear selection.
Here, I discuss the scientific legacy of the paper by Lande and Arnold, the discussions it gave rise to, and the criticisms their approach encountered. I also briefly suggest some profitable future directions of phenotypic selection studies in natural populations in light of the many methodological and statistical advancements that have been made in the four decades since 1983. The title of the present paper has been inspired by similar titles of Perspectives in Evolution on reproductive isolation and speciation (Gavrilets, 2003; Rice & Hostert, 1993). My rationale is that it often takes several decades to evaluate the impact of papers in a slow-moving and largely conceptual field like evolutionary biology.

**The new approach proposed by Lande and Arnold**

The importance of Lande and Arnold’s paper for studies of selection on multiple characters simultaneously (multivariate selection) cannot be overstated. Before their paper, field biologists had typically estimated selection in a univariate fashion and on a trait-by-trait basis (Boag & Grant, 1981; Endler, 1986). Estimating the strength of selection on a single character is relatively straightforward and can be done using the linear selection differential (Falconer, 1989; Figure 1A). When selection operates on a single trait, the evolutionary response to selection (R) is simply the selection differential (S) times the heritability (b²) following the classical breeder’s equation in quantitative genetics:

\[ R = b^2 S \]  

(1)

When selection operates on a single trait, the evolutionary response to selection (R) is therefore perfectly aligned with the direction of selection (S) and the population will move directly to the closest adaptive peak, the rate of evolution only being limited by the additive genetic variance which is part of b² (b² is the additive genetic variance Vₐ, i.e., V/Vₚ; Figure 1A).

However, when traits are correlated with each other, the population will not necessarily follow a straightest path towards the closest adaptive peak, although it might eventually end up there (Figure 1B). Instead, when traits are correlated, the rate of adaptive evolution towards the optimum will be delayed and the population will follow a curved trajectory through phenotype space (Schluter, 1996; Figure 1B).

In the case of such multivariate selection on two or more traits, the individual fitness surface (W) can be estimated as (from equation 3 in [Phillips & Arnold, 1989], modified from Equation 16 in [Lande & Arnold, 1983]):

\[ W = \alpha + \sum_{i=1}^{n} \beta_i z_i + \sum_{i<j} \frac{1}{2} \gamma_{ii} z_i^2 + \sum_{i<j} \sum_{d} \gamma_{ij} z_i z_j + \epsilon \]

(2)

Here, W is relative fitness (absolute fitness divided by mean absolute fitness), \( \alpha \) is a constant (an intercept in a multiple regression), \( \beta_i \) is the directional selection gradient for trait \( z_i \), \( \gamma_{ii} \) is the quadratic selection gradient (indicating concave or convex selection) for trait \( z_i \), \( \gamma_{ij} \) is the quadratic selection gradient for trait interactions between \( z_i \) and \( z_j \) (indicating correlational selection) and \( \epsilon \) is an error term. These selection gradients can be obtained from the partial regression.
coefficients in a standard parametric multiple regression (Lande & Arnold, 1983). Note that to obtain the quadratic selection coefficients (i.e., stabilizing and disruptive selection), the partial regression coefficient in front of the term $\gamma_j^2$ should be multiplied by two (Stinchcombe et al., 2008).

Historical context: the influence of the Spandrels paper by Gould and Lewontin

An adaptationist programme has dominated evolutionary thought in England and the United States during the past 40 years. It is based on faith in the power of natural selection as an optimizing agent. It proceeds by breaking an organism into unitary ‘traits’ and proposing an adaptive story for each considered separately. (Gould & Lewontin, 1979).

Critiques of the “adaptationist program” (Gould & Lewontin, 1979; Lewontin, 1978) stress that adaptation and selection are often invoked without strong supporting evidence. We suggest quantitative measurements of selection as the best alternative to the fabrication of adaptive scenarios...The essential fact is that selection and adaptation can be measured. (Lande & Arnold, 1983).

In motivating their study, Lande and Arnold referred to Stephen Jay Gould’s and Richard Lewontin’s famous paper “The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme” that was published only four years earlier (Gould & Lewontin, 1979). They refer to this paper on the first page in their introduction.

