Book chapter to appear in Voosholz, J. & Gabriel, M. (2021, eds.). *Top-Down Causation and Emergence*. MIT Press.

Making sense of top-down causation:

Universality and functional equivalence in physics and biology

Sara Green[1] & Robert W. Batterman[2]

Abstract

Top-down causation is often taken to be a metaphysically suspicious type of causation that is found in a few complex systems, such as in human mind-body relations. However, as Ellis and others have shown, top-down causation is ubiquitous in physics as well as in biology. Top-down causation occurs whenever specific dynamic behaviors are realized or selected among a broader set of possible lower-level states. Thus understood, the occurrence of dynamic and structural patterns in physical and biological systems presents a problem for reductionist positions. We illustrate with examples of *universality* (a term primarily used in physics) and *functional equivalence classes* (a term primarily used in engineering and biology) how higher-level behaviors can be multiple realized by distinct lower-level systems or states. Multiple realizability in both contexts entails what Ellis calls "causal slack" between levels, or what others understand as *relative explanatory autonomy*. To clarify these notions further, we examine procedures for upscaling in multi-scale modeling. We argue that simple averaging strategies for upscaling only work for simplistic homogenous systems (such as an ideal gas), because of the scale-dependency of characteristic behaviors in multi-scale systems. We suggest that this interpretation has implications for what Ellis calls *mechanical top-down causation,* as it presents a stronger challenge to reductionism than typically assumed.

Keywords: Functional equivalence class; Multiple realizability; Reductionism; Top-down causation; Universality; Constraint

[1] Section for History and Philosophy of Science, Department of Science Education, University of Copenhagen, Niels Bohr Bygningen, Rådmandsgade 64, 2200 Copenhagen, DK. Email: sara.green@ind.ku.dk.

[2] Department of Philosophy, University of Pittsburgh, 1029-H Cathedral of Learning Pittsburgh, PA 15260, USA.

1. Introduction

The problem of top-down causation refers to how and whether changes of higher-level variables can have causal effects on lower-level behaviors (Campbell 1974; Ellis 2005; 2008). Top-down causation remains a contested issue in science and philosophy of science alike (Auletta, Ellis and Jaeger 2008). A common assumption is that if macroscale systems consist of "no more than" physical-chemical components, it should be possible to describe higher-level phenomena bottom-up from more fundamental lower-level descriptions (e.g., Oppenheim and Putnam 1958; Crick and Clark 1994; Bedau 1997). This chapter comments on and adds to important insights from George Ellis' work that challenge this assumption. Ellis has been one of the key figures emphasizing that topics such as reductionism and top-down causation are not only philosophically interesting but have important practical implications for science and medicine. We examine further examples in support of this view, by stressing an even stronger interpretation of what Ellis (2012) calls *mechanical top-down causation*.

An important precondition for the existence of top-down causation is that explanations of phenomena at higher scales or levels are (relatively) autonomous of explanations at lower levels. If higher-level explanations and parameters were fully reducible to or derivable from more "fundamental" ones, appeals to top-down causation would be unnecessary or even misleading. In arguing against bottom-up determination of higher-level properties, Ellis (2008, 2012) appeals to the existence of multiple realizability, illustrated through the existence of *equivalence classes* in different scientific domains. In the following, we therefore examine the connections between multiple realizability, equivalence classes, and top-down causation.

Multiple realizability means that a higher-level state or property is realized by different heterogenous states or properties at a lower level. The term is often introduced in discussions about the ontological or explanatory autonomy of higher-level phenomena and models. For instance, Putnam argues against physical reduction of mental states by highlighting that mental kinds are multiple realized by distinct physical kinds (Putnam 1980). Others have appealed to multiple realizability in discussions about explanatory unification (Fodor 1974; Sober 1999; see also Brigandt and Love 2017). However, for the purpose of the discussion of top-down causation, the most important aspect of multiple realizability is that it supports the explanatory autonomy of more general higher-level models that capture similarity in behaviors of heterogenous systems (see also

Batterman 2000; 2018). Ellis connects the issues of multiple realizability and top-down causation as follows:

> Top-down causation takes place owing to the crucial role of context in determining the outcomes of lower level causation. Higher levels of organization constrain and channel lower level interactions, paradoxically thereby increasing higher level possibilities. A key feature here is *multiple realizability* of higher level functions, and consequent existence of *equivalence classes* of lower level variables as far as higher level actions are concerned. An equivalence class identifies all lower level states that correspond to the same higher level state. (Ellis 2012, p. 128, emphasis added)

We unpack with further examples the claims that i) multiple realizability supports explanatory autonomy of higher-level features, and ii) top-down causation can be interpreted as the effects of higher-level constraining relations that determine outcomes of lower-level causation.

Equivalence classes are also sometimes called *universality classes* in physics. Both concepts highlight how systems with distinct microstructures often display general or universal patterns of behavior. Describing these behaviors does not require reference to microscale details – in fact, generic models and explanations are often identified through procedures that abstract from or selectively leave out irrelevant details (Batterman 2000; Green and Jones 2016). We illustrate the relation between multiple realizability and universality through the example of thermodynamics near critical points in Section 2.1. We then examine what Ellis terms *functional equivalence classes* in biology, exemplified through feedback control (Section 2.2). Functional equivalence can be interpreted as an instance of universality that applies only to engineered and living systems, since equivalent behaviors here are characterized in functional terms such as information, robustness, homeostasis, control, etc. (Ellis 2008; 2012). Functional equivalence classes typically are more context-dependent than classical examples of universality in physics, a point we shall elaborate on further below.

For multiple realizability (or conditional independence) to be possible, Ellis holds, there must be *causal slack* between lower and higher levels (Ellis 2012). The notion of causal slack usefully highlights how the explanatory autonomy of higher levels is justified by empirical demonstrations of *conditional independence* of upper-level behavior on many lower-level details (see also Woodward, forthcoming and this volume). The term implies that the autonomy is relative to certain conditions that hold for a given equivalence class. Just like a sail can be slack within

certain length limits of the sheets, so do relations of conditional independence hold within certain boundaries or parameter spaces. Hence, the autonomy of macro-level explanations is not absolute, but relative (see also Batterman 2018). Critics may argue that this threatens the explanatory autonomy of higher-level explanations, and hence the possibility of top-down causation. Instead, we believe that the notion of relative autonomy avoids many problems afflicting strong accounts of top-down causation (such as difficulties in understanding how different levels are connected), and therefore also offers a better description of how scientists develop multi-scale models in practice.

The emphasis on *relative* explanatory autonomy parallels Ellis' distinction between his account of top-down causation and a "stronger" interpretation in which top-down causation is described as efficient causation operating across levels. A strong account has been criticized for giving rise to the problem of causal overdetermination or to mysterious cause-effect relations, which would violate the lower-level laws of physics (Kim 1998; 2000). However, if we interpret top-down effects as higher-level *constraining-relations* on the possible lower-level states of a given system, top-down causation becomes a matter of understanding how higher levels define the *boundary conditions* of lower-level dynamics (see e.g., Ellis 2008; 2012; 2016; Green 2018; Moreno and Mossio 2015; Mossio and Moreno 2010). Constraints are here understood as physical conditions that limit the degree of freedom of a dynamic process, thus enabling only selected system states (Christiansen 2000; Hooker 2013; Juarrero 1998; 1999). Constraints are typically regarded as being at a higher spatial scale than the entities and operations of the constrained processes. In biology, constraints are often further defined according to functional levels in a hierarchically organized system (Pattee 1971; 1973; Salthe 1985; Wimsatt 2007). For instance, the shape and size of blood vessels enable efficient circulation by limiting the degrees of freedom of liquid motion, or blood flow. The interpretation of top-down causation as the ability of upper-level variables to set the context for lower-level ones may be seen as a "weaker" form of top-down causation (Emmeche et al. 2000). Yet, it allows for an understanding of how constraints productively can channel system states that are not possible to reach for an unconstrained system.

The productive aspect of constraints can be illustrated by how an open respiratory system would not be able to provide sufficient gas exchange for a large organism. Similarly, the constraints provided by a sail on wind flow enable a sailboat to move. When one increases the drag by trimming the sail, one does not (effectively) change the operating cause (the wind). Rather, one modifies the structural constraints that channel a pressure difference across the windward and leeward side of the sail. Constraints thus have causal power by delimiting the space of possibilities

for lower-level causes. Without appeals to top-down constraints, we would not be able to explain why specific states are realized among multiple possible lower-level states and through such selections give rise to emergent properties.

