1 Introduction: From Mechanism to Vitalism and Back

This section provides a brief historical and philosophical framing for the central issues to be discussed in this Element, which have to do with experiment, causality, and explanation in developmental biology before and after its molecular turn. I begin by telling the story of a puzzling discovery in experimental embryology from the late nineteenth century, a phenomenon known as embryonic regulation or regulative epigenesis. Its discoverer, Hans Driesch, later came to believe that it could not be explained mechanistically and was thus led to postulate an immaterial vital force. His arguments for vitalism didn't win many supporters, yet his influence on biological thought was considerable. In any case, the phenomena discovered by Driesch and their possible mechanistic and molecular basis have engaged developmental biologists ever since. Following their fate into contemporary developmental biology provides insight into the workings of experimental science.

1.1 A Puzzling Discovery

Developmental biology today studies a vast range of biological processes that occur in animal and plant embryos as well as in adult organisms, including gamete formation and fertilization, embryonic pattern formation, cell differentiation, organogenesis, limb formation, regeneration, senescence and aging, as well as evolution (e.g., Gilbert and Barresi, 2016).[1] The historical origins of this science, which only emerged as an independent professional discipline in the 1930s and 1940s, are diverse and include experimental as well as anatomical and comparative research traditions (Hopwood, 2009). The experimental tradition began to flourish in the nineteenth century with a research program that was best known at the time as *Entwicklungsmechanik,* which is German for "developmental mechanics," but which was also referred to as "physiological embryology" or "causal embryology." Its advent is usually described as a turn from a natural history-based approach to an experimental science that seeks to identify the causes of embryonic development (Maienschein, 1991). Indeed, the hallmark of *Entwicklungsmechanik* was a thoroughly experimental approach. For example, one of its chief proponents, Wilhelm Roux, punctured single cells in frog embryos with a needle and obtained half frog embryos. He used this result to support his mosaic theory of development according to which embryonic cells divide unevenly such that their daughter cells will rigidly develop into different parts of the body.

[1] Evolutionary developmental biology or "evo-devo" is treated in a different Element by Alan Love.

It was later shown that Roux's results were actually due to the dead cells that remained attached to the manipulated embryos. Normally, rather than maintaining a rigid determination of their fate, embryonic frog cells in fact remain responsive to outside signals that adapt their fate to their location within the embryo, at least during a certain time window. Unlike what Roux sought to prove, frog development is an example of what was referred to as "epigenesis," which could be defined as the generation of new structure due to interactions between different parts of the embryo. By contrast, Roux's mosaic model rather looked like the unfolding of structures that preexisted in some of parts of the embryo, as so-called "preformationist" theories of development held.

Roux's experiments stimulated a lot of research that eventually firmly established the reality of epigenesis, also in frogs. But Roux's work was at least as important in promoting a specific approach: by experimentally manipulating embryos of a model organism used as a stand-in for other organisms (Ankeny and Leonelli, 2020), he sought to learn something about the dispositions of embryonic cells.

While the experimental approach to development was initially associated with mechanistic doctrines, this was soon met with resistance. A keen experimenter with a liking of German philosophy (especially Kant), Hans Driesch was well-known for his experimental work on sea urchins that seemed to be at odds with Roux's earlier frog findings. In fact, Driesch's sea urchin embryos looked like the exact opposite of Roux's frogs: when he separated the embryo at the two- or four-cell stage of development with a hairpin, each one of the cells formed a whole and perfectly happy (although somewhat smaller) sea urchin larva. Driesch argued that this result supported a "regulative epigenesis" rather than Roux's mosaic theory. (It was later found that frog embryos can do the same if the cells are properly detached rather than just punctured; see Maienschein, 1991: 50). But Driesch soon went further. He showed that some parts of the sea urchin embryo retain the potential to form a whole organism up to the 800-cell stage. In addition to his results with sea urchin embryos, Driesch was also experimenting with adult marine organisms that showed remarkable regenerative powers. For example, he showed that sea squirts of the genus *Clavellina* were able to regenerate large parts of the body after surgical removal, and that some parts (such as the branchial syphon) were even able to form a whole new sea squirt.