The Spandrels-paper is a highly cited paper in evolutionary biology, more cited than Lande and Arnold (1983), although it has also had four more years to accumulate citations (Figure 2A). Obviously, Lande and Arnold motivated their new approach to estimate selection with an aim increasing the scientific rigor in evolutionary biology. It was precisely this lack of rigor that Gould and Lewontin had criticized when they argued that many biologists just presented adaptive “Just So”-stories without strong evidence (Gould & Lewontin, 1979). Lande and Arnold clearly thought that this lack of rigor that Gould and Lewontin had criticized when they argued that many biologists just presented adaptive “Just So”-stories without strong evidence (Gould & Lewontin, 1979). Lande and Arnold clearly thought that this new method where selection could be estimated and quantified, rather than just vaguely inferred, they would increase rigor and empirical standards, thereby responding to the criticism by Gould and Lewontin. Thus, Lande and Arnold saw selection analyses as a constructive solution to the problem of documenting adaptation and selection.

General impact

Lande and Arnold’s paper had a huge impact, judged by the number of citations following its publication in 1983, particularly in ecology and field studies of selection (Figure 2A). Compared to the Spandrels paper published four years earlier, its main impact has been in empirical studies in ecology, evolution, plant sciences, and agriculture. In contrast, Gould and Lewontin’s paper has had less influence on empirical research in ecology and population biology but has instead influenced other areas of evolutionary biology, developmental biology, and philosophy of biology (Figure 2B–D). Their paper gave rise to a flurry of selection studies in natural populations. This increasingly popular research approach was...
even jokingly called “The Chicago School of Evolutionary Biology” (Grafen, 1988), alluding to the neoliberal economic school led by Milton Friedman that thrived simultaneously at the same university.

Empirical applications: why were Lande and Arnold so successful?

...natural selection is daily and hourly scrutinising, throughout the world, every variation, even the slightest; rejecting that which is bad, preserving and adding up all that is good; silently and insensibly working, whenever and wherever opportunity offers, at the improvement of each organic being in relation to its organic and inorganic conditions of life. We see nothing of these slow changes in progress, until the hand of time has marked the long lapse of ages. (Darwin, 1859).

The success and popularity of the new approach suggested by Lande and Arnold was probably not only because they provided the formal theory behind selection analyses but also because they demonstrated empirical and statistical solutions of how to quantify selection in natural populations. They provided two worked-through empirical examples illustrating their new method. First, they used an old dataset on the mortality of House Sparrows (Passer domesticus) that were found dead after a winter storm and that were collected by Hermon Bumpus and which was compared with live individuals (Bumpus, 1899). They complemented this with a new and similar dataset collected by Arnold along the shores of Lake Michigan on the mortality of pentatomid bugs (Euschistus variolarius), also following a storm. Both these datasets were so-called cross-sectional fitness data, as opposed to longitudinal data, where individuals are followed throughout their lives. Selection on different phenotypes is thus estimated by comparing survivors and nonsurvivors or comparing mated and nonmated individuals (in the case of sexual selection). Such cross-sectional selection analyses have been carried out many times afterwards (Campbell-Staton et al., 2017; Svensson & Friberg, 2007; Young et al., 2004) and they are often the only practical alternative available to estimate selection. In contrast, longitudinal data using life-time reproductive success (LRS) are only possible to obtain for a limited number of species, usually long-lived vertebrates where researchers can mark individuals and follow them over their entire lives (Grafen, 1988).

Using the cross-sectional fitness data on mortality, Lande and Arnold estimated the variance-standardized directional selection gradients ($\beta$-as) on size-related morphological traits to vary between −0.27 and −0.52 for size in the House Sparrows (i.e., selection for smaller birds) and between −0.74 (wing length) and 0.58 (thorax) in the bugs (Lande & Arnold, 1983). These surprising findings indicated unexpectedly strong selection. These estimates indicated that relative fitness would change between 27% and 74% for a change in standard deviation of these traits experiencing selection.