Ellis distinguishes between several types of top-down causation. The most basic form of top-down causation is *mechanical* top-down causation, exemplified by how the rigid boundaries of a gas container constrains the degree of freedom of the lower-scale movement of gas molecules (see also Christiansen 2000). Elsewhere, Ellis also refers to this form as *algorithmic* top-down causation, because top-down causation can be understood mathematically as the effects of boundary conditions on the solution to equations describing lower-level dynamics (Ellis 2012). A similar account is defended by the systems biologist Denis Noble in the context of multi-scale cardiac modeling (Noble 2012; 2017; see also Emmeche et al. 2000; Green 2018).

Whereas physical or chemical systems can exhibit *mechanical* or *algorithmic* top-down causation, living systems display multiple additional types such as non-adaptive information control, adaptive selection, and intelligent top-down causation (Ellis 2008; 2012). These are often considered as stronger forms of top-down causation, because biological functions must be understood through goals of whole organisms and species, which again depends on higher-level features such as the environmental and evolutionary background (Ellis and Kopel 2017). These "stronger" types of top-down causation specific to biological systems are not the focus of our chapter.[3] Rather, our aim is to show that the ideal of "bottom-up reductionism" (Gross and Green 2017) can also be challenged with examples of *mechanical top-down causation*. Our chapter responds to a common view or concern that in the contexts of physics and chemistry, "it is not always clear whether traditional reductionist point of view is actually overcome, since these [high-level variables] can again be understood as a complicated effect of more elementary processes" (Auletta, Ellis and Jaeger 2008, p. 1162). We draw on examples from both physics and biology to argue that higher-level variables used to model many multi-scale systems are, in fact, not reducible in the sense often assumed. Hence, we think that mechanical top-down causation presents a stronger challenge to reductionism than typically assumed.

The analysis is structured as follows. After clarifying the concept of universality in physics and relating it to Ellis' notion of equivalence class, we further elaborate on Ellis' suggestion that

[3] Readers interested in these types of downward causation, as well as debates on the metaphysical implications of downward causation, may find Paoletti and Orilia's (2017) comprehensive anthology on downward causation interesting. For examples of downward causation in ecology, see also (Allen and Star 1982; Ulanowicz 1986; 1997).

feedback control and network motifs are examples of functional equivalence classes (Section 2). We then compare top-down causation based on information control to mechanical or algorithmic top-down causation (Section 3). Drawing on examples of multi-scale modeling in physics and biology, we argue that high level variables used to describe heterogeneous systems cannot be derived from micro-scale details through coarse-graining. We end the chapter with reflections on the practical implications of the autonomy of scales and top-down causation for science and medicine (Sections 4 and 5).

2. Universality and functional equivalence

One way to expresses the challenge to the reductionist is to ask: "how can systems that are heterogenous at the microscale exhibit the same pattern of behavior at the macro-scale?" (Batterman 2018, p. 861). We can answer this question by demonstrating that details of the heterogenous realizers are to a large extent explanatorily irrelevant, and thus that we are justified in idealizing (and thereby effacing) many lower-scale details when our explanatory target is at higher scales. Similarly, Ellis highlights that if an upper-level behavior is multiply realized, we do not have to appeal to micro-level details but can explain higher-level patterns through the generic characteristics of the equivalence class. For Ellis, the features explaining the characteristics of equivalence classes are higher-level constraining relations that *channel similar outputs in heterogenous systems*. Such effects can be interpreted as instances of top-down causation (Ellis 2012; Section 3).

2.1. Universality and multiply realizability in physics

A paradigmatic example of universality is found in thermodynamic behavior near critical points. Various fluids consisting of different chemical elements will have different critical temperatures and pressures. That is to say, they will undergo so-called continuous phase transitions at pressures and temperatures that depend upon the micro details of the molecules. However, the behavior of many different fluids at the critical point of phase transitions are identical and can be characterized by the same critical exponents. During a phase transition, e.g., when water boils in a pot, the densities of the liquid water and the vapor (steam) will differ. And, in fact, there will be regions of

liquid and regions of vapor that coexist in the pot. If one plots the difference in the densities of the liquid and the vapor one notices that as the temperature approaches a critical value, this difference exhibits power law scaling behavior. Remarkably, when one plots this behavior for a certain class of fluids in dimensionless (reduced) coordinates $(\frac{\rho}{\rho_c}, \frac{T}{T_c})$, one can show that they all exhibit the same scaling behavior. See Figure 1 for a dramatic display of this behavior. Thus, there is a universal property (shared behavior at macroscales by systems distinct at microscales) for systems near their respective critical points. More remarkable still is the fact that systems like magnets near criticality also display very similar coexistence curves. (For magnets the order parameter is the net magnetization[4] but the scaling exponent is identical.) Part of the reason for this universal behavior is the fact, noted explicitly by Kadanoff (1971) that the closer the system is to criticality, the less the macroscopic/continuum properties depend on the dynamical details of the system.

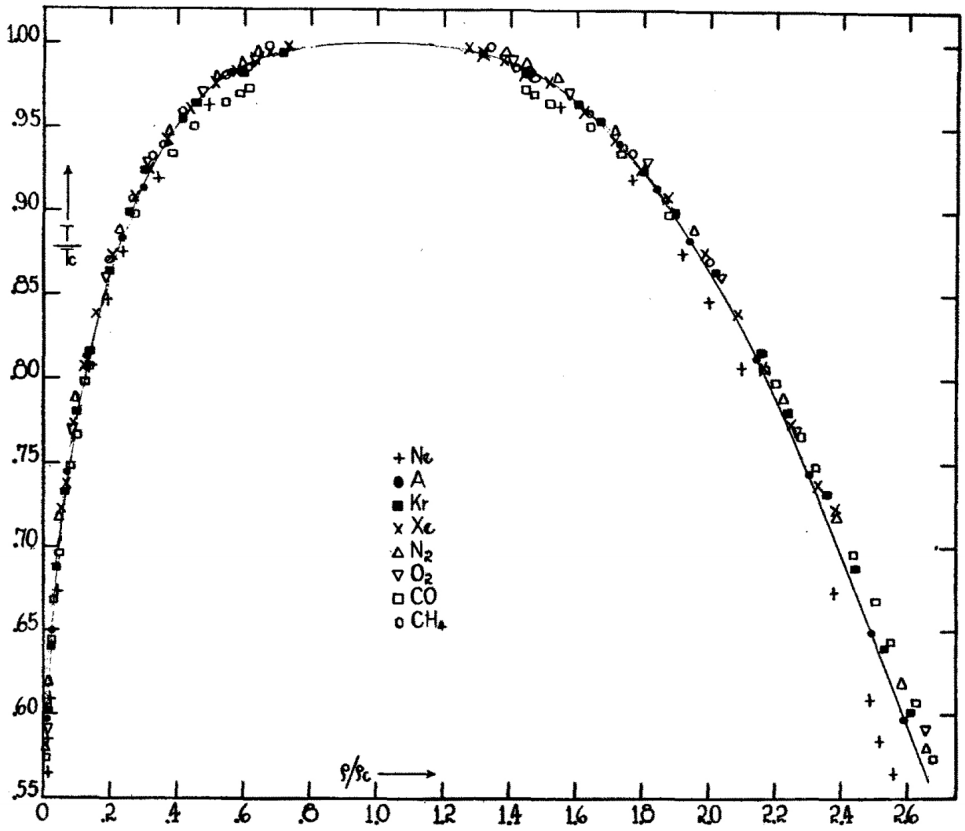

Figure 1. Vapor-liquid coexistence curve for various fluids. The figure shows the difference in densities at temperatures in reduced coordinates. At $(\frac{\rho}{\rho_c}, \frac{T}{T_c})$ = (1,1) liquid phase (left) and vapor phase (right) have the same density, thus their

[4] The net magnetization can also be understood as a difference in densities. The densities of up-spins vs. down-spins. This difference vanishes at the critical temperature, as the high temperature randomizes the directions of the spins.

