**Design Principles as Minimal Models**

**Abstract:** In this essay I suggest that we view design principles in systems biology as minimal models, for a design principle usually exhibits universal behaviors that are common to a whole range of heterogeneous (living and nonliving) systems with different underlying mechanisms. A well-known design principle in systems biology, *integral feedback control*, is discussed, showing that it satisfies all the conditions for a model to be a minimal model. This approach has significant philosophical implications: it not only accounts for how design principles explain, but also helps clarify one dispute over design principles, e.g., whether design principles provide mechanistic explanations or a distinct kind of explanations called *design explanations*.

**1. Introduction**

Systems biology recently has attracted a great deal of attention from philosophers of science (O’Malley & Dupré 2005; Boogerd et al. 2007; Bechtel & Abrahamsen 2010, 2013; Fagan 2012; Brigandt 2013, 2018; Green 2015, 2017; Matthiessen 2017). Its attraction is partly due to it seeming to provide a new platform for evaluating and discussing many important philosophical questions, such as scientific explanation, mechanisms, mathematical modeling, reductionism and emergence, interdisciplinary collaboration and integration, to list just a few.

In systems biology, generalizable patterns of network circuitry, also called *network motifs* or *design principles*, play a central role in providing explanation, prediction, understanding, manipulation and control. Accordingly, given their pivotal role in providing explanation in biological practice, there is a question of how they can do so. There are competing accounts proposed, for instance, proponents of mechanistic explanation attempt to develop an extended mechanistic explanation framework to accommodate design principles (Bechtel and Abrahamsen 2010, 2013; Brigandt 2013; Levy & Bechtel 2016; Matthiessen 2017). However, others—not necessarily enemies of the mechanistic explanation framework—contend that design principles provide a distinct kind of explanation differing from mechanistic explanation, which they call *design explanation* (Braillard 2010; Brigandt et al. 2018). The former position is advanced, by and large, on the grounds that constructing design principles belongs to “part of a broader practice of constructing and evaluating mechanism schemas” (Matthiessen 2017, 1). By contrast, the advocates of the latter argue that design principles usually provide more general and non-causal explanations, so that they should be subject to different norms of explanation than the mechanistic one.

In this essay I apply a well-developed approach to answering that question. This approach is the minimal model explanation view developed by Robert Batterman and Collin Rice (Batterman 2002a, 2010; Batterman and Rice 2014; Rice 2018, 2019, 2021, 2022). In applying this approach to systems biology, I suggest that a design principle is a minimal model, for it usually exhibits universal behaviors that are common to a wide range of heterogeneous (living and nonliving) systems with different underlying mechanisms.[[1]](#footnote-1) Following this approach, I then suggest that a design principle, as a minimal model, explains by showing that the principle falls into the same universality class as the diverse systems the model represents. Moreover, by showing that the universal behaviors exhibited by a design principle are common to a diverse range of (living and nonliving) systems with differing underlying mechanisms, this approach enables us to see why design principles tend to provide very general explanations. Furthermore, since a minimal model explains not via uncovering a system’s underlying mechanism or tracking a causal process, it belongs to the broad category of non-causal explanation. Situated in this way, this approach enables us to cut across a debate concerning the distinctive characters (i.e., generality and non-causalness) of explanations provided by design principles. Namely, as we shall see below, this approach helps us to clarify that design principles provide neither mechanistic explanations (because they provide non-causal explanations and never focus on mechanistic details) nor some distinct category of explanations called “design explanations” (because generality and non-causalness can be fully accommodated within the minimal model explanation framework without creating any extra category of explanation), but rather minimal model explanations.

The essay is organized as follows. Section 2 introduces what a design principle is and what a minimal model is. Section 3 discusses the reasons why we can treat design principles as minimal models, which brings along an account about how they explain. Section 4 examines the various philosophical implications obtained by viewing design principles as minimal models.

**2. Design Principles and Minimal Models**

*2.1. Design principles in systems biology*

Systems biologists, using mathematical and computational modeling tools, have discovered interesting generalizable patterns or principles, which they call *design principles* or *network motifs*, standing for “Patterns that occur in the real network significantly more often than in randomized networks” (Milo et al. 2002; Shen-Orr et al. 2002; Cf. Alon 2007, 27). These design principles seem to be simple building blocks of much more complicated biological systems, which usually recur over a whole range of organisms and species.

One such design principle, discussed by philosophers of science (e.g., Braillard 2010; Green 2015; Matthiessen 2017), is *integral feedback control* (IFC), which underlies a pattern of behavior called *robust exact adaptation* (REA) shown in bacterial chemotaxis. REA, also known as *robust perfect adaption*, is “the ability of a system to generate an output that returns to a fixed reference level (its ‘set point’) following a persistent change in input stimulus, with no need for tuning of system parameters” (Araujo & Liotta 2018, 2). Bacterial chemotaxis refers to bacteria’s biased movements in their environments upon sensing gradients of specific chemicals, e.g., moving towards them if the environment contains higher concentrations of beneficial chemicals (attractants) or swimming away from them if containing higher concentrations of toxic chemicals (repellents) (Wadhams & Armitage 2004). These movements are accomplished by bacterial flagella, which can rotate either clockwise or counterclockwise. When rotating counterclockwise, the bacterium is driven along one direction, e.g., traveling towards, or away from, higher concentrations of attractants or repellents. When rotating clockwise, on the other hand, the bacterium is driven to tumble about and travel randomly without orientation (Alon 2007, 137).

One striking feature of bacterial chemotaxis is robust exact adaptation (REA): “the tumbling frequency in the presence of attractant returns to the same level as before attractant was added. In other words, the steady-state tumbling frequency is independent of attractant levels” (Alon 2007, 138). The bacterium travels along a direction only when it senses gradients of chemicals; when the chemicals are kept uniform or simply disappear, it starts tumble again with the same tumbling frequency as before. This guarantees that the bacterium’s sensory system always stays active in response to changes of concentrations of attractants or repellents in its environment. Biologists have already worked out the underlying biochemical mechanism of exact adaptation, which involves a number of proteins (Barkai & Leibler 1997; Alon et al. 1999), as shown below in Figure 1.

The mechanism works in the following way. When an input signal (e.g., attractants or repellents) is present in the environment, it is sensed by a special type of proteins termed receptors spanning the bacterium’s membrane, which travel into the cell and thus deliver the input information inside. The receptors then are combined with the kinase CheA (A) and CheW (W) to form complexes. After this, CheA phosphorylates itself and relocates phosphoryl groups (P) to a messenger protein, CheY (Y). Then, the phosphorylated CheY interacts with the motors, which drive the flagella’s rotation, to elicit the bacterium’s tumbling. The phosphorylated CheY can also be dephosphorylated, facilitated by ChZ (Z). Notice that combining attractants with receptors can reduce ChY’s phosphorylation, thus reducing the bacterium’s tumbling. Finally, exact adaption is achieved via two opposite processes: methylation and demethylation of the receptors, processes that involve addition or removal of methyl (m) groups by two enzymes, CheR (R) and CheB (B), respectively. Methylation of the receptors enhances CheY’s phosphorylation, thus increasing tumbling. Since combining attractants with the receptors reduces tumbling while methylation of the receptors increases tumbling, methylation of the receptors must climb up to a level at which the effect of combining with the receptors can be counterbalanced or compensated. As such, a feedback system is formed by which exact adaptation is achieved.

A diagram of a motor cycle

Description automatically generated

Figure 1. Mechanism underlying chemotaxis of *E. coli.* Source: Rao et al. (2004, 240). Reprint with permission.