It is important to underscore how unexpected these results were in this era when the neutral theory in population genetics was well-established (Kimura, 1983) and when many were increasingly skeptical of the pervasiveness of natural selection. Neutral theory was preceded by and partly stimulated by a paradox discussed by the population geneticist J. B. S. Haldane’s about the demographic “costs of selection” (Haldane, 1937, 1957). Other findings of strong selection at about the same time as Lande and Arnold’s paper was published (Boag & Grant, 1981) raised the question how populations could persist in the long run in the face of such strong selection. As Lande and Arnold noted themselves, the persistence of a population in the long run requires that the total selective mortality should not exceed the reproductive rate, otherwise the population would go extinct (Lande & Arnold, 1983). They suggested that a solution to this dilemma is that most of the directional selection within a generation may be concentrated in a few relatively short periods of mortality (Lande & Arnold, 1983), whereas during other periods or in other generations selection might be weak or even nonexistent, allowing the population to recover demographically.

Research in the decades after Lande and Arnold have also revealed strong directional selection and sometimes rapid evolutionary change after brief intense selective episodes, such as winter storms (Campbell-Staton et al., 2017), in response to anthropogenic disturbances such as traffic (Brown & Brown, 1998; Price et al., 2000), hunting or fishing pressure (Allendorf & Hard, 2009; Campbell-Staton et al., 2021; Sanderson et al., 2022), or when organisms invade novel environments, such as cities (Santangelo et al., 2022). Natural or sexual selection is often strong, driving rapid evolutionary change in response to new predators, or when organism invade novel selective environments (Endler, 1980; Hendry & Kinnison, 1999; Reznick et al., 1997; Svensson, 2007; Svensson & Gosden, 2009).

The later emerging field of “eco-evolutionary dynamics” can also partly be traced back to the influence of Lande and Arnold and a growing awareness that evolutionary and ecological time scales are often similar and therefore that ecological and evolutionary processes interact and feedback on each other (Hendry, 2016; Schoener, 2011; Svensson, 2019). This is a marked change from 1983 when ecology and evolution were still largely separate fields. Back in 1983, it was often assumed that ecological processes were fast, relatively to evolutionary processes, and that therefore ecologists could largely ignore evolutionary processes in their day-to-day research. A central message from Lande and Arnold was that studies of selection, inheritance, and evolutionary response to selection are conceptually different and can be separated which made it possible for ecologists who only had access to data on fitness components and phenotypic trait data of individuals to contribute to the evolutionary literature by estimating selection using the common currency provided by evolutionary quantitative genetics (Barton & Turelli, 1989; Hansen & Pélabon, 2021; Lynch & Walsh, 1998; Walsh & Lynch, 2018). Lande and Arnold might have contributed to breaking up the barriers between the still separated fields ecology and evolutionary biology. It is worth underscoring that both Laurent Slobodkin, author of Growth And Regulation of Animal Populations (Slobodkin, 1961) and Eric Pianka, author of Evolutionary Ecology (Pianka, 1988) were two of several authors of influential ecology text books who emphasized the distinction between ecological and evolutionary time scales. This dichotomy only started breaking down several decades after Lande and Arnold (1983), catalyzed by an influential paper by Thomas Schoener and the growing field of eco-evolutionary dynamics (Schoener, 2011; Svensson, 2019).
Before Lande and Arnold, there were very few formal selection studies in natural populations and no studies estimating multivariate selection, simply because biologists did not have any statistical tools to carry out such studies. The major architects of the Modern Synthesis (primarily Mayr and Dobzhansky) argued for the pervasive role of natural selection as a major evolutionary process but interestingly none of them estimated selection themselves, presumably because they considered selection to be too weak for such an effort to be worthwhile (Antonovics, 1987; Endler & McLellan, 1988). Thus, it took almost four decades after the Modern Synthesis and after 1983 before biologists regularly started to estimate selection in natural populations. Presumably, many biologists—even those confident about the power of natural selection and its ability to evolutionarily transform populations—still implicitly adhered to Darwin’s view that natural selection was a slow process that it could only be inferred, but not observed directly (Darwin, 1859).

How representative were these strong selection gradients documented by Lande and Arnold? In the first major meta-analysis of published selection gradients in nature, Kingsolver and colleagues found that the average variance-standardized selection gradient across thousands of studies was 0.16 (Kingsolver et al., 2001b). Thus, relative fitness is expected to change by 16% for each standard deviation in a trait, indicating considerable evolutionary potential of natural populations, under the assumptions that traits are at least partly heritable, which they almost always are (Lynch & Walsh, 1998; Mousseau & Roff, 1987; Walsh & Blows, 2009). Issues have been raised, however, about the utility of the variance-standardized selection gradient and it has been proposed that the mean-standardized selection gradient is more appropriate (Hansen & Peabon, 2021; Hereford et al., 2004). A related methodological issue is at what spatial scale fitness should be relativized and traits should be standardized when one is interested in comparing selection among groups or populations (De Lisle & Svensson, 2017).