difference vanishes. Reprinted from Guggenheim, Edward A. 1945. "The principle of corresponding states." *The Journal of Chemical Physics* 13, 7, 253-261, with the permission of AIP Publishing

If we wish to understand why many different systems exhibit such similar behaviors, we are unlikely to find the answer in reductionist appeals to fundamental force laws for each chemical component. Rather, the answer is given by the so-called renormalization group explanation of the universality of critical phenomena (henceforth RG explanation). This explanation takes advantage of the fact that near criticality systems exhibit self-similar, fractal like, behavior. Thus, one can introduce a transformation on the space of Hamiltonians that throws away details via a kind of coarse-graining. Repeated application of this transformation eliminates details that genuinely distinguish the different systems from one another. The hope is to find a fixed point of the transformation from which one can determine the value for the scaling exponent. All systems that evolve under this transformation to the fixed point for the universality class of systems exhibit the same macroscopic scaling behavior. The RG explanation thus extracts structural features that stabilize macroscopic phenomena irrespective of changes in or perturbations of microscopic details. This example from physics not only gives a concrete interpretation of multiple realizability, but also allows for a better understanding of why some very simple models (Ising models, for example) can be used in quite varied contexts to help explain the behaviors of real systems (fluids and magnets) (Batterman 2000; 2018).

Phase transitions illustrate the emergence of new macroscopic features through the different characteristics of liquid water, steam, and crystalline ice. They involve discontinuous alterations in higher-level behaviors through the influences of higher-level variables (such as temperature and pressure) on lower-level interactions (Ellis 2016, 139). In Ellis' view, environmental variables triggering phase transitions should be interpreted as coarse-grained higher-level variables, because temperature and pressure cannot be attributed to isolated molecules. Like order parameters, these point to collective properties arising in a constrained system, such as a gas container (see also Christiansen 2000). Yet, pressure and temperature are interpreted as coarse-grained because they can be identified through averaging of lower-level details (see Section 3). Accordingly, phase transitions of this type represent the most basic (or weak) form of top-down causation (Ellis 2016, pp. 224-225). We shall return to this point in Section 3, after examining some examples from biology for comparison.

2.2. Functional equivalence and information control in biology

Ellis uses the term *functional equivalence class* when referring to models in the life sciences that identify correspondence (or equivalence) of lower-level states or systems with respect to the corresponding higher-level variables and behavior (Ellis 2008; 2012). Examples of these are recurring network motifs and feedback control, which exemplify multiple realizability in biology. We further unpack examples from systems biology that support this view.

A hallmark of living systems is *homeostasis*, i.e., the ability of organisms to maintain a relatively stable internal environment despite external perturbations (Bernard 1927/1957; Cannon 1929). The robustness of functional steady states in organisms is typically explained with reference to feedback control, a concept imported to biology from engineering in the 1920s and later formalized by the mathematician Norbert Wiener (Wiener 1948). Wiener's book was ground-breaking in suggesting that the same mathematical models can be used to describe feedback control in very different oscillatory systems, from electrical circuits to metabolic regulation in different organisms. The hope for generic systems principles was highlighted also in Rachevsky's mathematical biology and Bertalanffy's general system theory (Green and Wolkenhauer 2013). In recent years, systems biology has further strengthened and elaborated on this view by using generic models from control theory and graph theory to describe so-called *organizing* or *design principles* in living systems (Alon 2007; Green 2015).

An example of a design principle is integral feedback control, which is used to explain robust perfect adaptation in bacterial chemotaxis. The example is described in further detail in other publications (Green and Jones 2016; Serban and Green, 2020), and we shall here focus only on why integral feedback control can be seen as an instance of multiple realizability. An important question in biology is how various functions are maintained despite environmental perturbations. For instance, biologists are interested in understanding how motile bacteria can detect changes in the concentrations of nutrients or toxins in their environment and optimize their movements according to these. Remarkably, chemotactic bacteria have receptor systems that can detect and respond to concentration changes in their environments with the same precision before and after stimulus. In engineering terms, the receptor system is said to display *robust perfect adaptation* (RPA), i.e., the system will return to its pre-stimulus value and regain sensitivity over a large range of parameter values (Alon et al. 1999). Achieving this kind of robustness is a hard problem in engineering. Engineers are often interested in designing systems that asymptotically track a fixed steady-state

value, so as to maintain system function despite noisy input signals (or changes in initial conditions).

In engineering, robust adaptation to pre-stimulus steady-state values can be achieved through a design principle called integral feedback control (IFC). IFC refers to a quantifiable feedback relation in which the difference between the desired output (steady-state activity) and the actual output is fed back to the system as the integral of the system error. Strikingly, the mathematical description of bacterial receptor systems has been found to be equivalent to formal models of integral feedback control in engineering. Systems biologist John Doyle and colleagues derived the principle through a mathematical analysis in which they reduced a mechanistically detailed model of the receptor system to a generic description involving only relations sufficient for the higher-level property of robust perfect adaptation (Yi et al. 2000).[5] As a result, IFC was proposed as a *design principle* that generically constrains functional behaviors and enables robust perfect adaptation, regardless of the causal details of the heterogenous systems realizing this capacity. In Ellis' terms, we can say that there is sufficient "causal slack" between higher-level behaviors and lower-level realizers to allow for the *same principle* to apply to systems as different as bacteria and engineered thermostats. We shall comment further on this example below.

Similar examples of multiple realizability in biology are so-called network motifs in gene regulatory networks (Ellis 2012; Fang 2020). Network motifs are small sub-circuits of regulatory connections that have been found to be frequent in biological regulatory networks and have been hypothesized to display characteristic generic functions (Alon 2007).[6] For instance, a so-called coherent feedforward loop, cFFL, has been shown to implement a sign-sensitive delay of outputs in response to input signals. This function was first demonstrated mathematically through a simple Boolean input function (see Figure 2), and the hypothesized function was subsequently confirmed experimentally in living bacteria (Mangan et al. 2003). In biological systems, it can function as a persistence detector that can filter out noisy input signals, such as brief fluctuations in the concentration of nutrients available in a bacteria's environment. This ensures that protein synthesis of metabolizing enzymes is only activated when the activating signal (availability of sugars) is persistent (for further details, see Alon 2007).

[5] This aspect is analyzed in further detail in (Green and Jones 2016).

[6] Investigations of the stabilizing aspects of global constraints in networks have been explored much earlier, e.g., by Stuart Kauffman's demonstrations of how the structure of Boolean networks constrains the possible network states (Kauffman 1969; 1993).

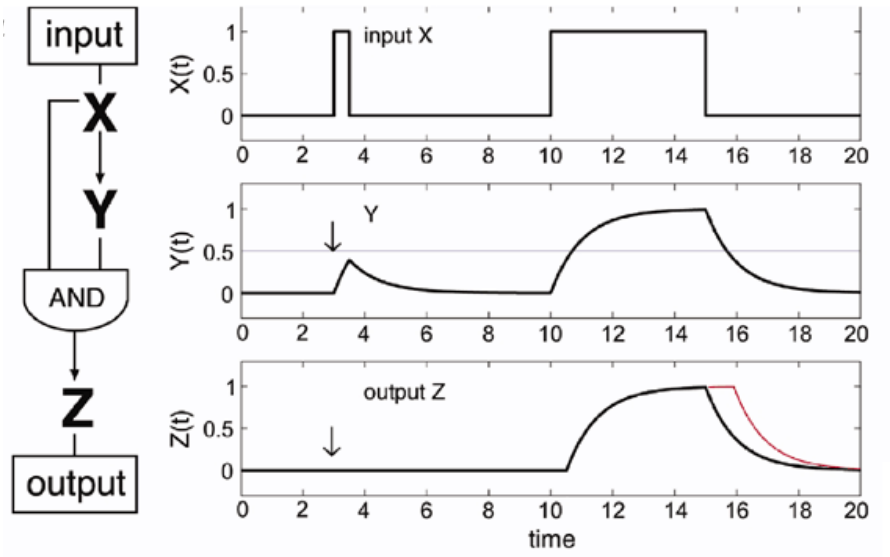

Figure 2. The design principle of the coherent FFL describes how the rate of a protein, Z, is a function of the activity of two transcription factors, X and Y, which are activated by two independent signal inputs (S_x and S_y). Detection of signal persistence arises as a result of the time difference between direct and indirect activation routes (X to Y, and X to Z via Y), and the requirement for persistent stimuli from both X and S_y before Y is activated (compare graphs for activation of X, Y, and Z). Hence, short pulses do not lead to activation of Z, but persistent activation of X and Y will. Reprinted by permission from Macmillan Publishers Ltd: Shen-Orr et al. Network motifs in the transcriptional regulation network of *Escherichia coli*, *Nature Genetics*, 31, 64-68, Copyright (2002).

