David Haig. *From Darwin to Derrida: Selfish Genes, Social Selves, and the Meanings of Life*. MIT Press 2020. 512 pp. $39.95 USD (Hardcover ISBN 9780262043786).

David Haig, a professor of Organismic and Evolutionary Biology at Harvard University, has written a challenging, wide-ranging, and broadly philosophical reflection on a range of fundamental issues in evolutionary biology. The challenge emerges in part from Haig's dense and idiosyncratic writing style. I suspect it also reflects the fact that *From Darwin to Derrida* did not originate as a single monograph, but as a series of published and unpublished texts that Haig has assembled in the form of a book. In keeping with its multiple origins, Haig's book has a number of overarching objectives. The first is to show that we need concepts such as purpose, meaning, final and formal causality in order to make sense of living things. A strictly mechanistic account of natural phenomena, he thinks, necessarily falls short as explanations. The mechanistic view of life also provides a foil for the second overarching objective of *From Darwin to Derrida.* Contrary to the idea that individual organisms can be understood as machines, with their different parts working in unison toward shared goals, Haig shows that organisms are marked by various kinds of internal conflict.

Haig's book gets underway with a discussion of final causality. As defined initially by Aristotle, the final cause of a thing is that *for which* the thing exists. This idea of causality is relatively uncontroversial in everyday contexts. I can explain to someone what an umbrella is by noting that it helps to keep me dry in the rain. As Haig recounts in chapter 1, however, final causes are effectively taboo within natural science, having been banished in the seventeenth century by the advocates of strictly mechanical science. The reason for their exclusion is straightforward. Within the mechanical worldview, all of the causal factors responsible for a given object or event *precede* that object or event in time. Since final causes necessarily come *after* the objects or events that they ostensibly cause, it is impossible to make sense of them in mechanical terms.

From this perspective, evolutionary biology looks somewhat anachronistic, according to Haig, since one of its central ideas—natural selection—is inextricably bound up with final causality. To explain a gene or a trait as having been *selected* is to explain it in terms of the effects that the gene or the trait brings about. Thus, to explain the prevalence of the FOXP2 gene in humans in terms of its role in language acquisition (and ultimately, in terms of its contribution to reproductive success) is to explain the FOXP2 gene in terms of effects for which FOXP2 is responsible. Albeit familiar, such explanations are backwards in the sense that they place the underlying material factors *after* the development or the behavior that those material factors make possible.

Haig inserts a helpful clarification here, noting that this apparent paradox results from a focus on *types*. In an isolated case, he explains, the priority of the underlying material factors is uncontroversial (xxv). As long as we are thinking of a particular gene in a particular individual, the FOXP2 gene *really does* precede language acquisition. Once we begin thinking in terms of whole populations, however, things are no longer straightforward. From this vantage point, it *is* correct to place language acquisition first in the causal order, because it is language acquisition *as a type* that explains the prevalence and durability of the FOXP2 gene *as a type*. In other words, it is because language acquisition *in general* is conducive to reproductive success that the FOXP2 gene *in general* has endured.

This suggests a thought. If we could dispense with types and attend strictly to individuals, then perhaps final causality would become unnecessary. By extension, if we could track every event and interaction within the world of living things, then the idea of natural selection would likewise become unnecessary. The evolutionary biologist could simply record the sequences of efficient

causes leading to the births and deaths of particular individuals (337). Haig concedes this, noting that final causation and natural selection are ultimately means of *summarizing* complex sequences of efficient causes (253, 284). In other words, they are not part of the metaphysical fabric of the universe; they are explanatory 'tools' that we bring to bear on the world. Nevertheless, according to Haig, they are 'indispensable' tools (268, 284). In their absence, we would not have any meaningful purchase on the world of living things. In order to understand why particular genes have proliferated rather than others, we need to understand the *difference* that those genes make in the world; and this means thinking in terms of final causality (337).