Another methodological issue is sampling error of the selection gradients. A rough estimate of the extent of sampling error can be obtained from temporal replicated selection studies (Morrissey & Hadfield, 2012). Analyses of a small subset of temporally replicated studies suggest that the mean variance-standardized selection gradient could be as low as 0.05 when sampling error is taken in to account (Morrissey & Hadfield, 2012). However, in population genetic terms, this is still strong selection and would indicate high evolutionary potential of most populations, especially in combination with the existence of large amounts of additive genetic variance of most phenotypic traits as well as fitness itself (Bonnet et al., 2022; Mousseau & Roff, 1987). Finally, the issue about fluctuating selection that was raised by Lande and Arnold as an explanation of their findings has gained some subsequent empirical support (Calsbeek et al., 2012; Gibbs & Grant, 1987; Gosden & Svensson, 2008; Grant & Grant, 2002; Siepielski et al., 2009). However, it is still unclear how much of such observed fluctuations in selection that are due to sampling error vs. real fluctuations (Morrissey & Hadfield, 2012).

**Criticisms and issues about causality**

*The question, “what is the causal relationship between fitness and the characters?” cannot be answered conclusively by an observational approach, simply because the paths of causation, particularly for life-history traits and fitness itself, are so numerous.* (Mitchell-Olds & Shaw, 1987)

The multivariate analysis of selection is insufficient for identifying the causal agents of selection. We discuss how the observational approach of multivariate selection analysis can be complemented by experimental manipulations of the phenotypic distribution and the environment to identify not only how selection is operating on the phenotypic distribution but also why it operates in the observed manner... The biotic and abiotic environment is the context that gives rise to the relationship between phenotype and fitness (selection). The analysis of the causes of selection is in essence a problem in ecology. (Wade & Kalisz, 1990)

In the decade following 1983, influential papers by Thomas Mitchell-Olds, Ruth Shaw, Michael Wade, and Susan Kalisz stand out in criticizing the regression approach to study selection (Mitchell-Olds & Shaw, 1987; Wade & Kalisz, 1990). These and other criticisms (Kingsolver & Schemske, 1991) did not only discuss technical and statistical issues but also raised the deeper question about causal inference. In particular, how can we know that a trait-fitness covariance relationship reflects a causal influence of the trait? It is important to note that the question about causality cannot be solved by statistical methods alone but would require additional biological and ecological data and ideally complemented with functional analysis, experiments, and natural history information (Figures 3–5).

Mitchell-Olds and Shaw (1987) emphasized than an observed selection gradient—even if statistically significant—would not in itself prove that the trait is target of selection, especially if the trait is correlated with other characters that are not included in the statistical analyses. They suggested that any documented selection gradient should be considered as a provisional hypothesis, in need of experimental verification. Experimental manipulations of suspected targets of selection—such as the sexually selected tail length in male widowbirds (Andersson, 1982) or egg size in lizards (Sinervo et al., 1992)—would complement any inferred selection on unmanipulated phenotypic variation (Figure 3). Alternatively, when traits could not be easily experimentally manipulated such as beak sizes in birds or body size, functional analyses (Opedal, 2021), and careful natural history observations are needed before any safe conclusions could be made (Mitchell-Olds & Shaw, 1987). Of particular concern are environmental covariances, such as when individuals vary in condition that independently affects both traits and fitness (Rausher, 2000). Such environmental covariances can lead to a false impression of directional selection on a trait (Price et al., 1988). One solution is to incorporate condition as a covariate in the selection analyses, in effect an additional trait (Rausher, 2000; Stinchcombe et al., 2002), although this is not always feasible. One can also try to verify causal relationships between traits using a combination path analysis, causal modeling, and/or structural equation modeling (Edelaar et al., 2022; Kingsolver & Schemske, 1991; Otsuka, 2019; Shipley, 2002). It is important to emphasize that the multiple regression approach proposed by Lande and Arnold is only a subset of all possible causal relationships between a set of traits and fitness (Figure 4). The underlying assumption in the multiple regression approach is that traits act on the same level in the biological
hierarchy (Figure 4A, cf. Figure 4B–D). Thus, the multiple regression approach is one specific causal model in a greater universe of alternative causal scenarios that can be captured by different path models (Figure 4).