Systems biologists more generally use the term *design principles* to highlight that generic network structures instantiate general dynamic patterns that i) are independent of specific realizations in different causal systems, and ii) serve functional or goal-oriented roles in engineering and biology (Green 2015). In other words, the characteristics of functional equivalence classes are explained with reference to *how network structures constrain dynamic outputs* to enable generic types of functions such as sustained oscillations, noise filtering, robust perfect adaptation, signal amplification, bi-stable switching, etc. (Doyle and Stelling 2006; Tyson et al. 2003; Tyson and Novák 2010).

The quest for design principles highlights the hope in systems biology that any network circuit with a specific structure, regardless of the specific details of its causal constituents, will belong to a more general functional equivalence class. If so, this would allow gene regulatory functions to be predicted and explained independently of detailed knowledge about the lower-level genetic and molecular details of specific systems. Generic functions of network motifs have been demonstrated in various contexts (Alon 2007). In the neighboring field of synthetic biology, multiple realizability through network motifs is exploited as a design heuristic for the synthesis of synthetic circuits with pre-defined functions (Koskinen 2017; 2019). Similarly, systems biologists

have recently explored the global properties of gene regulatory networks, following up on Kauffman's (1969; 1993) insight that even complex networks often converge to a limited set of stable states. Using the framework of dynamical systems theory, systems biologists have demonstrated that many different molecular mechanisms can lead to the same attractor states (representing biological functions or states of cell differentiation), thus moving the focus from the details of causal pathways to system trajectories (Huang 2011). This approach can potentially explain biological robustness (via top-down causation), because it can show how stable functional states are largely independent of the specific states (or initial conditions) of specific network nodes.

Functional equivalence classes in biology are, however, more contested than the classical examples of universality in physics (Auletta, Ellis and Jaeger 2008). Analyses of global network topologies and network motifs have shown that functions of genetic circuits are dependent on the contexts of the gene regulatory network, the environment, and organisms as a whole (Huang 2011; DiFrisco and Jaeger 2019). It has therefore been debated to what extent the structure of network motifs determines gene regulatory functions (Isalan et al. 2008; Jiménez et al. 2017). Similarly, systems biologists have debated whether biological systems exhibiting robust perfect adaptation necessarily realize integral feedback control (cf., Yi et al. 2000; Briat et al. 2014). Importantly, however, conditional independence is compatible with some degree of context-dependence within defined boundaries. We further clarify this below.

The discovered complexity has sparked an interest in understanding how wider system contexts can influence the characteristic functioning of specific network motifs. This can for instance be done through simulations where parameter spaces for the strengths of inputs and weighting of regulatory connections are varied (Tyson and Novák 2010). Hence, an aim here is to explore the conditions under which generic functions can be inferred from structural network types. Systems biologists have also used computer simulations to explore the possibility spaces for network structures that can realize specific functions of interest. For instance, they explore how many network topologies fall within a functional equivalence class and which structural features characterize the class. As an example, Ma et al. (2009) conducted a computational search for networks capable of performing robust perfect adaptation and investigated their regulatory wiring patterns. From a starting point of 16,038 possible network topologies, they found that only 395 were capable of performing RPA, and that they all fell into two generic structural classes (one is a negative feedback loop with a buffer node, the other is an incoherent feedforward loop). Interestingly, all known biological examples of RPA are instantiations of the negative feedback

control type, as described by Yi et al. (2000). The example thus highlights how structural constraints may realize functions that are multiply realized in distinct systems and thus "unify the organization of diverse circuits across all organisms" (Ma et al. 2009, 760).[7]

The complexity and diversity of biological systems presents a major challenge to provide an analysis in this context similar to the RG explanation in physics. But despite the limitations for universal laws in biology, generic models have proven useful for explaining why characteristic dynamic patterns arise in causally different systems. An ideal in systems biology is to shift the focus from inherent properties of specific genes or proteins to how those are interconnected through stabilizing regulatory structures that give rise to similar higher-level behaviors. In physics and chemistry as well as biology, an important part of scientific analysis is thus to determine "how many values of hidden variables can underlie the same higher-level description" (Ellis 2016, p. 120). In the following, we examine further how the causal slack of "hidden variables" supports a relative explanatory autonomy of higher-level models.

2.3. Causal slack and explanatory autonomy

The notion of hidden variables can be understood as a domain of lower-level causal details that would not change the output of a higher-level function, e.g., because the system trajectory would converge to the same fixed point or attractor in an abstract phase space.[8] As Ellis highlights, for a given equivalence class "it does not matter which particular lower level state occurs, as long as the corresponding higher-level variables are in the desired range" (Ellis 2008, p. 74). This has important implications for the way natural phenomena are represented and explained.

A notable feature of systems biology textbooks, compared to those of molecular biology, is that molecular details are almost absent in the figures and diagrams (cf., Alon 2007; Lodish et al. 2008). The use of highly abstract illustrations not only highlights how the functional descriptions are (relatively) independent of molecular details, but also that functional equivalence classes are identified through procedures of what Ellis (2012) calls *information hiding*. Akin to how we arrive at explanations for universal behaviors in physics by the use of RG explanations, systems biologists

[7] This necessarily involves abstraction from lower-level details. In the words of Ma et al.: "Here, instead of focusing on one specific signaling system that shows adaptation, we ask a more general question: what are all network topologies that are capable of robust adaptation?" They further state that the aim to "construct a unified function-topology mapping […] may otherwise be obscured by the details of any specific pathway and organism". (Ma et al. 2009)

[8] It goes (almost) without saying that this notion of "hidden variables" is not quantum mechanical.

must necessarily abstract from molecular details in order to make the generic patterns of network organization visible (Levy and Bechtel 2013; Green 2015).

Some would argue that information hiding strips the abstract network models of their explanatory power, and that these are explanatory only in so far as details are added back to explain the workings of specific systems. For instance, Matthiessen (2016) argues that design principles such as IFC are not explanatory if they do not allow us to distinguish between different species of bacteria or between organisms and thermostats. But while fine-graining by adding details would serve the explanatory purpose of how specific systems work (what we can call a type-I question), it would not address the (type-II) question of what these systems have in common or *why the same abstract model or principle applies to causally diverse systems* (Batterman 2002, p. 23). The reductionist would hence have to explain how higher-level descriptions can be relatively autonomous from changes of lower-scale details (see also Wimsatt 2007).

To respond to the type-II question, it is important to highlight two implications of what Ellis calls causal slack, and others understand through *scale separation* of multi-level systems. Scale separation refers to the manifestation of different dominant behaviors at different length scales, which accordingly must be described through different types of mathematical models (Batterman 2012). If, for instance, one is interested in modeling the bending properties of a material such as steel, it would not be useful for take the lowest (atomic) scale as the starting point. Instead, one typically starts with variables at the mesoscale and upscales to higher-level variables (such as elastic material parameters) that are used in continuum models. In fact, a detailed microscale models would often be irrelevant if the aim is to describe upper-scale continuum behaviors, as upper-level behaviors are literally invisible at the lowest scales. Scale separation explains why meso- and macroscale models often work well, despite ignoring or even misrepresenting many lower-level details. Indeed, mesoscale parameters are often dependent on some microscopic details, and these are accounted for through so-called homogenization strategies (see also Section 3). One can interpret the use of such strategies as a way to determine the degree and kind of causal slack between different spatial scales.

