Scientific Research Publishing

Open Journal of Philosophy, 2026, 16(2), 258-273
https://www.scirp.org/journal/ojpp
ISSN Online: 2163-9442
ISSN Print: 2163-9434

Critique of Williams' Opposition to Evolutionary Progress: An Argument for Evolutionary Progress from an Individual Life History Perspective

Xinyu Zou, Jianhui Li

School of Philosophy, Beijing Normal University, Beijing, China
Email: zouxy@mail.bnu.edu.cn

How to cite this paper: Zou, X. Y., & Li, J. H. (2026). Critique of Williams' Opposition to Evolutionary Progress: An Argument for Evolutionary Progress from an Individual Life History Perspective. *Open Journal of Philosophy, 16,* 258-273.
https://doi.org/10.4236/ojpp.2026.162016

Received: April 29, 2026
Accepted: May 16, 2026
Published: May 19, 2026

Abstract

George C. Williams put forward a famous critique of evolutionary progress in 1966. As the theories supporting evolutionary progress have gradually developed, most of Williams' viewpoints had been well responded to. For example, evolutionary progress can be discussed from a non-anthropocentric stance. However, there is still one argument in Williams' critique remaining unanswered: when specialization is used as the standard for evolutionary progress, specialization often manifests in an improvement in one adaptation accompanied by a decline in another. This characteristic makes it impossible for specialization to support cumulative progress. This paper argues that the core point of Williams' critique is that relying solely on the traditional population-level framework, there is no clear and feasible method to measure the positive and negative impacts brought about by functional improvement. Based on Howard's Selective Effect Theory, this paper constructs two idealized models: First, a single selective pressure model establishes the ontological reality of evolutionary progress, thereby countering Williams' claim that evolutionary progress lacks any empirical support; Second, by incorporating individual life history event frequencies into a multiple selective pressure model, we demonstrate that the benefits and costs of a given functional improvement can be quantitatively integrated. This approach ultimately resolves the challenge of identifying evolutionary progress amid concurrent adaptation enhancement and decline. This paper argues that based solely on the SE theory, we can only retrospectively identify progress. Only when we additionally have information about whether the environment has changed can we achieve a certain degree of rough prediction of evolutionary trends.

Keywords

Evolutionary Progress, George C. Williams, Selective Effect Theory, Natural Selection, Life Histories of Individuals

1. Introduction

Since the emergence of evolutionary theory in the early 19th century, debates surrounding "Evolutionary Progress" have persisted. This paper refrains from offering a universally accepted definition of the term, as definitions of the concept have been extraordinarily diverse since its emergence. Rosslenbroich have documented at least 31 distinct definitions of "Evolutionary Progress" throughout the 20th century (Rosslenbroich, 2006). Nevertheless, we may broadly understand it as addressing the following question: Over time, do biological organisms become better (whether in specific aspects or across the entirety of life's history)? If so, how should we characterize this process?

In the debate about "whether evolution is progressive" over the past half-century, well-known scholars who oppose evolutionary progress include George C. Williams, Stephen Jay Gould and others. In his 1966 masterpiece *Adaptation and Natural Selection*, Williams put forward a decisive critique of the view supporting evolutionary progress. He classified past definitions of evolutionary progress into five categories: as accumulation of genetic information; as increasing morphological complexity; as increasing physiological division of labor; as any evolutionary tendency in some arbitrarily designated direction; as increased effectiveness of adaptation. However, in the specific discussion, Williams' critique of evolutionary progress was only divided into four parts, because the two categories of definitions, as increasing physiological division of labor and as any evolutionary tendency in some arbitrarily designated direction, were integrated by him into as increasing of specialization (Williams, 1966: pp. 20-55).

After Williams' critique, scholars supporting evolutionary progress needed to respond to his arguments, even when not explicitly addressing him. Wimsatt noted that Williams' reductionist stance on genic selection was theoretically linked to his opposition to evolutionary progress. As the wave of reductionism waned in recent years, the concept of evolutionary progress reemerged in scholarly discourse. Wimsatt emphasized that Williams' critique did not cover all forms of evolutionary progress concept; it could still be used in alternative frameworks, but required some prior qualifications. Clarifying these qualifications helped avoid attacks from opponents like Williams (Wimsatt & Schank, 1988). Ayala argued that the evolutionary progress concept must be operationalizable, and he modified the progress standard proposed by Kimura Motoo and Huxley J., suggesting that a revised definition could yield meaningful insights into the history of life (Ayala, 1988). Ayala's work responded to Williams' view that "there is no empirical evidence to support the evolutionary progress." Dawkins, through his analysis of the evolutionary progress concept, observed that some meanings became blurred due to anthropocentric presuppositions. He contended that terms like "higher" and "lower" were too vague and should be replaced with precise terminology (Dawkins, 1992). Dawkins proposed the arms race model (for short-term progress) (Dawkins, 2006: pp. 178-193) and defined progress as "the accumulating number of features contributing towards whatever adaptation the lineage in question exemplified," (Dawkins, 1997) successfully sidestepping anthropocentric biases. Rosslenbroich differentiated between the essential and non-essential elements of the evolutionary progress concept through conceptual analysis. Building solely on essential elements, he argued that evolutionary progress must describe four levels: change, order, direction, and general patterns, while excluding normative (specifically anthropocentric) content (Rosslenbroich, 2006). The concept of progress that met Rosslenbroich's requirements could avoid Williams' doubt that supporting evolutionary progress is influenced by anthropocentrism.