Driesch suggested that all these systems had something in common, namely, they formed what he called "harmonious equipotential systems." Roughly, these are systems in which each part has the same potential as all the other parts and also the same potential as the whole. For example, all the cells in a four-cell sea urchin embryo have the same potential as the other cells and the same potential

as the whole, namely the potential to form a whole organism. The cell's actual fate (i.e., the structures it ends up actually forming by a series of cell divisions and movements) depends on the cell's location and on the size of the embryo. To explain this, Driesch postulated a factor "*E*" that is responsive to the cell's location and the embryo's size and which tells the cells what they should become. Formally, the fate *S* of any part of a harmonious-equipotential system can be written as a function: $S = f(\boldsymbol{a}, g, E)$, where $\boldsymbol{a}$ is a position vector, *g* a scalar expressing the size of the system, and *E* the factor that Driesch called "entelechy," an expression that he borrowed from Aristotle while admitting that he changed its meaning. What exactly it means in Aristotle is subject to debate among specialists, so it's better not to offer a translation. Let's just note that, etymologically, it is derivable from the Greek word *telos*, which means goal or end. Now the goal-orientedness of developmental processes was clearly an important aspect for Driesch, but his notion of entelechy contains more than this. I shall come back to his concept in Section 1.3.

Driesch then proceeded to provide a proof according to which the factor *E* cannot be a "machine" (his word for causal mechanism, as we shall see in Section 1.3). As a first premise in his proof, Driesch claimed that, in a harmonious-equipotential system, "each one of its parts behaves like the whole" [my translation] (Driesch, 1905: 207). Thus, if *E* were a machine, it would have to be contained as a whole in all of the parts. As a second premise, Driesch claimed that no machine contains itself in all of its parts. From these premises, it follows deductively that *E* is not a machine. Echoing one of Descartes' proofs for the immaterial nature of the soul from its indivisibility, Driesch claimed that *E* is an immaterial principle or an "intensive" as opposed to an "extensive" term (expressions he borrowed from Kant). Thus, Driesch defended *vitalism*, which is a form of *ontological antireductionism* (see Section 1.2).

Even though Driesch's proof is formally valid, the premises are problematic (Weber, 1999). Nonetheless, the impact of Driesch's reflections on the science of developmental biology was considerable. While the solution he offered, the theory of the immaterial entelechy, never had many followers, it has nonetheless influenced developmental biological thought. For example, Driesch's vitalistic ideas have inspired the concept of morphogenetic field (see Section 2.1) and the theory of positional information (see Section 3.1). Furthermore, it was clear that Driesch had identified a major problem for developmental biology and that the solution was not going to be simple. It should be noted that his argument cannot simply be dismissed by pointing out that all the cells contain a copy of the entire genome, for it still needs to be explained why different cells in an embryo activate different parts of this genome at different times in response to their

relative position. We will see what solution was eventually found for this problem in Sections 2 and 3.

This historical episode is very instructive for setting the stage for an engagement with developmental biology from a contemporary philosophy of science perspective. For while vitalism is currently not a live issue (at least in scientific and anglophone philosophical circles), mechanism, reductionism and the extent to which living organisms and their development can be understood like machines very much are.

Since Driesch's days, developmental biologists have learnt a great deal about the processes by which embryonic cells become committed to specific fates within the organism. Long before the molecular turn, experimental embryologists figured out at what stages embryonic cells become determined to form, for example, eyes or neural tissue. By grafting and other experimental techniques, they showed that some cells and tissues interact somehow in this determination process. They also showed in fruit flies that genes are involved. After about 1980, numerous molecules (mostly proteins and the genes that encode them) were identified that mediate such interactions, as well as proteins that control the cells' gene expression patterns such as to commit the embryonic cells to a specific developmental fate. A small number of so-called model organisms were instrumental in these discoveries, including the fruit fly *Drosophila melanogaster*, the African clawed frog *Xenopus laevis*, the zebrafish *Danio rerio*, the mouse *Mus musculus* and the water cress *Arabidopsis thaliana*.

Do these discoveries solve the puzzle raised by Driesch, and if yes, how? Do they constitute a reduction in some sense, or do they vindicate holistic notions such as emergence? What concepts do biologists use when attempting to explain developmental processes? What are the relevant concepts of mechanism and of cause? What is the relationship between the knowledge of classical experimental embryology and that of molecular developmental biology? Did the former provide explanations or was it merely descriptive? Why are molecular accounts deeper than higher-level explanations, if they are? These are the questions to be addressed in this Element. In the rest of this section, I shall outline some philosophical perspectives that could be used for such an undertaking, and eventually choose one.