Top-down causation implies that higher-level features are not just relatively autonomous from lower-level description but also influence the latter through constraining relations. An objection to this view may be that since higher-level features primarily select among possible lower-level states, higher levels are not really autonomous after all. If lower levels define the possibilities, and if emergent features are always realized through materials at lower levels, what

does autonomy really consist in?[9] Our response is as follows. First, causal slack implies that higher levels are not fully controlled or determined by lower-level features but also by the structuring of the system. Second, the requirement of higher-level boundary conditions to constrain the space of causal possibilities implies that some lower-level states cannot be accounted for without the boundary conditions imposed by higher-level structures. For instance, Noble (2012) highlights how a phenomenon such as the heart rhythm is not possible without top-down causation, because the constraining relations of a membrane (understood mathematically as cell voltage) are required to produce oscillations of ionic currents.

Noble's example illustrates how constraints, as indicated in the introduction, must be understood in an *enabling* as well as *limiting* sense, as they allow for stable system behaviors or robust functions that would be impossible to reach in an unconstrained system (Pattee 1971; 1973). Another example illustrating this point is how our rigid skeleton enables upright movement on land by delimiting the possible directions of muscle contraction (Hooker 2013). It would be misleading to say, in this context, that the top-down constraints of the bone primarily provide a (non-explanatory) background for the lower-level states or operations of molecular muscle cells. Rather, as in the case of cardiac cells, certain emergent properties become possible only when constraining relations are applied on lower-level states. The causal power of top-down constraints can also be clarified through "negative" examples, where the constraints are removed and functions as a result become impossible to obtain. Breaking a bone has immediate effects on the causal possibilities for muscle performance, just like a ripped sail immediately changes the speed and control of a sailboat. Similarly, Noble mathematically demonstrates how removal of the top-down constraints of cell voltage causes oscillations to cease in a simulation of the heart rhythm (Noble 2012). Generally, many biological processes would not be possible without inter-level constraining relations (see DiFrisco and Jaeger 2020 for further examples).

[9] This concern was for instance raised after a talk by Ellis entitled "On the Nature of Causality in Complex Systems", at the conference The Causal Universe, Krakow, Poland, May 17-18, 2012. Available online: https://www.youtube.com/watch?v=nEhTkF3eG8Q. In the following we further elaborate on a possible response to this question.

3. Top-down causation and high-level variables

For Ellis, a crucial feature of top-down causation is how coherent higher-level actions emerge from top-down constraints on lower-level dynamics (Ellis 2012, p. 128). Such constraints are often mathematically described as boundary conditions that delimit the set of lower-level variables (e.g., the set of initial conditions as inputs to the dynamics) within which a given function can be realized. Top-down causation should thus not be understood as causal effects that are completely autonomous from lower-level dynamics. The organizing or design principles defined at higher scales nevertheless have causal effects that constrain the distributions or values of lower-level constituents or states, thus enabling emergent behaviors of lower levels (cf. Emmeche et al. 2000). As mentioned in the introduction, Ellis (2008; 2012) distinguishes between several types of downward causation. Following up on sections 2.1. and 2.2, we focus only on two, namely top-down causation by information control and mechanical or algorithmic top-down causation.

Feedback control loops in biology are instances of what Ellis calls *top-down causation by information control* (Auletta, Ellis and Jaeger 2008; Ellis 2012). The IFC-principles and the proposed generic functions of network motifs highlight a basic condition for equivalence classes, namely that many different input situations or causal "realizers" would give rise to equivalent operational outcomes, as long as basic structural requirements are obeyed. Importantly, a feedback loop is here interpreted as a structural constraint that delimits the space of possibilities for lower-level dynamics. Hence, feedback control consists in *information selection*. For instance, the structure of a feedforward loop motif determines whether a genetic circuit should respond only to persistent input signals (coherent feedforward loop) or immediately to any nutrient detected (incoherent feedforward loop). As illustrated in Figure 2, coherent feedforward loops control transcription by introducing a time difference between direct and indirect transcription activation routes. Similarly, incoherent feedforward loops can minimize the response to disturbance by simultaneously activating the transcription of an output product and an intermediary transcription factor that inhibits output protein production. Hence, structures such as network motifs constrain the dynamic possibilities of gene regulation at a lower scale in a hierarchy (see also Bechtel 2017). Ellis defines top-down causation by information control as follows:

> Top-down causation by information control occurs thanks to the connection between equivalence classes and information control […] In this case, the feedback control circuits

> produce reliable responses to higher level information (Ellis 2006; 2008), allowing equivalences classes of lower level operations that give the same higher level response for a certain goal. (Auletta, Ellis and Jaeger, pp. 1169-1170).

Different lower-level operations are here considered as *controlled by information from above* in the sense that the control circuits are considered as higher-level entities in two senses. First, the functions are implemented by networks that cannot be reduced to the operation of lower-level entities in isolation (Ellis 2008). Second, functional goals are higher-level concepts referring to the properties of a whole system (an organism, metabolic system, or circuit of interacting processes or entities). The second feature requires some clarification. Ellis (2008) argues that: "[t]he goals in biological systems are "intrinsic higher-level properties of the system considered, and determine the outcome (unlike the usual physical case, where the initial state plus boundary conditions determine the outcome). […] The initial state of the system is then irrelevant to its final outcome, provided the system parameters are not exceeded" (Ellis 2008, p. 74).

Ellis stresses that top-down relations in biology include considerations of part-whole relations, which do not necessarily translate to the physical context. While we agree with this characterization, we do not view downward causation primarily as a compositional relation between parts and wholes (see also Woodward, forthcoming). Rather, we view top-down causation as relations between higher-level and lower-level variables. Thus understood, we find it potentially misleading to consider the relative independence from initial conditions as a prime feature that separates physical and biological systems. As we have seen, universality in physics is characterized through insensitivity to lower scale details, and both biological and physical systems can be described through equivalence classes within certain boundaries of system parameters. Accordingly, we suggest, that "bottom-up reductionism" can also be challenged by examples of mechanical or algorithmic top-down causation in both physics and biology. Although typically considered a "weaker form" of top-down causation, compared to top-down causation by information control, cases of multi-scale modeling in both domains highlight the limitations of a bottom-up approach (Batterman and Green 2020). Mechanical or algorithmic top-down causation refers to a ubiquitous form of top-down causation that occurs whenever "high-level variables have causal power over lower level dynamics through system structuring or boundary conditions, so that the outcome depends on these higher level variables" (Ellis 2012, p. 128). With Ellis, we believe that this form of top-down causation is much more common than typically recognized.

3.1. Revisiting mechanical top-down causation

As mentioned, Ellis views mechanical top-down causation as a phenomenon occurring also in physical and chemical systems. This was exemplified in the way changes in higher-level variables (such as pressure or temperature) can lead to changes in lower level interactions in gases and fluids, enabling new properties such as gas ignition or phase transitions (Section 2.1).[10] In these examples, Ellis seems to assume the correspondence between lower- and higher-level variables is given by a relatively simple relation between these, i.e., that higher-level variables can be derived from coarse-graining of lower-level ones (Auletta, Ellis and Jaeger 2008). Ellis broadly defines high level variables as follows:

> A high level variable is a quantity that characterizes the state of the system in terms of a description using high level concepts and language – it cannot be stated in terms of low level variables. Use of such variables involves information hiding, for they are the relevant variables for the higher level description, e.g., the pressure, temperature and density of a gas, without including unnecessary lower level details (such as molecular positions and velocities). (Ellis 2008, p. 70).

Ellis further distinguishes between *coarse-grained* higher-level variables and *irreducible* higher-level variables. The pressure of an ideal gas exemplifies the former, whereas feedback loops or the tertiary structure of protein folding illustrate the latter (see also Brigandt and Love 2017). Coarse-grained variables can be obtained by averaging over a set of lower level variables, and they are therefore in principle possible to derive from lower-level details (although this is often not done for practical reasons). This has implications for the strength of mechanical top-down causation because Ellis views all the high-level variables concerned with this type of top-down causation as coarse-grained. He further writes: "The resulting high level relations are then an inevitable consequence of the low level interaction, given both the high level context and the low level dynamics (based in physics)" (Ellis 2008, p. 72).

As a friendly amendment to this view, we suggest that the scope of mechanical or algorithmic top-down causation be expanded to include cases that go beyond instances of simply coarse-grained higher-level variables. While we agree that the high level variables are coarse-

[10] Other examples from physics are discussed in (Bishop 2012; Christiansen 2000; see also Ellis 2018).

grained in the case of homogeneous systems such as ideal gases, there are many multi-scale systems in physics where meso- and macroscale parameters cannot be obtained via simple averaging procedures (Batterman and Green, 2020). This presents a further challenge to the reductionist point of view.