Verifying the causality of trait-fitness covariance relationships is not sufficient, however, for a full understanding of selection. There is also the additional causal layer: the ecology of selection (MacColl, 2011; Wade & Kalisz, 1990). We thus also need to know why the trait-fitness covariance relationship looks like it does, i.e. what is the cause of selection? This is an ecological question: what agents or environmental factors cause this fitness-trait covariance? Natural selection and sexual selection are processes that arise due to interactions between individual phenotypes and their local selective environments (Hull, 1980; MacColl, 2011; Wade & Kalisz, 1990). A full understanding of selection therefore requires not only
knowledge about trait and fitnesses or even the causality of trait-fitness covariances, but also information about how ecological agents and causes of selection—such as competitors, mates, parasites, pollinators or parasites, or abiotic factors such as temperature and precipitation—give rise to trait-fitness covariances. This is typically multidimensional (White & Butlin, 2021), but for simplicity, I here illustrate only two environmental factors and agents of selection. It is assumed that higher predation risk favors smaller individuals, which is shown as weaker selection or even negative selection on body size with increasing predation pressure (from top to bottom). In contrast, higher conspecific density favors larger body size due to increased intraspecific competition, which is shown as steeper and more positive slopes of the fitness functions as one moves from the left to the right (and larger voles). Different combinations of predation and conspecific density can causally interact and shape local selective environments, resulting in different trait-fitness covariances in different populations. In this particular example, the selective environment is thus two-dimensional, but in nature selection is most likely multidimensional. The selective environment can also be described by both conspecific and heterospecific phenotype frequencies and the various social interactions that can arise from such interactions, sometimes in combination with path analytical tools (e.g., Figure 4; see De Lisle et al., 2022; McGlothlin & Fisher, 2022; Wolf et al., 2001). This example illustrates the importance of measuring not only phenotypic traits and fitnesses, but also to quantify and (when possible) experimentally manipulate the local selective environments to gain a full understanding of selection. Silhouettes of the kestrels reproduced with permission from Rebecca Groom under the Creative Commons CC-BY 3.0 license (https://creativecommons.org/licenses/by-sa/3.0/) and voles obtained from Phylopic (http://phylopic.org/). Example inspired by Wade and Kalisz (1990).
high-density populations simply for practical and logistical reasons and the need for large sample sizes.

In summary, Lande and Arnold stimulated some still ongoing discussions about causality in evolutionary biology and in philosophy of biology. These discussions include question at what level selection operates and whether there are “cross over effects” between different levels (Heisler & Damuth, 1987; Okasha, 2006), whether natural selection is a force or only a statistical outcome of lower-level events and the fates of individual organisms (Endler, 1986; Ottsuka, 2016; Sober, 1984; Walsh, 2015; Walsh et al., 2002) and whether genes (“replicators”) or phenotypes (“vehicles” or “interactors”) are the true targets of selection (Ågren, 2021; Dawkins, 1976; Hull, 1980; Lewontin, 1970). Many evolutionary biologists now view phenotypes as the true targets of selection, regardless of their heritable basis, in the spirit of Lande and Arnold (1983).

Alan Grafen’s critique: adaptation vs. selection in progress

In making their claims for their methods, Arnold, Wade and Lande do not always distinguish clearly between the analysis of adaptation and the detection of selection in progress. It is clear, however, that the design of their methods is to detect selection in progress...I believe that most evolutionists and behaviorists would say they were primarily interested in adaptation, as opposed to selection in progress, once the distinction is brought to their attention. Their primary concern is why male red deer have such big antlers, not whether there are genes now changing in frequency that affect antler size... The methods of analysis of LRS data proposed by Wade & Arnold (1984), Lande & Arnold (1983), and Arnold and Wade (1984a,b) seem primarily designed to study selection in progress—that is to say, gene frequencies changing now rather than adaptation. (Grafen, 1988).