To further understand REA, Barkai and Leibler (1997) built a two-state mathematical model for the bacterial chemotaxis network, with the receptor complex being either active or inactive. The average activity of the receptor complex is treated as the output while the concentration of attractants or repellents as the input of the network. This two-state model is accompanied by a set of differential equations that depict the interactions amongst the components. Barkai and Leibler showed quantitatively using their model that the chemotactic network involved is insensitive to the precise values of key biochemical parameters such as the reaction rate constants, the enzymatic concentrations and the attractants/repellents concentrations. In other words, the network obtains REA due to the network’s *connectivity* rather than any ‘fine-tuning’ of the biochemical parameters. This finding was later confirmed by Alon et al. (1999) via experiments.

Employing tools from control theory and dynamic systems theory, Yi’s group has mathematically shown that the design principle, i.e., IFC, is implicated in the chemotactic network identified by Barkai and Leibler’s model, and therefore is responsible for REA. The fundamental property of this principle relies on the fact that “[…] the time integral of the system error, the difference between the actual output and the desired steady-state output, is fed back into the system” (Yi et al. 2000, 4650), as shown below in Figure 2.

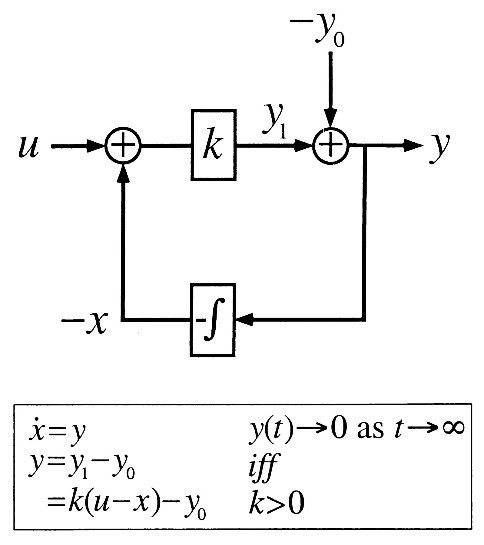


Figure 2. IFC. Source: Yi et al. (2000, 4650). Reprint with permission.

In Figure 2, refers to some input to the system, to a process mediating the input and output, to the system error, namely, the difference between the actual output and the desired steady-state output , and the time integral of . Integral control occurs when is fed back into the system via the feedback loop. This control system guarantees that “the steady-state error approaches zero despite fluctuations in the input or in the system parameters. The only required condition is that the closed loop system is stable” (Yi et al. 2000, 4650). According to Yi et al. (2000), this control structure is not only sufficient but also necessary to explain why robust exact adaption arises. This is because the presence of the control structure entails the presence of REA.

*2.2. Minimal models and minimal model explanation*

A minimal model is merely, metaphorically speaking, a *caricature* of real physical systems. It is usually not a faithful representation of its target physical system(s), for it typically does not pay attention to the actual causal mechanism underlying a particular phenomenon of interest. Rather, it often focuses on some *essential features*, potentially realized by heterogeneous underlying mechanisms, that are responsible for bringing about a common pattern exhibited by a whole range of diverse systems (Batterman 2002a; Batterman and Rice 2014; Rice 2021). So, even though a minimal model is only a caricature, it nevertheless captures the ‘essential features’ of some real physical systems. One such a familiar minimal model, already discussed in the literature (Batterman & Rice 2014; McGivern 2019), is the Lattice Gas Automaton (LGA), a model employed to probe interesting features of fluid flow.[[2]](#footnote-2)

Instead of representing fluids faithfully, the LGA model represents them as oversimplified lattice structures. Each such lattice structure is constructed from a set of point particles that are only allowed to move in one of six directions. Each point particle is attached a momentum vector. The rules of movement for these particles are quite simple (see Figure 3 below): “[…] the particles move in the direction of their arrow to their nearest neighbor site. Then, if the momentum at that site sums to zero, the particles undergo a collision resulting in a jump of 60°” (Batterman & Rice 2014, 358).

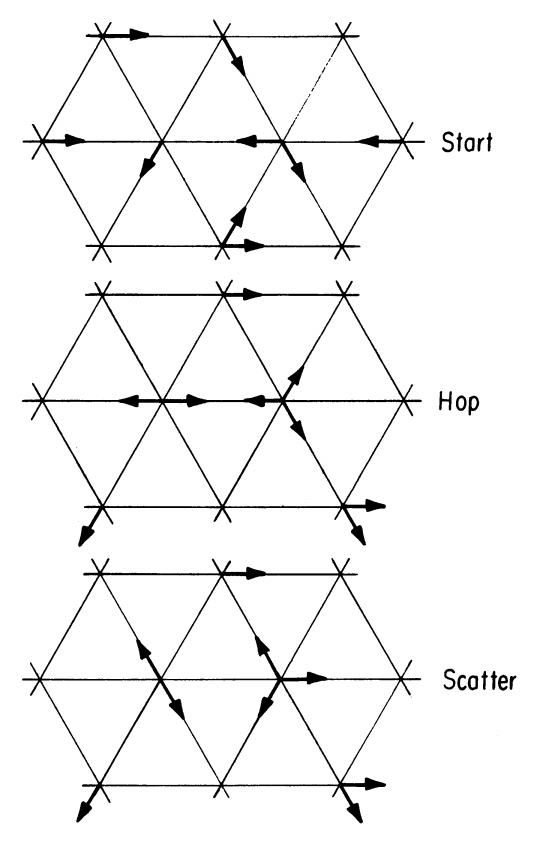


Figure 3. Update rules for movement in lattice structures. Source: Goldenfeld & Kadanoff (1999, 88). Reprint with permission.

At this point one might have already noticed how unrealistic this model is. To begin with, it goes without saying that particles of real fluids do not form such a lattice structure. Second, particles of any real fluid do not move in discrete steps, nor do they move only in six directions. Third, even though particles of real fluids can be associated with momentum vectors, they are not associated in the way stipulated by the model, i.e., each vector only has six directions. However, the most striking thing is that, after iteratively applying the update rules, the model is able to very accurately produce some characteristic macroscopic features of real fluid flow. For example, research has shown that “this model reproduces quite accurately the parabolic profile of momentum density that is characteristic of incompressible laminar flow through a pipe” (Kadanoff et al. 1989; Cf. Batterman & Rice 2014, 358–59; also see Batterman 2002a; Rice 2021). Hence, despite the huge differences between the model and real fluids with respect to their underlying microscopic mechanisms, the model is capable of producing accurate macroscopic results observed in the real fluids.

Given such, the next question is why such a caricature of reality is so useful for investigating real physical systems, or put it differently, why it offers insights to explain the arising of the macroscopic patterns characteristic of real fluids. One short answer to this question is that the model “most economically caricatures the essential physics” of physical systems such as fluid flow (Goldenfeld 2018, 33). However, it is important to note that offering a minimal model is just the first step towards providing a minimal model explanation. To provide a minimal model explanation, the minimal model must be equipped with a process that demonstrates the irrelevance of most of the features of the diverse systems the model intends to represent, as well as delimits a *universality class* in which the model and the systems all fall (Rice 2021, 69–70).[[3]](#footnote-3) And this is typically achieved through a method called the *Renormalization Group*, which “is a strategy for extracting stable macro-behaviors by eliminating irrelevant degrees of freedom from the mathematical description of the system and increasing the coefficients of relevant parameters” (Rice 2021, 70).[[4]](#footnote-4) Essentially, this is a way to demonstrate

“[…] the stability of these behaviors under changes in the lower-scale makeup of the systems. We would be able to show that one could *perturb* the very microstructural details of water to those of gasoline, say, without affecting the common behaviors they display at continuum scales.” (Batterman & Rice 2014, 362)

If, after applying the Renormalization Group, we could demonstrate that the LGA minimal model is a member of the class including diverse real fluids, then we could perturb the microstructural details of the model to that of water (or gasoline) without changing the model’s macroscopic patterns of behaviors. If this is the case, then we obtain an answer to why the model can be employed to represent and explain the macroscopic patterns of behaviors of real fluids such as water and gasoline, namely, this is because *the model falls in the same universality class as many other real fluids* (Batterman and Rice 2014; Rice 2021). This entails that one can use any member of the class to represent and explain the macroscopic patterns of behaviors of the rest members, e.g., one can use water flow as a model to represent and explain why the LGA minimal model exhibits the patterns of behaviors we are interested in. However, though this in theory is possible, using water as a model would be computationally much more inaccessible. That is the very reason we employ the LGA model, a somehow oversimplified or even distorted caricature of reality, to model real fluids but not the other way around.