Indeed, in the case of an ideal gas, we can assume that the system is homogenous, and we can therefore upscale to the higher-level thermodynamic behavior by relatively simple averaging over micro-scale details (such as molecular spatial and velocity distributions). However, whenever multi-scale systems are heterogenous and display more complicated limit behaviors, this approach would fail. For instance, if the task is to develop a multi-scale model of the bending behaviors and relative strength of heterogenous materials such as a steel beam or vertebrate bone, complex homogenizations strategies are typically adopted to account for mesoscale structures that are not observable at lower scales (Batterman and Green 2020). In both contexts, simple averaging over lower-scale variables would not enable scientists to predict macroscale material properties.

For multiscale systems that are heterogeneous (e.g., composites of materials with different conductivities or elastic behaviors), the aim of upscaling is to find effective (continuum scale) parameters (like Young's modulus) that code for microstructural details of the composites.[11] Typically this involves the examination of a representative volume element (RVE) that reflects the nature of heterogeneities at scales (mesoscales) where those structures are deemed to be important. One introduces correlation functions to characterize (primarily) the geometric and topological aspects of the mixture in the RVE (Torquato 2002). The mathematics that enables one to find the effective parameters that characterize the behavior of the composite at the continuum scale is called "homogenization theory." (Batterman and Green 2020 discusses some aspects of this in the context of materials science and biology. See also Batterman, forthcoming.) Note that the relative autonomy of the homogenized system (at the continuum scale) from the atomic lower scale details reflects a kind of emergence. This sense of emergence is weaker than that associated with higher-level variables characterizing the human mind or social phenomena. Yet, the higher-level parameters cannot be reduced to or derived from lower-level details.

Consider Young's modulus, an example of a higher-level parameter of central importance in materials science and biophysics. Young's modulus parameterizes the stiffness of a material and is identified as the slope (or coefficient of proportionality) of a stress-strain curve of a given material. Stiffness is understood as the resistance of a material to deformation in response to applied force

[11] Note that "microstructure" here refers to structures far above the atomic and far below the continuum.

and cannot be understood or defined at atomic scales. More generally, material parameters describe mechanical properties of a larger continuum of structure that cannot be measured or defined at the level of individual "parts". As stressed by developmental biologist (and biophysicist) Lance Davidson and colleagues, material parameters of relevance for modeling of the development of an embryo are inherently higher-level concepts:

> The capacity of the notochord to resist bending as it extends the embryo comes from the structure of the whole notochord. Measurements at the level of the individual collagen fiber or fluid-filled cells that make up the structure would not reveal the mechanical properties of the whole notochord. (Davidson et al. 2009, p. 2157)

We can interpret this as a form of (mechanical) top-down causation because biomechanical features influence the development of vertebrate embryos through constraints on motility and bending of cells (Green and Batterman 2017). As in the case of feedback loops, the higher-level material parameters cannot be achieved through coarse-graining but requires measurements where the whole structure must be intact. Similarly, the action potential in neurons is a mesoscale parameter that cannot be measured or understood at the molecular or sub-cellular level because the property depends on the whole cell structure (Noble 2012). We should therefore not think of higher-level variables in this context as merely "smeared-out versions" of a more fundamental lower-level description. At the same time, the explanatory autonomy should not be overemphasized as there is clearly some connection between microstructure and material parameter values. In both physics and biology, modeling of materials over large spatial scales requires that scales are bridged, e.g., via the identification of RVEs. In biology, the relations between higher and lower-scale variables are often illustrated in diagrams through feedback relations going both up and down (see e.g., Noble 2012; Lesne 2013). Thus, the idea of a scale dependent *relative autonomy* offers an alternative to more extreme positions (reductionism or anti-reductionism), while capturing aspects of how scientists deal with multi-scale systems.

In summary, examples of multiscale systems that are heterogenous (unlike ideal gases) support a stronger interpretation of the causal role of higher-level variables in the context of mechanical top-down causation. Higher-level variables in such contexts are not coarse-grained in the weaker sense that they can be *derived* from lower-level details via simple averaging. Rather, the requirement for homogenization strategies in multiscale modeling highlights how higher-level variables are *relatively autonomous* from lower-level descriptions. The examples thus challenge an

assumption in the definition of mechanical top-down causation (Ellis 2008; 2012), but in doing so they also strengthen Ellis' point about the importance and ubiquity of top-down causation and emergence in science.

4. The practical importance of top-down causation

Debates on the possibility or strength of top-down causation are often assumed to primarily be of theoretical interest to philosophers. However, as emphasized by Ellis and others, top-down causation and equivalence classes have important practical implications (Auletta, Ellis and Jaeger 2008; Ellis, Noble and O'Connor 2012; see also Wimsatt 1976; 2007). Top-down causation, for Ellis, is not a mysterious metaphysical concept, but an empirical phenomenon that can be demonstrated through experimental intervention. He defines the following operational criteria for top-down causation:

> To characterize some specific causal effect as a top-down effect, we must demonstrate that a change of higher level conditions alters the sequence of processes at lower levels; we do this by changing higher level conditions and seeing what happens at the lower levels (Ellis 2012).

Thus, top-down causation is here given a concrete interpretation as a relation between system variables at different spatial scales or levels (see also Ellis 2016, p. 16, and Woodward, forthcoming). If intervening on macroscale variables can change the dynamic states of microscale processes, it has important practical implications for the design of experiments, for multi-scale modeling, as well as for discussions about where and how to intervene to control and change future outcomes (such as disease definitions and treatment modalities in medicine, for instance). In the following, we present further examples of this view.

4.1. Top-down causation and multi-scale modelling

Scientific explanations often highlight molecular activities (biology) or the role of laws and initial conditions (physics), but boundary conditions are often equally important for understanding system behaviors. In the context of multi-scale cardiac modeling, Noble (2012) argues forcefully that that the models describing the processes at the lowest scale, i.e., ordinary differential equations

describing ionic currents, cannot be solved without the boundary conditions determined by the cell voltage. It is important to note here that cell voltage is a parameter that cannot be defined at the molecular or subcellular scale. Similarly, models of the action potential at the level of cells depend on inputs from models at the tissue scale (defined via partial differential equations), which describe how biophysical features of the tissue can influence the propagation of electrical currents through the 3D structure of the heart (Qu et al. 2011; Green 2018).

Although parameters such as the cell voltage, or the geometrical and electrical properties of different tissue types are "nothing but" properties of physical structures, they are not reducible to or derivable from lower-scale variables. In fact, one cannot measure or even conceptualize these variables at lower scales. They also cannot be reduced to explanatory background conditions for descriptions of causal efficacy at lower scales, because the boundary conditions and higher-level parameters in general are required to channel the lower-level behaviors in the first place. As highlighted by Noble: "without the downward causation from the cell potential, there is no [heart] rhythm" (Noble 2012, p. 58). A reductionist perspective thus faces great difficulties in terms of showing that a bottom-up analysis is itself sufficient.

In the context of developmental biology and cancer research, the importance of top-down causation is also increasingly acknowledged. New experimental techniques to manipulate higher-level biomechanical cues have revealed that macroscale biomechanical properties (e.g., tissue stiffness) can influence gene expression, molecular signaling pathways, as well as cell differentiation (Miller and Davidson 2013). This has important implications for understanding how biomechanical constraints can buffer genetic "noise", and how there is sufficient causal slack between macroscale biomechanical models and molecular details to allow modelers to efface many lower-scale details. Moreover, it has (negative) implications for the view that genetic or molecular causation has a privileged role in developmental biology (this is further discussed in Green and Batterman 2017). The biases of such perspectives can also have important social implications, as we now clarify.

4.2. Research and treatment modalities in medicine

Ellis (2012) highlights the existence of different treatment modalities in medicine, depending on whether one commits to a reductionist (bottom-up) perspective or to a more holistic view. This is particularly apparent in the case of mental disorders, where the focus can span from genetic susceptibilities and molecular dysfunctions to how states of the mind can impact physical health

(see also Ellis 2016). Philosophical assumptions concerning the possibility and relative importance of top-down causation (vs. bottom-up causation) does, in fact, influence which research questions and treatment regime are seen as most promising. Attention to top-down causation can fruitfully point to the missed opportunities of the prioritized focus on reductionist research programs – and positively highlight the potential benefits of a broader perspective (Fuchs 2018).