Did Lande and Arnold succeed in convincing Gould, Lewontin and other contemporary critics of naive adaptationism? Not really, according to British theoretical biologist Alan Grafen (Grafen, 1988).

Grafen criticized the regression approach suggested by Lande and Arnold for failing to address what he claimed that most biologists really are interested in: adaptation and the current utility of traits (Grafen, 1988). He argued that their approach was more designed to detect selection in progress than to identify adaptations. Grafen criticized such a purely correlational approach, relying on unmanipulated variation in phenotypic traits and fitness, and he argued that biologists interested in adaptation should rather carry out manipulative experiments to clarify the adaptive significance of traits (if any). Grafen’s criticism in a nutshell was thus that Lande and Arnold had conflated selection in progress (an evolutionary process) with adaptation (an optimum phenotypic state of a population) (Grafen, 1988). Following the logic of G. C. Williams (1966), he argued that fitness is a property of design, not a property of an individual, and that using too all-encompassing fitness measures such as Lifetime Reproductive Success (LRS) would not answer the question about the adaptive significance of traits that have their most important function during restricted parts of the life cycle, such as among juveniles or during mating (Grafen, 1988).

Grafen used a hypothetical example of the wing spots on the hindwing of the butterfly Maniola jurtina to illustrate his reasoning. He argued that the really interesting question was the adaptive significance of these hindspots, rather than if they were currently under selection, and he suggested that biologists would gain more insights by experimentally increasing or decreasing the number of spots instead of measuring natural variation in spot number or spot size, using Lande and Arnold’s approach (Grafen, 1988). That is, evolutionary biologists should focus on current utility of traits, rather than on selection in progress.

Although Grafen’s distinction between adaptation and current utility vs. selection in progress is important, it is not always that clearcut. Current utility of a trait implies that the current population trait mean (presumably located at some intermediate optimum) maximizes fitness, compared to alternative variants. This is just another way of saying that the trait is currently experiencing stabilizing selection, thus it is a claim about selection in progress! Moreover, Lande and Arnold’s regression approach was not only designed to detect directional selection, but could also reveal stabilizing and disruptive selection (Lande & Arnold, 1983), so it is strange that Grafen did not embrace this complementary approach to experimental manipulations.

Behavioral ecologists in the British research tradition that Grafen represents tend to focus only on evolutionary endpoints and equilibria, asking questions like: “Is this trait adaptive?” but also tend to ignore the equally interesting question “How did the trait end up here?”. This obsession with evolutionary endpoints and the adaptive significance of traits is quite evident in the research tradition Grafen belongs to, where phenotypic models based on optimization theory and game theory are valued more highly than dynamic quantitative and population genetic models aimed to detect selection in progress. Many evolutionary biologists—the present author included—are more interested in selection in progress than if a trait is an adaptation (or not). Thus, Grafen’s value-laden statement that evolutionary biologists are more interested in whether a trait is an adaptation than they are interested in selection in progress may well reflect his own cultural and scientific bias than the majority of the evolutionary biologists, but that is ultimately an empirical question for historians and sociologists of science to investigate. The historian Tim Lewens has characterized the British research tradition in behavioral ecology and phenotypic modeling as “Neo-Paleyean Biology” (Lewens, 2019), referring to the natural theologian William Paley who in the pre-Darwinian times saw adaptive design everywhere in nature, which he interpreted as a sign of God’s designing ability. Paley made famous the analogy with a watchmaker, and Richard Dawkins openly expressed his admiration of Paley in his book “The Blind Watchmaker” (Dawkins, 1986). The provocative title demonstrates how Dawkins was largely in agreement with Paley that adaptive design is the important question in evolutionary biology. Neo-Paleyean biology today, according to Lewens and Arvid Ågren, is primarily alive in Britain, and with Dawkins, Grafen, and Andy Gardner as its main representatives (Ågren, 2021; Lewens, 2019). Neo-Paleyean biology can be characterized as a research program focused on the adaptive design and current utility of traits but with little interest in evolutionary history or selection in progress (Reeve & Sherman, 1993). However, the distinction between adaptation and selection in progress largely disappears if we realize that claims about
adaptation and current utility are also implicit claims about selection in progress, namely stabilizing selection around a current local optimum (Hansen, 1997).