1.2 Reduction and Emergence

The topic of reductionism has always loomed large over developmental biology, as the debate between developmental mechanics versus vitalism outlined in the Section 1.1 has shown. Driesch opposed a form of what philosophers call *ontological* reductionism or sometimes physicalism, which the early

proponents of developmental mechanics endorsed. This is the idea that living matter consists of the same stuff and is subject to the same physical-chemical laws as ordinary matter. Vitalists, by contrast, believed that the vital forces such as Driesch's entelechy could interfere with physical-chemical laws. While most contemporary philosophers of science (and some philosophers of mind) accept ontological reductionism, many believe in addition that this is the *only* viable form of reductionism. Specifically, they reject two other forms of reductionism, widely known (after Ayala, 1974) as *methodological* and *epistemological* reductionism, respectively. The former, which has rarely been defended by philosophers of science, is the idea that all scientific inquiry should use the same approach (e.g., methods appropriate for the lowest level). The latter claims that there is some logical or epistemic relation (e.g., entailment or explanation) between two confirmed bodies of knowledge, one of which is more fundamental than the other. This is the most extensively debated form of reductionism in the philosophy of biology (Brigandt and Love, 2017).

Whatever arguments philosophers have exchanged on this topic, it is surely tempting to view developmental biology's recent spectacular successes in identifying the molecular basis of development as a triumph of some form of reductionism or another. But which form? By rejecting Driesch's vitalism we are only committed to an ontological reductionism. Can we also claim a form of epistemological reductionism?

First, we would need to identify something to be reduced. A part of the philosophical debate on reduction has been about the question of how some scientific theory is related to its successor (e.g., statistical mechanics and classical thermodynamics or wave optics and Maxwell's electromagnetic theory). Is the older theory reducible to the newer, more fundamental theory? This is called diachronic reduction (Nickles, 1973). Or are we speaking about whether some theory that describes a phenomenon at a higher level is reducible to a lower-level theory without these theories standing in a historical succession, which is called synchronic or inter-level reduction? I shall briefly examine these two kinds of reduction with an eye to our topic. As the knowledge of developmental biology is not normally perceived as consisting of theories (Love 2014), I will use the more neutral term "bodies of knowledge."

First, diachronic reduction. The two bodies of knowledge that might be candidates for a historical succession relation are those of classical experimental embryology and contemporary molecular developmental biology. The former was the research program of *Entwicklungsmechanik* already mentioned in Section 1.1. As we have seen, this was a science that experimented with embryos such as amphibians or marine invertebrates in order to study interactions between different parts of the embryo. However, it has been claimed that

classical experimental embryology never provided any explanations of the phenomena they studied; the experiments at best described the phenomena to be explained in the first place (Rosenberg, 1997; cf. Laubichler and Wagner 2001). On this view, such explanations only became available once the first molecules involved in development had been identified. This is a strong form of reductionism, which claims that all good explanations are reductive. I will refer to this as *explanatory reductionism*. I will argue in Section 2 that classical experimental biology did provide causal-explanatory knowledge and that there is much diachronic continuity between these two sciences; however, it consisted mainly in the experimental practices.

Second, inter-level reduction. I will focus here on *explanatory reduction*, which encompasses not only relations of a theory to a fundamental theory (theory reduction) but also relations between other knowledge items such as individual facts, generalizations of varying scope, fragments of theories or models of mechanisms (Brigandt and Love, 2017). A central question is the extent in which the properties of complex systems are explainable in terms of the properties of their parts and their organization and interactions. It is crucial to include the organization and interactions in the characterization of explanatory reduction to make this a viable notion. In addition, it is sometimes claimed that explanatory reduction means to appeal to parts that can be studied "in isolation." Kaiser (2015) has argued that this requirement is only viable if understood as "studied in a context other than *in situ*." For example, the standard explanation of the propagation of action potentials (nerve signals) by ion channels located in neuronal membranes is reductive not because these ion channels can be studied in complete isolation (they can't work without a membrane), but because they can be studied in vitro; for example, in small patches of membrane attached to the tip of a pipette.

A major obstacle to explanatory reduction that has been claimed is the existence of so-called *emergent properties* (i.e., such properties of complex systems that are *in principle* unexplainable from the properties of the parts and their interactions). A possible source of this kind of emergence is top-down causation (i.e., an influence of the whole over its parts). A standard way of arguing for top-down causation appeals to situations where what some part of an organism does depends on the activities of the organism as a whole (Dupré, 2021: 3–13). For instance, the movements of the heart valves are influenced by the whole body's physiological state. In addition, it is argued that the heart valves would rapidly decay if it wasn't for the vital activities of the rest of the organism. Thus, it appears that to understand living organisms it's not enough to look *down* to its parts; we always need to look *up* to the whole and beyond, the organism's connections to the world that surrounds it.