In my paper "Is There No Objective Standard of Evolutionary Progress? Williams' Refutation of Criterion of Evolutionary Progress and Criticism of Williams", which was published in 2025, I refuted most of the views put forward by Williams. Based on the SE (selective effect) theory proposed by Andrew S. Howard in 2020, I hold that the evolutionary progress is a norm existing in nature. I also established a simplified model to distinguish short-term progress from long-term progress, refuting Williams' core view that "no matter which standard of progress is adopted, there is no empirical evidence to support evolutionary progress". In addition to this most important criticism, I also pointed out that many of Williams' other views are not valid: He claimed that believing in the evolutionary progress and the inevitable trend of evolution is merely an embodiment of anthropocentrism and orthogenesis, which ignores the norms that exist in nature; He overemphasized the principle of simplicity and excluded a lot of scientific content that conforms to academic norms from biological research; He set the boundary of genetic information accumulation in the Cambrian period, and this demarcation has no basis; Without distinguishing between the predictive perspective and the retrospective perspective, he completely denied the existence of cumulative progress. I believe that we can still identify cumulative progress from a retrospective perspective, but from a predictive perspective, evolutionary progress does not necessarily exist (Zou & Li, 2025).

However, I also realized that my work in 2025 did not completely refute Williams. At that time, I evaded one of his very thorny core viewpoints. When criticizing "specialization[1] as a criterion for evolutionary progress", Williams put forward: "Such tissue specialization is apparently acquired at the price of regenerative

---

[1] Specialization is a specific adaptation of organisms to unique environments (such as specific foods, climates, or habitats), sometimes manifested as the over development of local structures or the specialization of functions. Darwin first proposed this criterion in "*On the Origin of Species* (3rd)" (Darwin, 2006: p. 222).

abilities. To a certain extent this implies the substitution of one adaptation for another, not merely additional adaptations." (Williams, 1966: p. 47) The meaning of this view is that when a certain trait change can bring both benefits and disadvantages to an organism, we cannot identify cumulative progress from it, nor can we assert that this change is overall better for the organism. A variant of this view is: when multiple selective pressures act on an organism simultaneously, how should we identify the evolutionary progress? For example, the larger and more conspicuous tail of a male peacock can help it gain an advantage in sexual selection, but at the same time, it is more likely to be discovered by predators, bringing a survival disadvantage. Obviously, there are two completely opposite selective pressures here. This example may not be rigorous enough, but it is undeniable that whether it is specialization or any other form of biological improvement, it is often accompanied by corresponding costs. Therefore, Williams' views really hit the fatal weakness of those who support evolutionary progress.

Since this view was put forward in 1966, it has not received a good response for a long time. Among the scholars who support evolutionary progress, Richard Dawkins proposed the metaphor of the arms race, which provided a strong argument for the objective reality of evolutionary progress on a short-term scale. Dawkins did realize that organisms may face multiple different arms races simultaneously (Dawkins, 2006: pp. 181-182), but he did not discuss how to analyze the process and results of the arms races in this situation. Howard modified the SE theory and discussed a kind of evolutionary progress defined as "functional improvement"[2] without using teleological expressions. Howard believes that we can only recognize the existence of evolutionary progress (i.e., a certain trend of functional improvement) from a retrospective perspective, and he also provided an analytical framework covering the ontological, epistemological, and methodological levels (Howard, 2020). However, like Dawkins, he failed to respond well to Williams' question: through Howard's theory, we still don't know exactly how organisms achieve progress under the action of multiple selective pressures; we can only confirm the existence of the result of functional improvement. Based on Howard's research, I constructed a "simplified model of single selective pressure". As I mentioned earlier, this model can already refute most of Williams' views in 1966, but to completely refute Williams, this model needs to be extended to the scenario of multiple selective pressures.

In the second section of this paper, I will first briefly review Howard's SE theory and two limitations I identified. The third section will present an idealized model of single selective pressure constructed based on this theory. Then, the fourth section will focus on the issue of identifying evolutionary progress under multiple selective pressures. This paper argues that by quantitatively integrating the occurrence frequencies of different types of events from the perspective of individual

---

[2] "Functional improvement" denotes the real and potential quantifiable improvement in the performance of a specific function by a trait, influenced by natural selection (Howard, 2020: pp. 87-89).

life history[3] on the basis of the causalism population-level perspective, this problem can be solved. In addition, the model proposed in this paper can avoid the two limitations of Howard's theory under specific additional conditions.