The Renormalization Group helps us not only delimit the universality class, but also identify the ‘essential features’ of all the members within the universality class. In our case of fluid flow, the essential features are locality, conservation and symmetry.[[5]](#footnote-5) These features correspond to Goldenfeld’s ‘essential physics’, for they “form minimal inclusion criteria for systems in the universality class” (McGivern 2019, 8). So, the heterogeneous systems (including the LGA model) can satisfy these inclusion criteria in very different ways due to their diverse microstructural details, “but as long as they are satisfied, the systems will exhibit the same universal behaviors” (Ibid.).

In summary, as Rice argues, to provide a minimal model explanation, we need not only offer a minimal model but also provide “a detailed demonstration of why most of the features of a class of systems are irrelevant to the explanandum” (2021, 67). And the latter comes down to a demonstration process through which to show that the following four conditions are satisfied:

(1). The behaviors to be explained are universal, i.e., they are common to a whole range of systems.

(2). The behaviors are realized by heterogeneous underlying mechanisms (so the difference-making causes do not matter).[[6]](#footnote-6)

(3). The mechanisms share the same essential features—these are the minimal inclusion criteria for systems in the universality class, but not the causal-making features of the mechanisms.

(4). A minimal model is a member of that universality class (i.e., it has the essential features and exhibits the universal behaviors), and possesses those essential features in some ‘economical’ way. (These conditions are adapted from McGivern (2019, 8))

Notice that in our LGA model, we make appeal to the Renormalization Group to demonstrate that all the above four conditions are satisfied. The Renormalization Group represents a particular way of demonstrating that, and there might be other ways to do the job. However, the more important point is that, no matter how the conditions are satisfied, an explanation is deemed a minimal model explanation insofar as the four conditions are satisfied.

**3. How to Situate Design Principles?**

Given the previous section’s descriptions of design principles and minimal models, an interesting question arises: what is the relationship between design principles and minimal models? In what follows I will suggest that we view a design principle as a minimal model—this is essentially because, as I shall demonstrate below, the explanation provided by a design principle operates exactly in the way that a minimal model explanation operates. Consequently, the explanation provided by a design principle satisfies the four conditions introduced in the last section. Put it in a different way, my claim is this: a design principle provides a minimal model explanation since we can show (through some way discussed below) that the principle and the diverse systems it models all fall into the same universality class. Now, let us consider how an explanation provided by a design principle meets the four conditions.

To begin with, the pattern of behavior, REA, is quite universal across a diverse range of systems. As scientists have noticed, REA “has been widely observed throughout biology at the cellular level (signal transduction, gene regulation, protein interaction networks), in sensory systems, at the whole-organism level in mammals, and during development” (Araujo & Liotta 2018, 2).[[7]](#footnote-7) Moreover, REA has not only been observed in living systems but also in nonliving engineering systems, such as the thermostat-controlled heating system, the vehicle steering system (Sugitani & Tsurumiya 2006), the human-constructed synthetic gene network (Ang et al. 2010), and so on.

Second, the universal behavior is realized by heterogeneous underlying mechanisms. To a first approximation, we might say that all these processes mentioned above in living and nonliving systems are associated with different microscopic mechanisms for they involve different components, different interactions amongst their components, and perhaps distinct organizations. For instance, no one would deny that signal transduction and gene regulation represent two different biochemical mechanisms. However, the devil is in the detail, so we must look at these cases more closely. For this purpose, let me briefly compare two systems, one associated with bacterial chemotaxis discussed in Section 2.1 and one associated with thermostat-controlled heating, one ordinary system we are familiar with.

The mechanism underlying chemotaxis of *E. coli* has been discussed in Section 2.1, which involves components such as receptors, kinases (A and W), a messenger protein (Y), phosphoryl groups (P), enzymes (R and B), etc., and interactions such as binding, phosphorylation, dephosphorylation, methylation, demethylation, etc. By contrast, the bimetallic thermostat, a simple thermostat-controlled heating system, is composed of entirely different components, e.g., it has an outer dial, a temperature sensor, an electrical circuit that connects all the other parts and by which electricity flows through the system, and some other parts. Also, the bimetallic thermostat involves a different set of interactions associated with electricity, which has nothing in common with the set of interactions involved in chemotaxis of *E. coli*. Finally, one can easily tell that chemotaxis of *E. coli* and the bimetallic thermostat have different organizations. So, it is unsurprising that the bimetallic thermostat exhibits a different mechanism: Initially when we connect the system to electricity, one part called ‘the strip bridge’ keeps straight so the system is not disconnected and electricity travels through the system—hence the systems starts to heat; as time goes on, ‘the strip bridge’ gets hotter and one strip (usually the left one) bends much more than the another one (usually the right one)—hence finally ‘the strip bridge’ bends toward the right side to such an extent that it disconnects to the circuit. At this point, the system switches off: it stops heating; then, ‘the strip bridge’ gradually cools down and returns to its initial status—hence the system starts to heat again.

Third, even though these diverse systems are constituted by heterogeneous underlying mechanisms, they share some essential features: homeostasis and adaptation. Homeostasis stands for “the dynamic self-regulation of a system to maintain essential variables within limits necessary for acceptable performance in the presence of external disturbances” (Ang et al. 2010, 723), and in a homeostatic system “the variable of interest is maintained at a constant level by manipulating the internal reaction fluxes to compensate for these perturbations” (Somvanshi et al. 2015, 301). Adaptation here refers to the fact that a system “responds to a persistent perturbation in the environment by a transient response of a sensory or signaling molecule. However, the signaling molecule eventually returns to its pre‐stimulus level despite the presence of the persistent perturbation, thereby, realizing a distinctly new phenotypic behavior” (Ibid.). In essence, homeostasis and adaption are system level properties that keep “the output signal at a desired level in the face of sustained changes in one or more of the input signals” (Ibid., 302). As discussed in Section 2.1, bacterial chemotaxis shows these essential features of homeostasis and adaptation, for “the tumbling frequency in the presence of attractant returns to the same level as before attractant was added. In other words, the steady-state tumbling frequency is independent of attractant levels” (Alon 2007, 138). Similarly, the bimetallic thermostat discussed in this section also shows these essential features, for the thermostat is able to keep the room temperature, the key variable of the system, maintained within a certain acceptable interval of levels despite the persistent perturbations from the environment. Furthermore, many other biochemical processes with heterogeneous underlying mechanisms such as calcium homeostasis (El-Samad et al. 2002), glucose homeostasis (Saltiel & Kahn 2001), and energy homeostasis (Pattaranit & Van Den Berg 2008; Cloutier & Wellstead 2010), all manifest these essential features. Remember that these essential features serve as the minimal inclusion criteria for those systems in the universality class, and satisfying these criteria is sufficient for all those systems within the class to exhibit a particular pattern of behavior.