It is beyond the scope of this Element to do full justice to such arguments; I only wish to point out here that they typically presuppose that, in attempts at explanatory reduction, what we are seeking are explanations of biological phenomena that are *complete*. When Dupré argues that, for example, his capacity to walk upstairs is not reducible to the capacities of any of his parts because "both the capacities of the parts and their very existence as the kinds of parts they are depend on the whole organism" (2021: 11), he assumes that a successful reduction would have to include everything that is causally relevant to his capacity to walk upstairs, which includes the whole organism and a lot of environmental conditions. But note that this does not preclude a *partial* reductive explanation that correctly identifies *some* lower-level causes or constituent capacities, perhaps even all the salient ones given our explanatory goals. Completeness in the sense of causal sufficiency is not a realistic goal for reductive explanations. According to Kaiser (2015), such explanations only need to appeal to a lower level and satisfy the isolation condition mentioned above. Thus, even if Dupré's metaphysics is correct, suitably understood reductive explanations are still possible.

Of course, that reductive explanations exist doesn't mean that *all* good explanations are reductive. For example, while the standard account of action potentials in neuroscience is reductive (Weber, 2005), most evolutionary explanations are not (Sober, 1999). We can accept some reductive explanations or explanatory reductions without committing to explanatory reduction*ism*. As we shall see, in developmental biology we find explanations of the reductive and of the non-reductive type. Classical experimental embryology provided non-reductive explanations, while the explanations of molecular developmental biology are typically reductive (but incomplete).

As we shall see in Sections 2 and 3, it is indeed the case that developmental biology succeeds by identifying only a small fraction of all the causally relevant factors that are present in a developing organism and by ignoring a vast range of other factors. Explanations in biology, whether they are reductive or not, often involve *abstractions*. This means that they leave out a lot of causal detail and focus on just some select causes that are deemed pivotal with respect to the goals of inquiry. Large parts of the organism and its environment are backgrounded. Furthermore, scientific explanations often represent causal relations in an *idealized* way, that is, by radically simplifying the way in which causes operate (see the example of morphogens in Section 3). Such simplification is not a defect. First, according to some philosophers, it can provide *understanding* (Potochnik 2017). Second, it has tremendous *heuristic* value for research (Wimsatt, 2007: Chapter 6). Once it is understood that biology doesn't aim at complete explanations, I suggest, notions such as top-down causality and

emergence are revealed as being irrelevant to scientific practice. Of course, it is still possible that some phenomena resist even partial reductive explanations. The phenomena discovered by Driesch (see Section 1.1) are a potential candidate, and I shall examine them in Section 3.4.

It will not be possible to give the topic of reduction a full treatment here. Nonetheless, I will point to some potential difficulties for explanatory reduction (see Section 2.5), even though I think there clearly are such explanations, however partial, in developmental biology.

1.3 Mechanism

As we have seen in Section 1.1, a central issue raised by Driesch and *Entwicklungsmechanik* has to do with mechanism or the doctrine that all biological phenomena can be explained mechanistically. But what exactly should we take this to mean? To start with, it is important to distinguish between (1) mechanism as the doctrine according to which life can be understood mechanistically, which should really be called "mechanicism," (2), the idea of a mechanism as a machine-like structure and (3) the notion of causal mechanism (Nicholson 2012). While all machine-like structures are causal mechanisms, the converse may not hold; for example, molecular diffusion is a causal mechanism, but it is not machine-like (Levy 2014). As we have seen in Section 1.1, Driesch's proofs primarily purported to show that no machine-like structure that is composed of different parts can be responsible for the phenomena of regulative epigenesis; but I will show now that his concept of machine was so broad that we must read him as being opposed to any kind of causal mechanism that is consistent with the laws of physics and chemistry.

Driesch's concept of a machine was that of an "extensive manifold" in contrast to entelechy, which he viewed as an "intensive manifold." "Extensive" here can be read in the Cartesian sense of spatially extended. Driesch characterizes extensive manifolds as a form of "causality that is based on spatial configurations" (Driesch, 1928: 142) and as a "physical-chemical structure" or a "tectonic," which contains "numerous physical and chemical substances and forces in a typical order" (Driesch, 1905: 206 [my translations]). It is not entirely clear what Driesch meant by "physical-chemical structure" and by a "tectonic," but the most natural reading would be that he meant any kind of causal mechanism that posits different physical or chemical causes in some spatiotemporal arrangement that act in accordance with physical-chemical laws, not just such mechanisms that resemble a human-made machine. Indeed, Driesch's characterization sounds a bit like contemporary

accounts of what causal mechanisms are.[2] That he wanted to exclude any kind of causal mechanism is also implied by the fact he thought that the only feasible alternative is an immaterial force.