2. Howard's Selective Effect Theory and Its Limitations

Howard's SE theory opens with a critique of the vague teleological formulation employed by previous accounts of SE theory. Traditional SE theory often draws on the analogy of artifacts, employing normative statement to describe biological traits and functions. However, this analogy carries a clear flaw: artifacts are designed by human agents, while there is no conscious designing agent behind the biological functions. When we judge whether an artifact is "functioning properly", we are essentially evaluating whether it meets the designer's psychological expectations, rather than merely describing its physical properties. Directly applying this mode of expression to biological traits will lead to unclear reference regarding the source of function—since there is no conscious agent to "prescribe" what the function of a given biological trait must be, normative statements such as the "obligation" of a trait, or what function a trait "supposed" perform, become difficult to make sense of (Howard, 2020: pp. 73-78).[4]

To address this problem, Howard does not follow the "naturalising normativity" approach of traditional SE theory, which interprets evolutionary phenomena in normative terms. He instead turns to "normativising nature", the position that natural norms are inherent to evolutionary phenomena themselves. This work is grounded in the further taxonomy of normativity proposed by Christensen in 2012 (Christensen, 2012). Christensen divides normativity into evaluative norms and non-evaluative norms. The latter category has received very little philosophical attention, as it is typically treated as merely an issue of measurement. Non-evaluative norms can be further subdivided into performance norms and constructive norms: "The ranking of 50-meter sprinters by their race results" falls into the category of performance norms, which inherently includes evaluations of better or worse performance. "The competition rules of a 50-meter sprint" belongs to constructive norms, which refer to rules and customs.

Based on Christensen's work, Howard argues that when two populations share a common selective history[5], this is analogous to 50-meter sprinters competing under the same set of race rules—uniform natural selective pressures also form a constructive norm, much like the rules that govern a competition. On this basis,

---

[3] This is a more detailed perspective than the individual perspective. From birth to reproduction, organisms will experience a variety of specific events, such as "trait a performs function x to obtain enough food", "trait b performs function y to avoid being detected by predators", "trait c performs function z to successfully mate", etc.

[4] Although Howard didn't explicitly point out who the supporters of the old SE theory he was criticizing were. We can search the literature before 2020 for verification. For example, in a literature from 2017: "Something has gone wrong for these mice—they are not the way they ought to be." (Matthewson & Griffiths, 2017: p. 453). There is indeed a teleological expression here.

[5] Common selective history means that two populations have been under identical selective pressures over a period in the past.

comparing the functional capacity of different variants of a given trait counts as a measurement at the level of performance norms. As long as the measurement approach is valid, the result—just like the ranking of 50-meter sprinters—should be objective and uncontroversial. When there are differences in functional performance between trait variants, two possible scenarios arise: The common case is gradual functional improvement driven by natural selection; The rarer case is evolutionary novelty arising from non-selective factors, where the trait performs the same function at a higher level through a different approach. Only the first scenario can be invoked by Howard's SE theory to support the reality of evolutionary progress (Howard, 2020: pp. 79-98). In sum, within this framework, terms such as "better" and "worse" do not carry the type of normative meaning with ambiguous reference that plagued earlier versions of SE theory; instead, they are objective descriptions of changes in functional capacity over the course of evolution.

Howard next turns to address methodological issues. In empirical case studies, researchers must always isolate a specific segment from the vast expanse of life's evolutionary history for analysis, which inevitably raises questions concerning the delineation and demarcation of evolutionary lineages. First, the concept of lineage Howard employs is broadly defined, including but not limited to lineages formed by species, individual organisms, and traits. Second, he advances a refutation of Gould's position on this issue: Gould argues that all lineage demarcation is purely human-constructed and arbitrarily subjective. Howard points out that Gould is actually criticizing the view that "some lineages are inherently more important or more central than others", and that Gould's metaphor of evolutionary process as a "labyrinthine route" is misplaced. Although it is true that researchers may fail to reconstruct the true lineage due to insufficient evidence in practical research, we can still maintain at the ontological level that a clear, continuous path connects modern species to their distant ancestors, stretching from the origin of life to the present day (Howard, 2020: pp. 147-151).

The second core methodological question concerns what approach researchers should adopt for further lineage subdivision, which also requires a response to Gould's critique that lineage demarcation must avoid researchers' subjective arbitrariness. Two approaches to lineage subdivision exist: The first approach demarcates lineages on the basis of time or generational counts, which has the advantage of being fully objective. However, since trait evolution rarely proceeds at a constant rate, a second approach is also used, which subdivides lineages in line with the researcher's specific topic. It is true that the second approach carries the risk of subjective arbitrariness that Gould identified. Howard argues that researchers can nevertheless avoid this problem: including extraneous samples irrelevant to the research topic, or omitting samples that should be included, will produce a reconstructed lineage that delivers misleading or incomplete information, giving researchers a clear standard to correct for bias. The second approach also gives rise to another common misunderstanding: when an early sample is included in a curated lineage sequence, this inclusion is not based on researchers' ability to foresee that the lineage will inevitably evolve toward the target set by the research topic. Instead, the decision whether to include the sample is made retrospectively, based on evolutionary outcomes that occurred in the period after the sample first emerged (Howard, 2020: pp. 152-161).