Meaning has a similar status for Haig. While we could in principle represent nature in terms of the playing out of meaningless efficient causes, the result would be illegible. To understand nature is to see it as imbued with meaning at the most fundamental level. Take for example a houseplant growing in the direction of the available sunlight. Plausibly, there is a complete mechanical account of this phenomenon, referring strictly to molecules, photons, chemical reactions, and other physical factors. However, to genuinely grasp the phenomenon, according to Haig, we need to shift into a different conceptual register. We need to regard the plant as having generated an *interpretation* of a certain item of *information* in its environment, with the result of this interpretative act being a publicly accessible *meaning* (285). In effect, we need to regard the plant as having made a decision, such that we can inquire into the purposes or preferences animating its decision.

The analysis is largely the same when we turn to genes. While there is a purely physical story to be told about any given gene, there is a more illuminating story to be told in Haig's semantic idiom. According to this story, a gene is the product of a *choice* made by natural selection. Surveying a set of variants, natural selection *chose* the gene in question on the basis of a salient difference for which the gene was responsible. More than simply a sequence of nucleotides, the gene is therefore a *text* that *represents* a history of natural selection (246, 269).

The notion that genes embody the history of natural selection reflects Haig's commitment to a perennially controversial view within evolutionary biology—the view that genes are the primary object of natural selection. Associated with Richard Dawkins, this view has a number of rivals, including the idea that natural selection operates at multiple 'levels,' including that of genes, cells, individuals, groups, and species (60). Haig's contribution to this debate consists in a pair of refinements to Dawkins' famous image of the 'selfish gene' (74). The first of these is definitional. Looking to capture the fact that natural selection targets *types* of genes, Dawkins himself defines the selfish gene as '*all replicas* of a particular bit of DNA' (61). According to Haig, though, this is not quite right, because it is not generally the case that *all* replicas of a particular gene benefit from the phenotypic effects of that gene. What generally happens, rather, is that a *set* of gene tokens that is related on the basis of common descent will benefit from the gene's phenotypic effects (24). The 'selfish gene,' Haig thinks, should be identified with such sets. He proposes to call the gene as so described the 'strategic gene,' situating it between the 'informational gene' (meaning the abstract type) and the 'material gene' (meaning any particular gene token).

The other refinement that Haig proposes has to do with the idea of the organism. In *The Selfish Gene* (1976), Dawkins famously describes the organism as a 'gigantic lumbering robot,' the suggestion being that the organism is a machine blindly pursuing the objectives of its underlying genes. Albeit vivid, this metaphor is misleading for a crucial reason, according to Haig. Unlike machines, organisms are characterized by conflict between their constituent parts (108). This begins at the level of genes, where, for example, we see conflict between genes of maternal and paternal origin. It extends to conflict between 'genes and memes,' for instance, between my genes' preference to reproduce and my personal preference to avoid reproduction (159). Finally, it includes conflicts

at the level of ideas. As Haig argues, our ethical experience is frequently a matter of negotiating between multiple internal 'voices' (182).

Before concluding, the presence of the name 'Derrida' in the title of a book on evolutionary biology warrants comment. Haig himself is playfully elliptical on this point, declining to explore Derrida's work in any depth, but pointing to a few noteworthy parallels between his own thought and that of the late French philosopher. The most conspicuous of these concerns the public nature of meaning—the proposition that meaning *just is* the physical product of the act of interpretation, with no additional layer of metaphysical depth, or 'intention,' lurking behind it. For my part, I would have liked to see more engagement with another well-known philosopher, namely, Immanuel Kant. The reason is that a number of Haig's views seem to resolve ultimately to Kantian positions. For example, Haig's view that final causality and natural selection are not real forces, but "summaries of complex concatenations of efficient causes" (284), can be readily reframed in Kantian terms as the notion that final causality and natural are *regulative ideas.* Leaving aside the fact that 'From Darwin to Kant' would not have worked as a title for Haig's book, this philosophical kinship strikes me as important and worthy of further exploration.

Evan Clarke, Woods College of Advancing Studies (Boston College)