Finally, consider the fourth condition. To a first approximation, I think we have good reasons to believe that the focal design principle, IFC portrayed in Figure 2, is a member of that universality class, for we have already seen in Section 2 that it has the essential features (homeostasis and adaptation), exhibits the universal behavior of interest (REA), and possesses those essential features in a very ‘economical’ way. This should not be surprising, for the model was just built for the purpose of reproducing and understanding how the pattern of behavior, i.e., REA, would emerge. And to the extent that the model precisely reproduces the pattern of behavior, and to the extent that this is not a coincidence but rather a traceable outcome resulted from possessing the same essential features, saying that IFC does not fall into the universality class would be rather surprising.

There is one point worth noting. One might wonder whether, as in the case of the LGA model, there are ways to demonstrate that perturbing the underlying mechanisms of those systems discussed in this section does not change their essential features and thus does not alter their macroscopic behavior. Recall in the last section we appealed to the Renormalization Group to demonstrate that we can do so for fluids such as water as well as for the LGA model. In other words, we can perturb the microstructural details of, e.g., the LGA model, to that of water (or any other fluids) without changing the model’s macroscopic patterns of behaviors. Notice that these fluids are physical systems that “can be described very precisely in formal terms, together with a rigorous mathematical technique for demonstrating that a group of systems belong to the same universality class” (McGivern 2019, 8). Yet, it remains unclear how the same level of descriptive precision can be realized in living systems, nor is it clear how the living systems can always be described in similar formal terms. What also remains unclear is whether we have similar rigorous mathematical techniques for doing the same job.

Nevertheless, it does not follow that there is nothing we can do.[[8]](#footnote-8) Rather, based on contemporary scientific practice, we can indirectly demonstrate that perturbing the underlying mechanisms of some systems (within the same universality class) does not change their essential features (and thus their macroscopic behaviors). One such a case comes from synthetic biology. In a study aimed to employ IFC to build a perfectly adapting synthetic gene network, a group of scientists first construct a two-promoter genetic regulatory network to explore whether it manifests REA (Ang et al. 2010). After confirming that the two-promoter genetic regulatory network exhibits REA, they go on to perturb the network in various ways, e.g., transforming the two-promoter network into a three-promoter network, and explore whether the new network still manifests REA. Surprisingly, they find that this three-promoter network retains the essential feature (i.e., adaptation) of the two-promoter network and is capable of exhibiting REA. Thus, this is a clear case in which perturbing a two-promoter genetic regulatory network’s underlying mechanism to that of a three-promoter genetic regulatory network does not change their essential features as well as their macroscopic behaviors.

Serban and Green (2021)’s recent work reinforces my point. They point out that there is *mathematical equivalence* between the set of ordinary differential equations (ODEs) provided by Barkai and Leibler (1997) and the IFC characteristic equation provided by Yi et al. (2000). More specifically,

“The starting point of the engineering model (Yi et al. 2000) was Barkai and Leibler’s set of ODEs that was reduced by a series of analytic techniques to a single equation, which corresponds in the control theoretic framework to one of the formulas for computing IFC. […] This equivalence guarantees that there is a formal analogy between the internal organization of engineered and biological systems, which in turn help explain why both systems achieve robust properties in virtue of instantiating specific type of organization described by the IFC design principle” (Serban & Green 2021, 148-149)

Serban and Green also point out that “changing some of the empirical or mechanistic assumptions about the target chemotactic network might still lead to a dynamical description which can be reduced to an IFC formula” (Ibid., 150). Therefore, their analysis has shown that biologists do have some mathematical techniques—perhaps not as powerful as the Renormalization Group—to demonstrate that perturbing (or transforming) the microstructural details of one biological system to that of another does not necessarily alter their macroscopic patterns of behaviors.

Given these considerations, I think there are already good grounds to suggest that design principles in systems biology serve as minimal models (and they provide minimal model explanations accordingly). This is because, as discussed in this section, the design principle in systems biology, e.g., IFC, works in the same way as those in physics, e.g., the LGA model.

**4. Philosophical Implications**

Viewing design principles in systems biology as minimal models has philosophical implications for various philosophical issues surrounding design principles. For instance, a different story about how design principles explain can be told: as a minimal model, a design principle provides a minimal model explanation for the phenomenon (or pattern of behavior) of interest. Moreover, viewing design principles as minimal models helps us resolve one dispute in the literature, one linking to whether design principles provide a distinct kind of explanation called *design explanation*, or provide a sort of mechanistic explanation. In this section I will first discuss whether design principles provide mechanistic explanations or some sort of ‘design explanations’, and then go back to the question of how design principles actually explain.

*4.1.* *Are design principles mechanistic explanations?*

Regarding design principles, there is a dispute in the literature over whether the explanations provided by design principles are mechanistic explanations, or some other kind of explanations. On the one hand, many philosophers think that building design principles, as a type of mathematical modeling, constitutes “part of a broader practice of constructing and evaluating mechanism schemas” (Matthiessen 2017, 1; Levy & Bechtel 2013; Green et al. 2018; Fang 2022).[[9]](#footnote-9) On the other hand, other authors disagree with this, contending that design principles provide a distinct type of explanation which they call *design explanation* (Braillard 2010; Brigandt et al. 2018). This section considers the first horn of the debate (i.e., design principles as mechanistic explanations), and the next section examines the second (i.e., design principles as design explanations).

Let me first target the view that building design principles is part of a broader practice of constructing mechanistic explanations. For two reasons, I think this view is untenable. To begin with, notice the relationship between the explanation provided by IFC and the mechanistic explanation outlined in Figure 1. The mechanistic details of REA associated with bacterial chemotaxis were known a few years before the design principle, i.e., IFC, was proposed—the mechanistic details were worked out by Barkai and Leibler (1997) and the design principle was proposed by Yi et al. (2000). Had the goal been to provide a mechanistic explanation, the explanatory goal would have been satisfied by Barkai and Leibler (1997)’s detailed explanation. However, that was not the case, for systems biologists continued to ask questions such as whether there are general principles governing the discovered mechanism. Note that here they were not searching for another more extended (or still lower-level) mechanism that incorporates (or underlies) the initially discovered mechanism. Instead, they aimed at something more general, something not confined to any specific mechanism or system; they did not ‘look around’ or ‘look down’, but rather ‘look up’.

If the time gap (and the associated focus shift) between the discovery of the mechanism underlying bacterial chemotaxis and the later proposition of the design principle does not convince you, here is something more substantial: a design principle is typically detached from any particular mechanism. To use Batterman (2000)’s terminology, we may say that a common pattern of behavior is multiply realized by heterogeneous underlying mechanisms. Since there is usually only one design principle explanatorily responsible for the arising of a pattern of behavior, and the pattern is multiply realized by different underlying mechanisms, the principle stands in a one-to-many relationship to these underlying mechanisms. Namely, one design principle typically corresponds to many different mechanisms. Recall we discussed in Section 3 that IFC is implicated in a diverse range of systems with heterogeneous underlying mechanisms, e.g., in scaling of morphogen gradients, calcium homeostasis, glucose homeostasis, energy homeostasis, etc. (see Footnote 6). Hence, this is the deeper reason why we cannot simply equate a design-principle-based explanation to any particular mechanistic explanation, which lends support to the claim that building design principles is not part of a broader practice of constructing mechanistic explanations.