Later, Howard cites two classic cases from prior research—the evolution of the Cit+ trait in bacteria and the evolution of the eye lineage—to demonstrate that even when lineage demarcation is conducted on the basis of a specific research topic, the curated lineage can still provide meaningful information about lineage evolution (Howard, 2020: pp. 164-174).

The final step of Howard's argument extends the scope of SE theory. He introduces an independent method external to SE theory, designed to identify non-species-specific functions that transcend species boundaries. Even when traits from different species differ dramatically (such as the various traits that make up the eye lineage), as long as these traits consistently perform the same function, the premise that different species are governed by the same set of natural rules remains intact. As a result, interspecific comparison of biological functions can be achieved to a certain extent (Howard, 2020: pp. 175-179).

In summary, Howard provides a framework for the objective comparison of functional capacity, and thereby argues that evolutionary progress is a necessary process-property of the complex adaptive traits ubiquitous in the biological world. Because the comparisons generated by this framework belong to non-evaluative norms, it transforms a question vulnerable to anthropocentric bias into a straightforward problem of measurement and comparison. It is fair to conclude that Howard has achieved preliminary success in reaching the goal of objective comparison of functional capacity.

Nevertheless, Howard's SE theory has two limitations: First, in Howard's definition of evolutionary progress, progress refers only to functional improvement, but he never connects functional improvement to individual fitness. This omission leaves many concrete evolutionary scenarios impossible to explain coherently. We can easily imagine a plausible case: a trait in Species A undergoes functional improvement driven by natural selection and becomes more structurally specialized. But increased specialization comes at the cost of reduced regenerative ability: injured individuals struggle more to recover to full health, which clearly reduces individual fitness. In this scenario, functional improvement does not necessarily bring greater adaptiveness, so why would natural selection still drive a directional trend toward this kind of functional improvement? Second, Howard's identification of evolutionary progress relies exclusively on a retrospective perspective, and the theory has no predictive capacity. This greatly limits the theory's explanatory power; it would obviously be far more robust if it could be endowed with predictive ability. For these two reasons, I aim to remedy the shortcomings of Howard's SE theory by constructing a simplified model to compensate for its gaps.

3. A Model of Single Selective Pressure

Based on Howard's SE theory, the logical structure of our model is as follows: First, we already know that a certain trait is performing function a. From the perspective of the SE theory, the identification of the function is retrospective. Second, based on this knowledge, we trace back the evolutionary lineage of the target species. According to the theme of "gradual improvement of function a", we include the ancestral trait samples that match this theme into the editing sequence. Third, according to the editing sequence we constructed, from the early to the later samples, function a is gradually improving. Fourth, according to the SE theory's definition of evolutionary progress, the evolution of this trait lineage is considered progressive because of the continuous improvement of the function. In the third step, Howard omitted a large number of non-negligible details: namely, how exactly did function a of the early samples evolve into the function a with better execution ability today. From a retrospective perspective, these details are skipped, and we can only get the fact that "function a with a high execution level already exists". However, our model needs to provide predictive guarantees and also explain the relationship between function improvement and individual fitness. Therefore, such omissions are unacceptable to us.

The argument in this paper needs to solve these problems. The simplified model used no longer emphasizes the first steps, two and the focus is on presenting the specific process of the third step. As is well known, evolution is a complex phenomenon, which is the result of the combined action of factors such as the physical environment (temperature, air pressure, humidity, etc.), the biological environment (competitors, parasites, predators, etc.), the organism's own mutations, natural selection, random forces, and the existing physiological structure. If we want to comprehensively consider all the above factors, it will make the argument difficult to advance. Therefore, we start with the most simplified model: based on the adaptationism stance of the SE theory, starting from natural selection, we presuppose that there is only one selective pressure and ignore other factors driving evolution. In our model, we never consider the differences in mutations and the existing physiological structure. As for the physical environment, we adopt Dawkins' presupposition that the environment remains unchanged when discussing short-term progress. As for the biological environment, to make the argument more concise, we adopt the intraspecific arms-race model and only need to consider the individuals within the target population.