But what is a causal mechanism? And what is a mechanistic explanation? Even though these questions are related, the second one may be easier to answer. We can distinguish between at least two senses of "mechanistic explanation": first, the explanation of an activity of a system that breaks this activity down into the activities of components and shows how the interactions between these components produce the system's activities. For example, the activity of neural cell membranes of transmitting action potentials is broken down into the opening and closing of selective ion channels located in the cell membrane. Such mechanistic explanations involve some system (here: a neural cell membrane) with an activity (here: transmitting action potentials) and a set of components (here: selective sodium and potassium ion channels) with their own activities (selective ion transport, voltage-dependent opening and closing) that together produce the activity of the system. Some philosophers of science known as "New Mechanists" think that scientific explanations essentially describe such mechanisms.[3]

According to an alternative and perhaps also more common conception, a mechanistic explanation is simply a causal explanation that identifies some *mediating causal variables* for a cause–effect relation. Simply put, a mechanistic explanation in this sense is a causal explanation that shows how some cause–effect relation is mediated by causal variables that lie causally in between the cause and the effect. For example, when it is shown that smoking causes lung cancer due to certain carcinogens damaging the DNA of lung cells, this amounts to a mechanistic explanation in the second sense. Smoking causes the release of carcinogens into the airways and their uptake by lung cells, which causes DNA damage in the cells' nucleus, which destroys some of the cells' systems that control their division, which causes uncontrolled cell division (i.e., cancer). Of course, such mechanistic explanations may involve not only linear causal chains but also more complex causal networks including feedback and dynamics. This alternative conception (e.g., Baetu, 2019) is distinguished from the New Mechanism approach by its non-verticality, that is, by its not referring to distinct levels of organization. Both kinds of mechanistic explanation may be found in developmental biology (Baedke, 2020).

[2] For example, "Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions" (Machamer et al., 2000: 3).

[3] Craver (2007) is still the most elaborate articulation and defense of this view.

The two notions of mechanism also have implications for experimental methodology. For according to New Mechanism, the components of a mechanism and their activities do not *cause* the phenomenon to be explained; they rather *constitute* it. To test for such constitutive relations, Craver (2007) proposed the criterion of mutual manipulability or MM. Roughly, MM is the idea that some component (e.g., a molecule) belongs to a mechanism for some phenomenon if (i) an experimental "bottom-up" intervention on that component changes the phenomenon and (ii) a "top-down" intervention on the phenomenon brings about a change in the component. Clause (ii) is meant to exclude factors that affect the operation of a mechanism but are causally too remote to belong to the mechanism (such as the effect of blood sugar levels on cognitive activities in the brain). MM has been criticized especially for the idea of top-down interventions, which may not be possible in principle because a system and its parts cannot be manipulated independently (Baumgartner and Casini, 2017). In response, New Mechanists (Craver et al., 2021) have recently revised the mutual manipulability condition in a way that may move their account closer to a level-independent mediating-variable account of mechanism. Thus, the two different conceptions of mechanistic explanation may turn out not to be so different in the end.

Other critics of New Mechanism have presented various examples of scientific explanations that do not appear to be mechanistic, for example, natural selection explanations in evolutionary biology (Skipper and Millstein, 2005) or systems-biological explanations (see the essays by Mekios and Gross in Braillard and Malaterre, 2015). Especially explanations that use dynamical equations are thought to represent a completely different kind of explanation (Stepp et al., 2011). However, this may just be too narrow an understanding of mechanistic explanation (Kaplan and Bechtel, 2011). Silberstein and Chemero (2013) argue that there are neurological systems that exhibit such a high degree of interaction that they cannot be decomposed and localized into separately operating parts. To the extent that mechanistic explanation requires such delocalization and decomposition, such systems are not mechanistically explainable.

Other critics have attempted to show that there exists an important class of biological explanations that cite *pathways*, which differ from mechanisms in several respects (Ross, 2020). One major difference is that pathways track the flow of some specific entity (e.g., a metabolite) through a series of steps without paying attention to much of the other causal factors that are necessary for these steps to occur. According to such critics, it is not illuminating to assimilate concepts such as the pathway concept to the mechanistic framework because by doing so we lose sight of the diversity of explanations that exist in biology.