There is another reason why we cannot regard building design principles as part of constructing mechanistic explanations: their representational ideals differ, with one aiming *more* at generality while another aiming *more* at realism.[[10]](#footnote-10) Accordingly, we may say that they stand for two different epistemic goals that should be better taken apart and examined separately. One good way to see this is through the lens of scientific models. We might think of creating an explanation as building an explanatory model, so that creating a mechanistic explanation can be thought of as building a mechanistic model.[[11]](#footnote-11) Now the question is whether creating a design principle model (design model for short hereafter) differs from creating a mechanistic model in terms of their representational ideals, so that we have good grounds to claim that they stand for two different kinds of practice. I think the answer is affirmative, for design models typically aim *more* at generality while mechanistic models strive *more* for realism. In other words, in building a design model systems biologists intend to build something general which potentially can be applied to a diverse range of systems (Green 2015), while in building a mechanistic model biologists are primarily cared about building something that is more concrete and realistic. The distinction made by the advocates of the neo-mechanistic philosophy between *mechanism sketches, mechanism schemas* and *complete mechanistic models* perhaps tells us something (Machamer et al. 2000; Darden 2002; Craver 2007). A mechanism sketch is “an incomplete representation of a mechanism that specifies some of the relevant entities, activities, and organizational features but leaves gaps that cannot yet be filled” (Craver & Tabery 2019), which therefore usually “leaves various gaps or employs filler terms for entities and processes whose nature and functioning is unknown” (Weiskopf 2011, 316). In contrast, A mechanism schema is a relatively more complete but “abstract description of a mechanism that can be filled with more specific descriptions of component entities and activities” (Darden 2002, S356). Correspondingly, a complete mechanistic model is one that does not leave gaps or employ filler terms, nor does it abstract away relevant details about the component entities and activities.

However, we must be cautious when making this distinction. First, as many authors have pointed out (e.g., Levy & Bechtel (2013),[[12]](#footnote-12) Green (2015)), building mechanistic explanations is not always a one-way process as championed by some neo-mechanists, namely, not a process that only “involves movement along […] the sketch-schema-mechanism axis” (Craver 2007, 114). Rather, it can sometimes depart from this simplistic one-way path and involve some degree of generality. This point leads to my second point: generality is always a matter of degree (so is realism), so that there is usually a tradeoff between generality and realism. This means that, other things being equal, a more general explanatory model is often less realistic than a less general model, and a more realistic explanatory model is often less general than a less realistic model (Weisberg 2006; Orzack & Sober 1993; Matthewson & Weisberg 2009). In other words, a model cannot maximize the two ideals at the same time: being general means it must sacrifice some degree of realism, and vice versa. Hence, it does not follow that a realistic model can by no means be general, nor does it follow that a general model can never be realistic; Instead, even a mechanistic model has some room for improving generality.[[13]](#footnote-13)

Given this more nuanced picture, I think my position is this: although there might be no clear-cut boundary between generality and realism, we do notice in scientific practice that there is a difference in focus: scientists aim more at generality when building design models and more at realism when building mechanistic models.[[14]](#footnote-14) And the fact that there are boundary cases should not lead one to conclude that there is no difference at all.

Due to these considerations, I think building design principles should not be treated as part of constructing mechanistic explanations. Making this distinction clear is not merely terminologically important but also philosophically rewarding, for now we get a better grip on both what the limits of mechanistic explanations are (and thereby a better understanding of mechanistic explanation) and what the kind of explanation provided by design principles is.

*4.2.* *Not mechanistic explanations, but design explanations?*

By correctly noticing that explanations provided by design principles differ from mechanistic explanations, some authors go even further; they argue that design principles provide a distinct, new category of explanation which they term *design explanation* (Braillard 2010; Brigandt et al. 2018).

One prominent advocate of this view is Braillard (2010), who argues that this is due to two reasons. First, design principles provide “a non-causal kind of explanation that does not show how a function is produced by a mechanism but illustrates how a system’s function determines its structure” (Braillard 2010, 43). Second, design-principle-based explanations tend to be more general than their mechanistic counterparts. Notice that Braillard’s focus is also on IFC, so that his claim is that IFC provides a distinctive design explanation.

Let me consider his reasons more carefully. First, I have no quarrel with his second reason, for I agree that one of the grounds on which to distinguish design-principle-based explanations from mechanistic explanations lies in their representational ideals, i.e., generality vs. realism.[[15]](#footnote-15) For his first reason, however, I only partially agree. The part I agree with is that design principles provide “a non-causal kind of explanation that does not show how a function is produced by a mechanism”. This part dovetails well with, as well as reinforces, my own view. As I said in Section 3, a design principle explains by showing that the principle falls into the same universality class as those target systems the model represents. The other part of his first reason, nevertheless, is what I disagree with. This part proposes that a design principle explains by illustrating “how a system’s function determines its structure”.

As a first step, let me dwell a little longer on what he means by saying “a system’s function determines its structure”. He argues that “while the function determines the structure, it does not cause that structure” (2010, 50–51). For him, this determination relationship is a kind of *non-causal functional constraint*, wherein a function constrains a structure by showing “why a given structure or design is necessary or highly preferable in order to perform a function” (Ibid., 55). For the particular case of IFC, he insists that “it shows a kind of necessity” (Ibid., 55)—even though he admits that “design explanations are not always so strong and should be thought in terms of probability rather than necessity” (Ibid., 55). Notice that his strong interpretation of IFC is inspired by Yi et al.’s claim that “integral control is not only sufficient but also necessary for robust perfect adaptation” (2000, 4652).

His reasoning can be reformulated in the following way: since a given structure (or design) enables a biological function to be performed, the request for the function explains why the structure must be present. Hence, the reasoning runs from the request for the function to the necessary presence of the structure, and it is in this sense Braillard argues that the function determines the structure. However, the trouble with his reasoning is that he takes Yi et al. (2000)’s claim too seriously. Admittedly, it is a common fact in biology that a function constrains the structure by delimiting the space of alternative structures that can realize the given function, and it is ultimately up to the evolutionary processes to select which one (or ones) is present. But constraining is different from determination, with the latter being a matter of necessity while the former not. In other words, we can say that a function constrains the space of possible structures to be selected, but cannot legitimately say that a function strictly determines which structure must be selected. At this point, one might argue that we should look at empirical evidence to adjudicate whether REA determines the presence of IFC, or only weakly constrains the space of alternative structures in which IFC is a member.

Unfortunately, evidence from recent empirical research speaks against Braillard. Ma et al. (2009) show that, after computationally searching all possible three-node enzyme network topologies and examining their adaptation properties over a wide range of kinetic parameters, two major topologies emerge as solutions for robust perfect adaptation: a negative feedback loop with a buffering node (i.e., an IFC-like design principle) and an incoherent feedforward loop with a proportioner node. Shinar and Feinberg (2010), after identifying quite simple yet subtle structural attributes that confer concentration robustness upon any mass-action network, portray a large group of robustness-inducing networks. Moreover, “other researchers commenting on the paper by Yi et al. have cautioned against the idea that RPA necessarily entails the modelled form of integral control” (Briat et al. (2016); Cf. Serban and Green (2021, 156)).

Given this set of empirical evidence, I think at least one thing is certain: the determination relationship, i.e., the defining feature of Braillard’s account, simply does not hold (or at least does not hold as widely as he supposes). If determination is dismissed, then Braillard’s argument loses its steam, for now what is left is at most constraining—for this reason, I think his account can be assimilated to the *constraint-based* account of explanation, which differs in kind from his design explanation. According to Green and Jones (2016), constraint-based explanations encompass a large class of non-causal explanations, and a typical constraint-based explanation is one in which some formal constraints explain by identifying general dependence relationships between functions and structures (see Green and Jones (2016) for details). However, the interesting thing is that, despite that Green and Jones do not say explicitly, minimal model explanation can be viewed as a subtype of the broader class of constraint-based explanations.[[16]](#footnote-16) [[17]](#footnote-17)

Given these considerations, I think Braillard is not justified to claim that design principles provide a distinct kind of explanation. Yet, he is insightful to point out that design-principle-based explanation involves some *modal* force, though not in the way suggested by his account. From the point of view of minimal model explanation, his insight can be interpreted as this: a design principle explains by showing that the principle *must* fall into the same universality class as the diverse systems the principle represents.