Suppose in a certain area, the function of the body color of species a is camouflage. At time t0, the similarity between the body color of ordinary individuals in the population and the living environment is 15%. At the same time, there are also mutant individuals, and the similarity between their body color and the living environment luckily reaches 16%. We infer from the similarity between the body color and the environment that the higher the similarity, the better the camouflage ability, the lower the probability of being noticed by predators, and the higher the survival rate of individuals. Since this is an intraspecific arms-race, predators do not specifically prey on species a. Species a only accidentally enters the predators' diet. Therefore, predators will not correspondingly improve their own recognition ability because species a gradually improves its camouflage ability. We only need to consider the situation of species a. Since the survival rate of ordinary individuals is slightly lower than that of mutant individuals, we can reasonably infer that mutant individuals with a 16% similarity to the environment will gradually spread within the population under the action of natural selection and eventually become the dominant genotype in the population. Suppose that by time t1, individuals with a 16% similarity have become the new ordinary individuals in the population, and then at time t2, new mutant individuals with a 17% similarity will be born. This process of mutant individuals replacing ordinary individuals will be repeated continuously. By analogy, we can see the trend that the body color of species a becomes more and more similar to the living environment, that is, the gradual improvement of the camouflage ability.

Since we presuppose that the environment where species a is located remains unchanged, it means that the situation of "the camouflage ability of species a suddenly declines due to environmental changes" will not occur. Under the premise of an unchanged environment, we can completely predict that the body color of species a can eventually be perfectly adapted to the living environment, and it is extremely difficult for predators to identify them only by eyesight.

Regarding the presupposition of an unchanged environment, I need to clarify the difference between my theory and Howard's. Howard adheres to the retrospective perspective, which I think is reasonable because the SE theory itself does not contain a commitment to environmental stability. The predictability of the model in this paper that I emphasize needs to meet an additional condition outside the SE theory to be realized: we need to obtain information about the future environmental change trend first. In this case, the living environment of species a always remains unchanged, which means that natural selection has formed a set of stable competition rules. This set of competition rules and the rules of the human 50-meter sprint competition both belong to the constructive norms in non-evaluative norms. The only difference is that the rules of the 50-meter sprint are artificially formulated, while the competition rules of "who has better camouflage ability" are "formulated" by the unconscious natural selection mechanism.

When we describe who is the champion of the 50-meter sprint, this is, first of all, a factual description, but it also contains a certain value judgment (performance norm): the implied meaning of this value is that under this set of rules, the champion is the best. Similarly, when we describe that "the proportion of mutant individuals of species a in the population is gradually increasing", we not only describe the fact that "mutant individuals have a higher survival rate than ordinary individuals because of their better camouflage ability", but also contain the value judgment that "mutant individuals are superior to ordinary individuals". This value judgment only makes sense under the rules of the camouflage competition. This analogy can help us intuitively understand what the "norms in nature" are, but we need to note the core difference between the 50-meter sprint and the camouflage competition: in human society, the evaluation of people is often multi-dimensional. The 50-meter sprint champion is not equal to "the best human individual" because in the multi-dimensional evaluation system, the definition of "better" is usually not clear. However, in our model, since there is only one selective pressure, there is also only this unique evaluation system. In this case, the superiority or inferiority of the function can be directly equivalent to the superiority or inferiority of individual fitness.

So far, we have broken through the two limitations of Howard's theory in the model of single selective pressure. Next, we will turn to the discussion of multiple selective pressures.

4. A Model of Multiple Selective Pressures

Why cannot we transition directly from the single selective pressure model to a multiple selective pressures model? The core difficulty has already been revealed at the end of third section: it is difficult to evaluate a person objectively and fairly under a multi-dimensional evaluation system. We do not need to introduce many evaluation rules, simply considering the combination of two evaluation systems at the same time is enough to illustrate this difficulty.

Suppose that in a certain society, there are only two criteria for evaluating whether a student is excellent: physics grades and PE (physical education) grades. How can we rank students objectively? The first fair solution that comes to mind is to treat the two criteria as equally important, meaning each grade accounts for 50% of the weight in the final comprehensive evaluation. But it is not hard to imagine that if this society values physical fitness more, the public may argue that the weight of PE grades should be increased rather than split equally: for example, physics would account for 30% and PE for 70% in the final comprehensive evaluation. If there are 100 students in this society, we can fully predict that the ranking of all 100 students will change between the two evaluation systems.

Here we encounter a problem: unlike the 50-meter sprint case with a single evaluation system, once we try to synthesize multiple one-dimensional evaluations, the appropriate weight for each dimension is largely subjective and arbitrary. Students who are better at physics will obviously prefer a system that assigns a higher weight to physics grades, while students better at PE prefer a system with a higher weight for PE. As a result, which comprehensive evaluation system to adopt depends entirely on the subjective attitude of the rule-maker. Therefore, an objective comprehensive evaluation of ability under a multi-dimensional system seems impossible.

If we apply this conclusion to the discussion of evolutionary progress, then Williams' critique seems correct. If Williams were to critique our concept of evolutionary progress, he would most likely argue as follows: functional improvement (referring to cases that lead to further specialization of a trait) only improves one adaptation of an organism, but at the cost of decreasing another adaptation. Scholars who believe in the existence of evolutionary progress only focus on the side where one adaptation is improved, while ignoring the side where other adaptations decline. This essentially reflects that the evaluator has already taken humans as the reference point of the evolutionary endpoint, which is a typical manifestation of anthropocentrism.