In closing, we provide a few examples to illustrate this view. Cancer is often understood as a disease caused by accumulation of somatic mutations. However, increasing evidence suggests that tissue-scale properties can sometimes overrule genetic instructions (Soto et al. 2008; Bissel and Hines 2011). The constraints of the tissue can either promote or reduce cell proliferation and motility, depending on the biomechanical properties of the tumor microenvironment, which is another important example of top-down causation (Green, forthcoming). Hence, the reductionist perspective may create unfortunate blind spots, such as the opportunity to develop treatment strategies that target tissue-scale properties (Stylianopoulos 2017).

Similarly, the criticism of reductionism is highly relevant in the context of preventive medicine. With the promotion of precision medicine, the research focus on genetic factors that increase an individual's susceptibility for developing complex diseases like cancer, depression, or dementia has intensified. But a focus on genetic factors is neither sufficient for understanding and treating such complex diseases, nor is it necessarily more precise. An important concern is that genetic risk profiling at the individual level shifts attention away from structural causes at the population level, such as socio-economic disparities, that may be more efficient to intervene on (Hoeyer 2019; Olstad and McIntyre 2019). Top-down causation is therefore not only of theoretical philosophical interest, but it is an empirical phenomenon with profound scientific and social implications.

5. **Concluding remarks**

Explanatory autonomy of levels or scales is often defended with reference to the existence of universality or functional equivalence classes. Universal or functionally equivalent behaviors are described through macroscale models and parameters that cannot be reduced to lower-level models. Examples examined in this chapter include thermodynamics near critical points as well as feedback control in biology. The examples illustrate how models can be explanatory without specifying how a behavior is causally realized in any specific system. In fact, generic models are explanatory *because* they show how many causal details are explanatorily irrelevant as long as stabilizing structures defining an equivalence class are in place. The notion of information hiding highlights

that there is enough "causal slack" at the bottom to make the inclusion of all possible lower-level details irrelevant or even counterproductive for the purpose of explaining higher level or multiscale systems.

A key issue in discussions of reductionism and top-down causation is whether higher-level variables can be derived from lower-level details. Mechanical top-down causation is often seen as a weak form of top-down causation, because the high-level variables are taken to be nothing but coarse-grained correlates of lower-level variables. Thus understood, it is unclear whether reductionism is overcome. We have argued that upscaling of variables via coarse-graining only works for simple homogenous systems, such as an ideal gas. For systems with complex microstructures at the mesoscale (such as steel or bone), more involved upscaling techniques are required. The reason is that physical systems at different scales display distinct physical structures and behaviors, and that higher-level behaviors are dependent on some microstructural details that are best studied at the mesoscale.

Mesoscale parameters (such as material parameters) differ from what Ellis calls coarse-grained high-level variables in being identified via homogenization strategies. The need for such strategies signals a stronger explanatory autonomy of high-level variables also in physical examples than often assumed. Hence, attention to scale-dependency of characteristic behaviors in multi-scale systems offers support to Ellis' account by further extending the scope and significance of mechanical top-down causation. Top-down causation is not a suspicious, rare form of causation, but is ubiquitous in physical and biological systems alike. This has important practical implications not only for scientific modeling and explanation, but also for how we best approach complex socio-scientific problems.

References

Allen, Timothy FH, and Thomas B. Starr. 1982. *Hierarchy: perspectives for ecological complexity*. Chicago: University of Chicago Press.

Alon, Uri, Michael G. Surette, Naama Barkai, and Stanislas Leibler. 1999. "Robustness in bacterial chemotaxis." *Nature* 397, no. 6715: 168.

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

Auletta, Gennaro, George F.R. Ellis, and Luc Jaeger. 2008. "Top-down causation by information control: from a philosophical problem to a scientific research programme." *Journal of the Royal Society Interface* 5, no. 27: 1159-1172.

Batterman, Robert W. 2000. "Multiple realizability and universality." *The British Journal for the Philosophy of Science* 51, no. 1: 115-145.

Batterman, R. W. (2002). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence*. New York: Oxford University Press.

Batterman, Robert W. 2012. "The tyranny of scales." In *Oxford Handbook of Philosophy of Physics*, ed. Robert W. Batterman, 225–86. Oxford: Oxford University Press.

Batterman, Robert W. 2018. "Autonomy of theories: An explanatory problem." *Noûs* 52, no. 4: 858-873.

Batterman, Robert W. forthcoming. *A Middle Way: A Non-Fundamental Approach to Many-Body Physics*. Oxford: Oxford University Press.

Batterman, Robert W., and Sara Green. 2020. "Steel and bone. Mesoscale modeling and middle-out strategies in physics and biology." *Synthese*, published online: https://doi.org/10.1007/s11229-020-02769-y.

Bechtel, William. 2017. Top-Down Causation in Biology and Neuroscience. Control Hierarchies. In *Philosophical and scientific perspectives on downward causation* pp. 203-224, eds. Paoletti, Michele Paolini, and Francesco Orilia, New York, NY: Taylor & Francis.

Bedau, Mark A. 1997. "Weak emergence." *Noûs* 31: 375-399.

Bernard, Claude. 1927/1957. *An Introduction to the Study of Experimental Medicine*. New York: Dover Publications.

Bishop, Robert C. 2012. "Fluid convection, constraint and causation." *Interface Focus* 2, no. 1: 4-12.

Bissell, Mina J., and William C. Hines. 2011. "Why don't we get more cancer? A proposed role of the microenvironment in restraining cancer progression." *Nature medicine* 17, no. 3: 320.

Briat, Corentin, Ankit Gupta, and Mustafa Khammash. 2014. "Integral feedback generically achieves perfect adaptation in stochastic biochemical networks.", *arXiv*:1410.6064v3.

Brigandt, Ingo, and Alan Love. 2017. "Reductionism in biology." In the *Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. https://plato.stanford.edu /entries/reduction-biology/.

Campbell, Donald T. 1974. "Downward causation in hierarchically organised biological systems." In *Studies in the Philosophy of Biology: Reduction and Related Problems*, ed. Francisco Ayala and Theodosius Dobzhansky, 179–86. London: Macmillan.

Cannon, Walter B. 1929. "Organization for physiological homeostasis." *Physiological reviews* 9, no. 3: 399-431.

Christiansen, Peder V. 2000. “Macro and micro-levels in physics.” In *Downward Causation: Minds, Bodies and Matter*, eds. Peter Bøgh Andersen, Claus Emmeche, Niels Ole Finnemann, and Peder Voetmann Christiansen, 51–62. Aarhus: Aarhus University Press.

Crick, Francis, and J. Clark. 1994. "The astonishing hypothesis." *Journal of Consciousness Studies* 1, no. 1: 10-16.

Davidson, Lance, Michelangelo von Dassow, and Jian Zhou. 2009. "Multi-scale mechanics from molecules to morphogenesis." *The International Journal of Biochemistry & Cell Biology* 41, no. 11: 2147-2162.

DiFrisco, J., & Jaeger, J. 2019. “Beyond networks: mechanism and process in evo-devo. *Biology & Philosophy*, 34(6), 1-24.

DiFrisco, J., & Jaeger, J. 2020. “Genetic causation in complex regulatory systems: an integrative dynamic perspective.” *BioEssays*, 42(6), 1900226.

Doyle, Francis J., and Joerg Stelling. 2006. "Systems interface biology." *Journal of the Royal Society Interface* 3, no. 10: 603-616.

Emmeche, Claus, Simo Køppe, and Frederik Stjernfelt. 2000. “Levels, emergence, and three versions of downward causation.” In *Downward Causation: Minds, Bodies and Matter*, eds. Peter Bøgh Andersen, Claus Emmeche, Niels Ole Finnemann, and Peder Voetmann Christiansen, 13–34. Aarhus: Aarhus University Press.

Ellis, George FR. 2005. "Physics, complexity and causality." *Nature* 435, no. 7043: 743.