*4.3.* *How design principles explain*

If we view design principles through the lens of minimal models, it clarifies how they are able to explain. Roughly speaking, as I have already indicated above, they explain in the same way as those minimal models in physics: first, we offer a minimal model, and second, we demonstrate the irrelevance of most of the features of a class of systems. Put it differently, a design principle explains by showing that *the principle falls into the same universality class as those target systems the model is intended to represent.*

More specifically, the minimal model explanation provided by design principles proceeds in the following steps. First, a universal pattern of behavior (or a phenomenon) is noticed and identified. In the case of IFC, REA in bacterial chemotaxis is the phenomenon or pattern of behavior that intrigues scientists’ interest. Of course, at this stage the pattern of behavior might not be as clearly defined as REA; instead, usually a rough idea of what something interesting is happening is enough to get the research off the ground. Scientists may notice that the pattern of behavior occurs in certain organisms, or in a range of different organisms. At this initial stage of investigation, it does not matter whether the pattern under question occurs only in a few organisms or in a wide range of different organisms. What matters is why and how the pattern of behavior arises. So, the next step is to investigate the underlying mechanism(s) of the pattern observed in those organisms.

In the case of REA in bacterial chemotaxis, the mechanism underlying it is discovered, as shown in Figure 1. This mechanism features a mechanistic explanation, for it involves showing how various components of the system (i.e., the bacteria) are spatiotemporally organized and interact with one another to produce the pattern of behavior in question (Bechtel & Richardson 1993; Machamer et al. 2000; Glennan 2002; Craver 2007). If the explanation were to stop just here, then it would be legitimate to say that it is genuinely a mechanistic explanation. However, as a distinctive ‘system-level’ science searching for generalizable principles (Green 2015), systems biologists usually are not satisfied by discovering any particular mechanism underlying the pattern. Rather, they aim at something more general, viz., generalizable design principles that can be ‘distilled’ from any particular mechanisms. Hence, they go on to ask a further question: going one level higher, could there be any potentially more general principle that governs this or that particular mechanism? This question propels systems biologists to, usually using complicated mathematical modeling techniques, build an abstract and often unrealistic model, and to investigate whether the model could potentially reproduce the same pattern of behavior as that observed in the target systems, as well as the same set of essential features resulted from these systems’ underlying mechanisms. If it turns out that the model is able to reproduce the pattern of behavior and the same set of essential features, then a promising design principle responsible for the pattern of behavior (and for the set of essential features) is on the horizon.

Now, we finally obtain a very general explanatory apparatus that is detached from any specific mechanism. That is a minimal model, and the corresponding explanation provided is a minimal model explanation, rather than a mechanistic explanation. Once this minimal model is established, explaining a given pattern of behavior (or phenomenon) exhibited by some systems of interest essentially amounts to showing that these systems fall into the same universality class as the minimal model. So, the fact that they fall into the same universality class as the minimal model explains why these systems exhibit the pattern of behavior they do, and why they manifest the set of essential features as the minimal model does. This is precisely the way that IFC explains why bacterial chemotaxis exhibits REA (and manifests the set of essential features such as homeostasis and adaptation). In a nutshell, IFC explains by showing that bacterial chemotaxis falls into the universality class where IFC is also a member.

**5. Conclusion**

In this essay I suggested that we view design principles in systems biology as minimal models, so that explanations provided by design principles can be thought of as minimal model explanations. In doing so, we obtained a new approach to understanding how design principles explain: a design principle explains by showing that the principle (as a minimal model) falls into the same universality class as those target systems the model is intended to represent. Advocating such a view has another significant philosophical payoff: it helps resolve one dispute in the literature by showing that explanations provided by design principles are neither canonical mechanistic explanations nor some sort of design explanations. Rather, they are cases of a well-established type of explanations: minimal model explanations.

**References**

Alon, Uri. 2007. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. Chapman and Hall/CRC.

Alon, Uri, Michael G. Surette, Naama Barkai, and Stanislas Leibler. 1999. “Robustness in Bacterial Chemotaxis.” *Nature* 397 (6715): 168.

Ang, Jordan, Sangram Bagh, Brian P. Ingalls, and David R. McMillen. 2010. “Considerations for Using Integral Feedback Control to Construct a Perfectly Adapting Synthetic Gene Network.” *Journal of Theoretical Biology* 266 (4). Elsevier: 723–38.

Araujo, Robyn P., and Lance A. Liotta. 2018. “The Topological Requirements for Robust Perfect Adaptation in Networks of Any Size.” *Nature Communications* 9 (1). Nature Publishing Group: 1–12.

Barkai, Naama, and Stan Leibler. 1997. “Robustness in Simple Biochemical Networks.” *Nature* 387 (6636): 913.

Batterman, Robert W. 2000. “Multiple Realizability and Universality.” *The British Journal for the Philosophy of Science* 51 (1): 115–45.

———. 2010. “On the Explanatory Role of Mathematics in Empirical Science.” *The British Journal for the Philosophy of Science* 61 (1): 1–25. doi:10.1093/bjps/axp018.

———. 2002a. “Asymptotics and the Role of Minimal Models.” *The British Journal for the Philosophy of Science* 53 (1): 21–38.

———. 2002b. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford University Press. https://books.google.com.au/books?hl=en&lr=&id=EiIM5koj\_J0C&oi=fnd&pg=PA3&dq=The+devil+in+the+details&ots=fuQa2Lbqwl&sig=hrxCrd4oxM4IYdSawQcTLCz1kwE.

Batterman, Robert W., and Collin C. Rice. 2014. “Minimal Model Explanations.” *Philosophy of Science* 81 (3): 349–76.

Bechtel, William, and Adele Abrahamsen. 2010. “Dynamic Mechanistic Explanation: Computational Modeling of Circadian Rhythms as an Exemplar for Cognitive Science.” *Studies in History and Philosophy of Science Part A* 41 (3): 321–33.

Bechtel, William, and Adele A. Abrahamsen. 2013. “Thinking Dynamically about Biological Mechanisms: Networks of Coupled Oscillators.” *Foundations of Science* 18 (4): 707–23.

Bechtel, William, and Robert C. Richardson. 1993. *Discovering Complexity*. Princeton: Princeton University Press.

Ben-Zvi, Danny, and Naama Barkai. 2010. “Scaling of Morphogen Gradients by an Expansion-Repression Integral Feedback Control.” *Proceedings of the National Academy of Sciences* 107 (15). National Academy of Sciences: 6924–29. doi:10.1073/pnas.0912734107.

Boogerd, Fred, Frank J. Bruggeman, Jan-Hendrik S. Hofmeyr, and Hans V. Westerhoff. 2007. *Systems Biology: Philosophical Foundations*. Elsevier.

Braillard, Pierre-Alain. 2010. “Systems Biology and the Mechanistic Framework.” *History and Philosophy of the Life Sciences* 32 (1): 43–62.

Briat, Corentin, Ankit Gupta, and Mustafa Khammash. 2016. “Antithetic Integral Feedback Ensures Robust Perfect Adaptation in Noisy Biomolecular Networks.” *Cell Systems* 2 (1). Elsevier: 15–26. doi:10.1016/j.cels.2016.01.004.

Brigandt, Ingo. 2013. “Systems Biology and the Integration of Mechanistic Explanation and Mathematical Explanation.” *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 44 (4): 477–92.