But we still have a way to make an objective argument for evolutionary progress. Do not forget that we have already pointed out: the core difference between "human sports rules" and "the rules of arms-race formed by natural selection" is that the latter does not have an artificial rule-maker. Under multiple selective pressures, individual fitness[6] equals the sum of fitness advantages and disadvantages brought by each function. Therefore, in the case of multiple selective pressures, the quality of a single function cannot be equated with the overall quality of individual fitness. This is probably the reason why Howard never managed to connect the two concepts.

The way out of this problem is that natural selection itself has no rule-maker, which means that the contribution proportion of each part to individual fitness does not depend on researchers' subjective preferences, but is determined by natural selection itself. As long as we can find a way to objectively and clearly identify the actual contribution of each part to individual fitness, it is possible to give an objective description of the quality of individual adaptation.

I have read Richard Dawkins' *Climbing Mount Improbable* in full, and I am deeply impressed by the content in the first chapter, where he simulates the evolutionary process under natural selection via a computer program (Dawkins, 1996). I originally intended to follow Dawkins' example and design a more complex program to achieve this goal, presenting all the details of simulated life history clearly so that the contribution proportion of each function to individual fitness would be completely objective. However, I have zero programming foundation and do not have enough time to complete this work, so this path is not feasible for me.

Later, I learned about the debate between causalism and statisticalism in natural selection research. Howard stands on the causalist position and has criticized statisticalism, which he refers to as non-causalism. Howard points out that in most cases, the difference between causalism (the population-level perspective) and non-causalism (the individual-level perspective) is only a difference in explanatory grain, rather than a difference in causal reality. But in frequency-dependent models, the limitations of the non-causalist individual perspective are fully ex-

---

[6] In this model, the definition of "individual fitness" is derived from the classic Propensity Interpretation of Fitness (abbreviated as PIF), which refers to the average number of offspring expected to be produced by individuals of a certain genotype (Bourrat et al., 2024). Although this definition has flaws in more complex situations, these flaws will not be exposed in the simplified model of this paper. This is only the definition from the population-level perspective. Since the perspective of individual life history is added in this model, "individual fitness" can also refer to the actual number of offspring of a particular individual. To avoid confusion between the two levels of definitions, I will make distinctions by clarifying the context.

posed (Howard, 2020: pp. 27-34).

After that, I tried to sort out the core arguments of non-causalism: in the 2009 paper Selection and Causation, Matthen and Ariew argue that the proposition "Variation with respect to trait T caused evolutionary change in population P" cannot be rephrased as "natural selection caused evolutionary change in population P". This is because natural selection is merely an instrumental term: it is a mathematical product derived from physical events at the individual level (such as predation and reproduction), and does not possess independent causal efficacy of its own (Matthen & Ariew, 2009). Since I share the same position as Howard, even if natural selection is indeed just a mathematical entity as Matthen claims, I cannot accept the view that "natural selection has no real causal efficacy". This paper will not further expand on the debate between causalism and statisticalism. What I need to emphasize here is that I gained inspiration from the formulations of statisticalism, which made me realize that this problem can be solved at the level of individual life history.

From the causalism position at the population-level, I will restate the definition and calculation method of "individual fitness" mentioned above: the "individual fitness" we discuss actually refers to the average fitness of a specific genotype in a population. To facilitate calculation and reasoning, we do not need to assume multiple independent functions. We only need to assume that a single functional improvement has both positive and negative effects. Next, I will continue to discuss the male peacock tail case mentioned in Section 1.

I simplify the entire process of a male peacock from birth to the verge of reproduction into two types of individual life history events: one is "being favored by females and successfully reproducing", and the other is "being detected by predators and dying". We further simplify the setting of the tail, only considering tail size and ignoring other traits such as tail feather color richness. Even if all male peacocks have the same tail size, different individuals may still experience different life history events. There are two opposing selective pressures here: On one hand, a larger tail is conducive to gaining an advantage in sexual selection and helps individuals successfully reproduce offspring; On the other hand, a smaller tail reduces the probability of being detected by predators, helps individuals survive successfully, and thus gives them the opportunity to reproduce in the future. Therefore, the average fitness of a male peacock individual is equal to the sum of the adaptive advantage (mating opportunity) and adaptive disadvantage (predation risk) brought by the tail.