**Conclusion**

Lande and Arnold’s paper had a long-lasting impact on evolutionary biology, particularly in field ecological studies (Figure 2). It stimulated several discussions about the nature and limitations of statistical tools vs. experiments and general issues about causal inference (Figures 1–5) and it uncovered both the power but also the limitations of selection and adaptation. The main influence of their paper was providing a useful empirical tool that resulted in hundreds of field studies documenting and quantifying natural selection (Figure 2). This stimulated several influential meta-analyses that have enriched our understanding about the strength and variability of phenotypic selection in natural populations (Kingsolver & Diamond, 2011; Kingsolver et al., 2001b; Siepielski et al., 2009, 2011, 2013, 2017). Given this large body of empirical work, what remains to be done and what is the future of selection studies in natural populations? Five remaining challenges come to my mind.

First, measuring and analyzing individual phenotypes is time-consuming and a major bottleneck. New automated data collecting techniques and high-throughput phenotyping (“phenomics”) combining digital data with tools from machine learning and Artificial Intelligence, including Computer Vision can hopefully overcome some of the bottlenecks of limiting sample sizes in selection studies (Lürig et al., 2021). However, formidable challenges remain to quantify fitness or fitness components in the field.

Second, our knowledge about multivariate selection—including various forms of nonlinear selection—still lags behind our knowledge about directional selection. In particular, how common is stabilizing vs. disruptive selection? How common and strong is correlational selection, i.e., selection for trait combinations, compared to selection on traits in isolation and what are the genomic, developmental, and evolutionary consequences of such selection (Sinervo & Svensson, 2002; Svensson et al., 2021)? There is still no major meta-analysis of correlational selection, largely because this form of selection is seldom quantified in field studies, in spite of statistical tools being available (Blows, 2007; Blows et al., 2003, 2004; Phillips & Arnold, 1989; Svensson et al., 2021).

Third, what are the demographic and life-history consequences of phenotypic selection on individuals for population growth rate, extinction risk (Martins et al., 2018), or evolutionary rescue (Bell, 2017)? This is an area that is largely unexplored empirically, although the theoretical framework has been available for decades (Lande, 1982).

Fourth, how can we better integrate ecological selection studies in the field with research on phenotypic plasticity and development? A theoretical and analytical framework is available to quantify selection on function-valued traits (Stinchcombe & Kirkpatrick, 2012), such as reaction norm slopes and intercepts, but empirical studies are still few, largely because of the need of large sample sizes (Chevin et al., 2010; Kingsolver et al., 2001a; Lande, 2009; Svensson et al., 2020).

Finally, how can we connect short-term ecological studies of selection on fitness components to the macroevolutionary time scales that are the focus of phylogenetic comparative studies (Uyeda et al., 2011)? In particular, how are phylogenetic signatures of multiple optima that are often interpreted as stabilizing selection (Beaulieu et al., 2012; Hansen, 1997) related to the estimates of stabilizing selection in microevolutionary studies? Solving these challenges will require close collaborations between experimental and comparative evolutionary biologists, empiricists, and theoreticians with complementary expertise.

**Funding**

My research has been funded by the Swedish Research Council (VR; grant no. 2020–03123).

**Data availability**

The only data in this paper are from the citation analyses in Figure 2. These citation data are available on Dryad Digital Repository (https://datadryad.org/stash): https://datadryad.org/stash/dataset/doi:10.5061/dryad.cjsxksnb6.

**Author contribution**

E.I.S. came up with the idea to write this paper, carried out all the analyses, made all the figures and wrote the manuscript.

**Conflicts of interest:** The author declares no conflict of interest.

**Acknowledgments**

I thank Editor Tracy Chapman for encouraging me to write this paper and for kindly inviting me to submit it as a Perspective-article in *Evolution*. I am grateful for constructive comments by Arvid Ågren, Masahito Tsuboi, Stephen De Lisle, Laura Sophie Hildesheim and several members of “Svensson Lab” on an early version of the manuscript.

**References**


Downloaded from https://academic.oup.com/evolut/article/77/14/1392/14837146 by guest on 28 September 2023