———. 2018. “Explanation of Molecular Processes without Tracking Mechanism Operation.” *Philosophy of Science* 85 (5): 984–97.

Brigandt, Ingo, Sara Green, and Maureen O’Malley. 2018. “Systems Biology and Mechanistic Explanation.” In *The Routledge Handbook of Mechanisms and Mechanical Philosophy*, by Stuart Glennan and Phyllis McKay Illari, 362–74. New York: Routledge.

Cloutier, Mathieu, and Peter Wellstead. 2010. “The Control Systems Structures of Energy Metabolism.” *Journal of The Royal Society Interface* 7 (45). The Royal Society: 651–65.

Craver, Carl. 2007b. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.

Craver, Carl, and James G. Tabery. 2019. “Mechanisms in Science.” In *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/science-mechanisms/#toc.

Darden, Lindley. 2002. “Strategies for Discovering Mechanisms: Schema Instantiation, Modular Subassembly, Forward/Backward Chaining.” *Philosophy of Science* 69 (S3). The University of Chicago Press: S354–65.

El-Samad, H., J. P. Goff, and M. Khammash. 2002. “Calcium Homeostasis and Parturient Hypocalcemia: An Integral Feedback Perspective.” *Journal of Theoretical Biology* 214 (1). Academic Press: 17–29.

Fagan, Melinda Bonnie. 2012. “Waddington Redux: Models and Explanation in Stem Cell and Systems Biology.” *Biology & Philosophy* 27 (2): 179–213.

Fang, Wei. 2022. “Design Principles and Mechanistic Explanation.” *History and Philosophy of the Life Sciences* 44 (4): 55. doi:10.1007/s40656-022-00535-6.

Glennan, Stuart. 2002. “Rethinking Mechanistic Explanation.” *Philosophy of Science* 69 (S3): S342–53.

Goldenfeld, Nigel. 2018. *Lectures on Phase Transitions and the Renormalization Group*. CRC Press.

Goldenfeld, Nigel, and Leo P. Kadanoff. 1999. “Simple Lessons from Complexity.” *Science* 284 (5411). American Association for the Advancement of Science: 87–89.

Green, Sara. 2015. “Revisiting Generality in Biology: Systems Biology and the Quest for Design Principles.” *Biology & Philosophy* 30 (5): 629–52.

———. 2017. *Philosophy of Systems Biology*. Springer.

———. 2022. “Philosophy of Systems and Synthetic Biology.” *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=systems-synthetic-biology.

Green, Sara, and Robert W. Batterman. 2021. “Making Sense of Top-Down Causation: Universality and Functional Equivalence in Physics and Biology.” In *Top-Down Causation and Emergence*, edited by Jan Voosholz and Markus Gabriel, 39–63. Cham: Springer International Publishing. doi:10.1007/978-3-030-71899-2\_2.

Green, Sara, and Nicholaos Jones. 2016. “Constraint-Based Reasoning for Search and Explanation: Strategies for Understanding Variation and Patterns in Biology.” *Dialectica* 70 (3): 343–74. doi:10.1111/1746-8361.12145.

Green, Sara, Maria Şerban, Raphael Scholl, Nicholaos Jones, Ingo Brigandt, and William Bechtel. 2018. “Network Analyses in Systems Biology: New Strategies for Dealing with Biological Complexity.” *Synthese* 195 (4). Springer: 1751–77.

Kadanoff, Leo P., Guy R. McNamara, and Gianluigi Zanetti. 1989. “From Automata to Fluid Flow: Comparisons of Simulation and Theory.” *Physical Review A* 40 (8). APS: 4527.

Levins, Richard. 1966. “The Strategy of Model Building in Population Biology.” *American Scientist* 54 (4): 421–31.

Levy, Arnon, and William Bechtel. 2013. “Abstraction and the Organization of Mechanisms.” *Philosophy of Science* 80 (2): 241–61.

———. 2016. “Towards Mechanism 2.0: Expanding the Scope of Mechanistic Explanation.” http://philsci-archive.pitt.edu/12567/.

Ma, Wenzhe, Ala Trusina, Hana El-Samad, Wendell A. Lim, and Chao Tang. 2009. “Defining Network Topologies That Can Achieve Biochemical Adaptation.” *Cell* 138 (4). Elsevier: 760–73. doi:10.1016/j.cell.2009.06.013.

Machamer, Peter, Lindley Darden, and Carl F. Craver. 2000. “Thinking about Mechanisms.” *Philosophy of Science* 67 (1): 1–25.

Matthewson, John, and Michael Weisberg. 2009. “The Structure of Tradeoffs in Model Building.” *Synthese* 170 (1): 169–90.

Matthiessen, Dana. 2017. “Mechanistic Explanation in Systems Biology: Cellular Networks.” *The British Journal for the Philosophy of Science* 68 (1): 1–25.

McGivern, Patrick. 2019. “Active Materials: Minimal Models of Cognition?” *Adaptive Behavior*. SAGE Publications Sage UK: London, England, 1059712319891742.

Miller, Paul, and Xiao-Jing Wang. 2006. “Inhibitory Control by an Integral Feedback Signal in Prefrontal Cortex: A Model of Discrimination between Sequential Stimuli.” *Proceedings of the National Academy of Sciences* 103 (1). National Acad Sciences: 201–6.

Milo, Ron, Shai Shen-Orr, Shalev Itzkovitz, Nadav Kashtan, Dmitri Chklovskii, and Uri Alon. 2002. “Network Motifs: Simple Building Blocks of Complex Networks.” *Science* 298 (5594): 824–27.

Odenbaugh, Jay. 2003. “Complex Systems, Trade-Offs, and Theoretical Population Biology: Richard Levin’s ‘Strategy of Model Building in Population Biology’ Revisited.” *Philosophy of Science* 70 (5). The University of Chicago Press: 1496–1507.

———. 2006. “The Strategy of ‘The Strategy of Model Building in Population Biology.’” *Biology and Philosophy* 21 (5): 607–21.

O’Malley, Maureen A., and John Dupré. 2005. “Fundamental Issues in Systems Biology.” *BioEssays* 27 (12): 1270–76.

Orzack, Steven Hecht, and Elliott Sober. 1993. “A Critical Assessment of Levins’s The Strategy of Model Building in Population Biology (1966).” *The Quarterly Review of Biology* 68 (4). University of Chicago Press: 533–46.

Patel, Anilkumar K., Sharad Bhartiya, and K. V. Venkatesh. 2014. “Analysis of Osmoadaptation System in Budding Yeast Suggests That Regulated Degradation of Glycerol Synthesis Enzyme Is Key to Near-Perfect Adaptation.” *Systems and Synthetic Biology* 8 (2). Springer: 141–54.

Pattaranit, Ratchada, and Hugo Antonius Van Den Berg. 2008. “Mathematical Models of Energy Homeostasis.” *Journal of The Royal Society Interface* 5 (27). The Royal Society London: 1119–35.

Rao, Christopher V., John R. Kirby, and Adam P. Arkin. 2004. “Design and Diversity in Bacterial Chemotaxis: A Comparative Study in Escherichia Coli and Bacillus Subtilis.” *PLOS Biology* 2 (2). Public Library of Science: e49. doi:10.1371/journal.pbio.0020049.

Rice, Collin. 2018. “Idealized Models, Holistic Distortions, and Universality.” *Synthese* 195 (6): 2795–2819. doi:10.1007/s11229-017-1357-4.

———. 2019. “Models Don’t Decompose That Way: A Holistic View of Idealized Models.” *The British Journal for the Philosophy of Science* 70 (1). The University of Chicago Press: 179–208. doi:10.1093/bjps/axx045.

———. 2021. *Leveraging Distortions: Explanation, Idealization, and Universality in Science*. Cambridge, Massachusetts: The MIT Press.