In the tail lineage we constructed, the genotype corresponding to the starting sample of the lineage is called genotype a: suppose there are 100 male peacocks carrying genotype a in the target population. 70 of them experienced the event of "being favored by females and successfully reproducing", and each individual produced 2 offspring; the remaining 30 individuals experienced "being detected by predators and dying" and failed to reproduce. For successfully reproducing individuals, their fitness is recorded as 2[7]; for dead individuals that failed to reproduce, their fitness is recorded as 0. From this we can calculate that the average individual fitness of male peacocks with genotype a in the target population is 1.4[8]. In this way, we have initially realized the objective identification of the contribution proportion of each factor to individual fitness: for male peacocks carrying genotype a, the current tail size determines that they have a 70% probability of successful mating and reproduction, and a 30% probability of being preyed on. The contribution proportion of the two types of events to average individual fitness is 7:3. This proportion is not determined by the subjective preference of researchers, but is a statistical result of the real-life history experiences of 100 genotype a individuals, and only reflects the occurrence frequency of the two types of events.

Our argument is not over yet. The purpose of building this model is to present the specific process of the evolutionary trend of the tail. Now we assume that a male peacock carrying genotype b appears in the target population: compared with genotype a, b has a larger tail. A larger tail increases both the probability of successful mating and the probability of being detected by predators; we assume that the increase in predation probability is smaller than the increase in mating success rate. This means that individuals with genotype b have a higher frequency of reproductive events than 70%, and a lower frequency of death events than 30%, so the average fitness of genotype b is higher than that of genotype a. As long as this assumption always holds—that the increase in predation probability is always smaller than the increase in mating success rate—the evolutionary trend of continuous tail enlargement can be predicted.

Regarding the trend of gradual tail enlargement and continuous improvement of the function of being favored by the opposite sex, this paper needs to supplement three explanations to further improve the interpretation of the model: 1) The trend of tail enlargement cannot continue indefinitely. There are two reasons for the cessation of growth: First, it is constrained by physical laws and gravity. On the premise that other parts of the body remain unchanged, there is a theoretical maximum for tail volume; when the tail grows to the maximum volume that the current body structure can support, it will no longer continue to increase. Second, after a certain round of functional improvement, our previous assumption no longer holds: the increase in predation probability becomes larger than the increase in mating success rate. At this time, continuing to increase tail volume will reduce the average individual fitness, so the growth trend will stop. 2) When genotype b first emerges in the population via mutation, it may be lost by chance due to the small number of individuals encountering excessive death events. But as long as the genotype b mutation can appear repeatedly in the genotype a-dominated population, the spread of b in the population is only a matter of time. 3) We

---

[7] Each of these 70 individuals left 2 offspring, and the fitness of each individual is 2.

[8] Calculation process: "The individual fitness of 70 individuals with genotype a is 2" is recorded as "70 × 2", and "The individual fitness of 30 individuals with genotype a is 0" is recorded " as "30 × 0". Add these two types of events together and divide by the total number of individuals with genotype a, which is 100, to get the result of 1.4.

need to re-emphasize the first step of the logical structure of our model: identification of a function can only be completed after the improvement of this function has lasted for a period of time. Therefore, if we presuppose that the average individual fitness of b is lower than that of a, it means that there is no trend of tail enlargement and functional improvement, because a cannot be both the starting point and the end point of the tail lineage. When we place a sample at the starting point of the lineage (on the premise of sufficient sample size), it already marks the beginning of a new function. If the new function never changes under natural selection after it emerges, this function cannot be identified from the perspective of SE theory.

In summary, by counting the occurrence frequency of various life history events, we can quantitatively integrate the different effects of the same function, or the effects of multiple functions, and calculate the final individual fitness. Introducing the individual life history perspective is a useful supplement to the evolutionary progress identification method in complex situations, but we still need to return to the population-level to complete the final explanation. From the individual perspective, the type of events that a male peacock with genotype a wil will encounter is accidental, and cannot provide effective information about evolutionary trends. But from the population perspective, the 70% mating probability and 30% predation probability are jointly determined by the number and aesthetic preferences of females, and the number and recognition ability of predators. If we repeatedly clone the target population (100 genotype a individuals) 10,000 times, with the number and aesthetic preference of females, and the number and recognition ability of predators remaining unchanged, the frequency of the two events in each cloned lineage will fluctuate around the 7:3 ratio. Therefore, from the population perspective, whether genotype b will gradually replace genotype a is not an accidental process: since the former has a higher average fitness, natural selection exerts real causal efficacy in this process.

5. Conclusion

Through the demonstration of the multiple selective pressure model, this paper proves that Williams' criticism of evolutionary progress is unfounded. The "functional improvement" progress criterion adopted in this paper has many similarities with the "specialization" progress criterion criticized by Williams: functional improvement may also lead to an increase in adaptation in one aspect and a decrease in another, but this does not mean that we are unable to identify the existence of a certain progress trend. As long as we take the common selective history as the premise (which ensures that all individuals are under the same constructive norms), use Howard's SE theory tool, expand the research perspective from the population-level to the individual life history event level, and quantitatively integrate the occurrence frequencies of different types of events, we can respond to Williams' criticism.

Compared with Howard's SE theory, this paper makes supplements in two directions: First, the model in this paper has predictability, which can only be achieved by introducing relevant information on environmental changes outside the SE theory framework; Second, this paper explicitly associates functional improvement with individual fitness, thus avoiding the situation where we cannot understand that "functional improvement may make the individual's situation worse". In the single selective pressure model, functional improvement can be directly associated with individual fitness. However, in the multiple selective pressure model, taking the male peacock's tail we discussed as an example: Continuing to increase the size of the tail will increase the probability of the individual being preyed upon. But as long as the increase in the risk of being preyed upon is less than the fitness gain brought by sexual selection, we can consider that the continued increase in the size of the tail (the function of attracting females is improved) is better for the individual's average adaptation[9]. Of course, to make the multiple selective pressure model meet the predictability requirement, additional information is needed. For example, we need to know the number and aesthetic preferences of females in the population, as well as the number and recognition ability of predators, in order to roughly judge whether the tails of male peacocks will continue to grow or shrink in the current environment in the coming period.

The ideas put forward in this paper are just a preliminary simplified framework, and all the above deductions are just paper calculations based on simple assumptions. I originally hoped to achieve this deduction goal through a computer program, but I was unable to complete it due to my lack of programming skills. Therefore, I hope that the ideas in this paper can inspire scholars who are also interested in the topic of evolutionary progress and have programming skills to further promote this work. In the future, if a more comprehensive model is to be constructed, more factors affecting evolution need to be incorporated into the framework, including more selective pressures, random forces, differences in mutation rates, existing physiological structures, physical environment, and biological environment. Only by comprehensively incorporating various factors can we solve the mystery of how a certain early life form gradually develops into a late-stage life form.

Funding

This paper is the periodic achievement of Project "Research on the Theory Conundrum and Solutions of the Progressive Thinking in Evolution Theory (21BZX004)," supported by the National Social Science Fund of China.

---

[9] It should be noted that when discussing whether an individual is better adapted under multiple selective pressures, it is still necessary to take "common selective history" as a premise. If the average fitness of genotype A consists of functions X and Y, and the average fitness of genotype B consists of functions X and Z. In this case, only the advantages and disadvantages of function X can be compared, and the advantages and disadvantages of the individual fitness of A and B cannot be inferred from those of function X.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Ayala, F. J. (1988). Can "Progress" Be Defined as a Biological Concept? In M. H. Nitecki (Eds.), *Evolutionary Progress* (pp. 75-96). University of Chicago Press.

Bourrat, P., Deaven, K., & Villegas, C. (2024). Evolvability: Filling the Explanatory Gap between Adaptedness and the Long-Term Mathematical Conception of Fitness. *Biology & Philosophy, 39,* Article No. 15. https://doi.org/10.1007/s10539-024-09951-3

Christensen, W. (2012). Natural Sources of Normativity. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences, 43,* 104-112. https://doi.org/10.1016/j.shpsc.2011.05.009

Darwin, C. (2006). *The Origin of Species: A Variorum Text.* University of Pennsylvania Press.

Dawkins, R. (1992). Progress. In E. Keller, & E. Lloyd (Eds.), *Keywords in Evolutionary Biology* (pp. 263-272). Harvard University Press.

Dawkins, R. (1996). *Climbing Mount Improbable.* W. W. Norton & Company.

Dawkins, R. (1997). Human Chauvinism. *Evolution, 51,* 1015-1020. https://doi.org/10.1111/j.1558-5646.1997.tb03686.x

Dawkins, R. (2006). *The Blind Watchmaker.* Penguin Books.

Howard, A. S. (2020). *Evolution, Progress and Biological Function: A Naturalist Perspective.* Ph.D. Thesis, University of Bristol.

Matthen, M., & Ariew, A. (2009). Selection and Causation. *Philosophy of Science, 76,* 201-224. https://doi.org/10.1086/648102

Matthewson, J., & Griffiths, P. E. (2017). Biological Criteria of Disease: Four Ways of Going Wrong. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine, 42,* 447-466. https://doi.org/10.1093/jmp/jhx004

Rosslenbroich, B. (2006). The Notion of Progress in Evolutionary Biology—The Unresolved Problem and an Empirical Suggestion. *Biology & Philosophy, 21,* 41-70. https://doi.org/10.1007/s10539-005-0957-4

Williams, G. C. (1966). *Adaptation and Natural Selection.* Princeton University Press.

Wimsatt, W. C., & Schank, J. C. (1988). Two Constraints on the Evolution of Complex Adaptations and the Means for their Avoidance. In M. H. Nitecki (Eds.), *Evolutionary Progress* (pp. 231-271). University of Chicago Press.

Zou, X. Y., & Li, J. H. (2025). Is There No Objective Standard of Evolutionary Progress? Williams' Refutation of Criterion of Evolutionary Progress and Criticism of Williams. *Philosophical Analysis, 16,* 156-169.