———. 2022. “Modeling Multiscale Patterns: Active Matter, Minimal Models, and Explanatory Autonomy.” *Synthese* 200 (6): 432. doi:10.1007/s11229-022-03885-7.

Saltiel, Alan R., and C. Ronald Kahn. 2001. “Insulin Signalling and the Regulation of Glucose and Lipid Metabolism.” *Nature* 414 (6865). Nature Publishing Group: 799–806.

Serban, Maria, and Sara Green. 2021. “Biological Robustness: Design, Organization and Mechanisms.” In *Philosophical Perspectives on the Engineering Approach in Biology*, 141–64. Routledge.

Shen-Orr, Shai S., Ron Milo, Shmoolik Mangan, and Uri Alon. 2002. “Network Motifs in the Transcriptional Regulation Network of Escherichia Coli.” *Nature Genetics* 31 (1): 64–68.

Shinar, Guy, and Martin Feinberg. 2010. “Structural Sources of Robustness in Biochemical Reaction Networks.” *Science* 327 (5971). American Association for the Advancement of Science: 1389–91. doi:10.1126/science.1183372.

Somvanshi, Pramod R., Anilkumar K. Patel, Sharad Bhartiya, and K. V. Venkatesh. 2015. “Implementation of Integral Feedback Control in Biological Systems.” *WIREs Systems Biology and Medicine* 7 (5): 301–16. doi:10.1002/wsbm.1307.

Sugitani, Nobuo, and Osamu Tsurumiya. 2006. Vehicle steering system with an integral feedback control, issued February 28, 2006.

Wadhams, George H., and Judith P. Armitage. 2004. “Making Sense of It All: Bacterial Chemotaxis.” *Nature Reviews Molecular Cell Biology* 5 (12): 1024.

Weisberg, Michael. 2006. “Forty Years of ‘The Strategy’: Levins on Model Building and Idealization.” *Biology and Philosophy* 21 (5): 623–45.

Weiskopf, Daniel A. 2011. “Models and Mechanisms in Psychological Explanation.” *Synthese* 183 (3): 313–38.

Yi, Tau-Mu, Yun Huang, Melvin I. Simon, and John Doyle. 2000. “Robust Perfect Adaptation in Bacterial Chemotaxis through Integral Feedback Control.” *Proceedings of the National Academy of Sciences* 97 (9): 4649–53.

1. Green (2022) mentions this idea in passing, but does not develop it. Also, Green and Batterman (2021) has discussed the relationship between universality in physics and design principles, even though their focus is not on explanation. [↑](#footnote-ref-1)
2. For a more comprehensive and accessible discussion of the model, see Goldenfeld & Kadanof (1999). [↑](#footnote-ref-2)
3. I thank one anonymous referee for helping me make clear the distinction between a minimal model and a minimal model explanation. [↑](#footnote-ref-3)
4. For details of this technique, see Goldenfeld (2018) and Batterman (2002b, 2010). [↑](#footnote-ref-4)
5. “Locality: A fluid contains many particles in motion, each of which is influenced only by other particles in its immediate neighborhood. Conservation: The number of particles and the total momentum of the fluid is conserved over time. Symmetry: A fluid is isotropic and rotationally invariant” (Goldenfeld & Kadanoff 1999, 87; Cf. Batterman & Rice 2014, 360). [↑](#footnote-ref-5)
6. This reminds us of multiple realizability, namely, systems with heterogeneous microscopic mechanisms can realize the same macroscopic property, behavior or function. Actually, Batterman argues that “The multiple realizability of the properties of the special sciences […] is best understood as a kind of universality” (2000, 115). [↑](#footnote-ref-6)
7. More specifically, REA has been observed in scaling of morphogen gradients (Ben-Zvi & Barkai 2010), calcium homeostasis (El-Samad et al. 2002), glucose homeostasis (Saltiel & Kahn 2001), energy homeostasis (Pattaranit & Van Den Berg 2008; Cloutier & Wellstead 2010), neuron firing in prefrontal cortex (Miller & Wang 2006; Patel et al. 2014), and son on. [↑](#footnote-ref-7)
8. Note that Batterman and Rice (2014) has already offered a biological case, in which they argue that Fisher’s sex ratio model provides a minimal model explanation. [↑](#footnote-ref-8)
9. Levy and Bechtel (2013) treat building design principles in systems biology as constructing abstract connectivity models (that emphasize the patterns of causal connections) within the mechanistic framework, and Green et al. (2018) view building design principles as a kind of network analysis which supports and extends traditional mechanistic strategies. For both, thus, constructing design principles is still part of a broader practice of constructing mechanistic explanations. [↑](#footnote-ref-9)
10. These representational ideals (plus *prediction precision* not discussed in this essay) were first introduced by Levins (1966), which later received a great deal of attention from philosophers of science (Orzack & Sober 1993; Odenbaugh 2003, 2006; Weisberg 2006; Matthewson & Weisberg 2009). According to these discussions, generality concerns how many actual (or possible) physical systems to which a model applies, whereas realism concerns the degree to which a model faithfully represents the actual physical system(s). [↑](#footnote-ref-10)
11. In fact, many authors in the mechanical philosophy literature explicitly view building mechanistic explanations as building mechanistic models, e.g., Craver (2007b) makes a distinction between *how-possibly, how-plausibly* and *how-actually* (mechanistic) models, and Weiskopf (2011) directly treats mechanistic explanations as mechanistic models, etc. [↑](#footnote-ref-11)
12. Notice that Levy and Bechtel (2013) use the term ‘abstraction’ rather than ‘generality’, and that they denote slightly different things, i.e., generality concerns how many actual or possible physical systems to which a model applies, whereas abstraction concerns the degree to which specific details about parts and connections are left out. Yet, when contrasted with realism, they are generally interchangeable for omitting details, and thus making a model more abstract, usually results in a more general model. [↑](#footnote-ref-12)
13. The real tension is that, even though a mechanistic model can be general in some way, what makes it general are not its causal-mechanistic details, for these realistic causal-mechanistic details are exactly what lead the model to be less general. In other words, adding realistic causal-mechanistic details can improve its realism but cannot at the same time improve its generality. For a relevant discussion, see Rice (2022). I thank one anonymous referee for alerting me to this point. [↑](#footnote-ref-13)
14. Notice that when talking about a design model’s generality (or realism), we are only talking about the general features of the systems (including the model itself) in the universal class, not the general features of any particular mechanisms or causes of those systems. I thank one anonymous referee for letting me clarify this. [↑](#footnote-ref-14)
15. See Section 4.1 above for this distinction. [↑](#footnote-ref-15)
16. From the quotation below, we can easily see how similar their characterization of constraint-based explanation is to the characterization of minimal model explanation described in Section 2.2: a constraint-based explanation involves: “1. A claim that systems within some class C, differing with respect to some range of mechanistic detail, share some organizational or structural property O and some dynamic or functional property D. 2. A demonstration that formal constraints F, applicable to systems within C despite their heterogeneous mechanistic detail, define or limit the range of possible dependence relations between O and D. 3. An inference that all systems in C exhibit the permitted O-D dependency relations, regardless of differences of mechanistic detail among members of C” (Green & Jones 2016, 361). [↑](#footnote-ref-16)
17. Notice that Green’s most recent work (Serban & Green 2021) conflicts with her earlier work (i.e., Green & Jones 2016), for the former suggests a *structural-causal* account of design explanations whereas the latter argues that design principles provide constraint-based explanations which are typical non-causal. For the limitations of space as well as the uncertainty of Green’s position, I will not discuss Serban & Green (2021) in this essay. [↑](#footnote-ref-17)