Ellis, George FR. 2006. "On the nature of emergent reality.” In *The re-emergence of emergence,* Eds. Philip Clayton and Paul Davies, Oxford: Oxford University Press, 79-107.

Ellis, George FR. 2008. "On the nature of causation in complex systems." *Transactions of the Royal Society of South Africa* 63, no. 1: 69-84.

Ellis, George FR. 2012. "Top-down causation and emergence: some comments on mechanisms." *Interface focus* 2, no. 1: 126-140.

Ellis, George FR. 2016. *How can Physics Underlie the Mind? Top-Down Causation in the Human Context*. Berlin: Springer.

Ellis, George FR. 2018. "Top-down causation and quantum physics." *Proceedings of the National Academy of Sciences* 115, no. 46: 11661-11663.

Ellis, George FR, Denis Noble, and Timothy O'Connor. 2012. "Top-down causation: an integrating theme within and across the sciences?" *Interface focus*: 1-3.

Ellis, George FR, and Jonathan Kopel. 2017. "On the difference between physics and biology: Logical branching and biomolecules." https://arxiv.org/abs/1709.00950

Fang, Wei. 2020. "Multiple realization in systems biology." *Philosophy of Science* 87: 663-684.

Fodor, Jerry A. 1974. "Special sciences (or: The disunity of science as a working hypothesis)." *Synthese* 28, no. 2: 97-115.

Fuchs, Thomas. 2018. *Ecology of the Brain: The Phenomenology and Biology of the Embodied Mind*. Oxford, UK: Oxford University Press.

Green, Sara. 2015. "Revisiting generality in the life sciences: Systems biology and the quest for general principles." *Biology and Philosophy* 30: 629-652.

Green, Sara. 2018. "Scale dependency and downward causation in biology." *Philosophy of Science* 85, no. 5: 998-1011.

Green, Sara. Forthcoming "Cancer beyond genetics: On the practical implications of downward causation." In *Biological Levels: Composition, Scale and Evolution in Complex Systems*, Eds. Daniel S. Brooks, James DiFrisco & William C. Wimsatt, MIT Press.

Green, Sara, and Robert Batterman. 2017. "Biology meets physics: Reductionism and multi-scale modeling of morphogenesis." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 61: 20-34.

Green, Sara, and Nicholaos Jones. 2016. "Constraint-based reasoning for search and explanation: Strategies for understanding variation and patterns in biology." *dialectica* 70, no. 3: 343-374.

Green, Sara, and Olaf Wolkenhauer. 2013. "Tracing organizing principles: Learning from the history of systems biology." *History and Philosophy of the Life Sciences* 35: 553-576.

Gross, Fridolin, and Sara Green. 2017. "The sum of the parts: large-scale modeling in systems biology." *Philosophy & Theory in Biology*, 9 no. 10.

Guggenheim, Edward A. 1945. "The principle of corresponding states." *The Journal of Chemical Physics* 13, no. 7: 253-261.

Hooker, Cliff. 2013. "On the import of constraints in complex dynamical systems." *Foundations of Science* 18, no. 4: 757-780.

Hoeyer, Klaus. 2019. "Data as promise: Reconfiguring Danish public health through personalized medicine." *Social Studies of Science* 49, no. 4: 531-555.

Huang, Sui. 2011. "Systems biology of stem cells: three useful perspectives to help overcome the paradigm of linear pathways." *Philosophical Transactions of the Royal Society B: Biological Sciences* 366, no. 1575: 2247-2259.

Isalan, Mark, Caroline Lemerle, Konstantinos Michalodimitrakis, Carsten Horn, Pedro Beltrao, Emanuele Raineri, Mireia Garriga-Canut, and Luis Serrano. 2008. "Evolvability and hierarchy in rewired bacterial gene networks." *Nature* 452, no. 7189: 840.

Jiménez, Alba, James Cotterell, Andreea Munteanu, and James Sharpe. 2017. "A spectrum of modularity in multi-functional gene circuits." *Molecular Systems Biology* 13, no. 4.

Juarrero, Alicia. 1998. Causality as constraint. In *Evolutionary Systems: Biological and Epistemological Perspectives on Selection and Self-Organization*, eds. G. Van de Vijver, S. Salthe and M. Delpos (pp. 233-242), Dordrecht: Springer.

Juarrero, Alicia. 1999. *Dynamics in Action: Intentional Behavior as a Complex System*. Cambridge, MA: MIT Press.

Kadanoff, Leo P. 1971. "Critical behavior, universality and scaling". In *Proceedings of the International School of Physics "Enrico Fermi" Course LI,* ed. M. S. Green, 100—117, New York: Academic Press.

Kauffman, Stuart A. 1969. "Metabolic stability and epigenesis in randomly constructed genetic nets." *Journal of Theoretical Biology* 22, no. 3: 437-467.

Kauffman, Stuart A. 1993. "The Origins of Order: Self-organization and Selection in Evolution. Oxford: Oxford University Press.

Kim, Jaegwon. 1998. "Mind in a Physical World." Cambridge, MA: MIT Press.

Kim, Jaegwon. 2000. "Making sense of downward causation." In *Downward Causation: Minds, Bodies and Matter*, eds. Peter Bøgh Andersen, Claus Emmeche, Niels Ole Finnemann, and Peder Voetmann Christiansen, 305–21. Aarhus: Aarhus University Press.

Koskinen, Rami. 2017. "Synthetic biology and the search for alternative genetic systems: Taking how-possibly models seriously." *European Journal for Philosophy of Science* 7, no. 3: 493-506.

Koskinen, Rami. 2019. "Multiple realizability as a design heuristic in biological engineering." *European Journal for Philosophy of Science* 9, no. 1: 15.

Lesne, Annick. 2013. "Multiscale analysis of biological systems." *Acta biotheoretica* 61, no. 1: 3-19.

Levy, Arnon, and William Bechtel. 2013. "Abstraction and the organization of mechanisms." *Philosophy of science* 80, no. 2: 241-261.

Lodish, Harvey, Arnold Berk, Chris A. Kaiser, Monty Krieger, Matthew P. Scott, Anthony Bretscher, Hidde Ploegh, and Paul Matsudaira. 2008. "*Molecular Cell Biology*." Macmillan.

Ma, Wenzhe, Ala Trusina, Hana El-Samad, Wendell A. Lim, and Chao Tang. 2009. "Defining network topologies that can achieve biochemical adaptation." *Cell* 138, no. 4: 760-773.

Mangan, Shmoolik, and Uri Alon. 2003. "Structure and function of the feed-forward loop network motif." *Proceedings of the National Academy of Sciences* 100, no. 21: 11980-11985.

Matthiessen, Dana. 2017. "Mechanistic Explanation in Systems Biology: Cellular Networks." *The British Journal for the Philosophy of Scienc,* 68, no. 1: 1-25.

Miller, Callie Johnson, and Lance A. Davidson. 2013. "The interplay between cell signalling and mechanics in developmental processes." *Nature Reviews Genetics* 14, no. 10: 733.

Moreno, Alvaro and Matteo Mossio. 2015. *Biological Autonomy: A Philosophical and Theoretical Inquiry*. Dordrecht: Springer.

Mossio, Matteo and Alvaro Moreno. 2010. "Organisational closure in biological organisms." *History and Philosophy of the Life Sciences*, 32:269-288.

Noble, Denis. 2012. "A theory of biological relativity: no privileged level of causation." *Interface focus* 2, no. 1: 55-64.

Noble, Denis. 2017. "Systems biology beyond the genome." In *Philosophy of Systems Biology*, ed. Sara Green, 227-235. Cham: Springer.

Olstad, Dana L. and Lynn McIntyre. 2019. "Reconceptualising precision public health." *BMJ open* 9, no. 9: e030279.

Oppenheim, Paul, and Hilary Putnam. 1958. "Unity of science as a working hypothesis." In *Minnesota Studies in the Philosophy of Science*, ed. Herbert Feigl, Grover Maxwell, and Michael Scriven, 3–36. Minneapolis: University of Minnesota Press.

Pattee, Howard H. 1971. "Physical Theories of Biological Co-ordination." *Quarterly Reviews of Biophysics* 4: 255–76.

Pattee, Howard H. ed 1973. *Hierarchy Theory*. New York, NY: Braziller.

Paoletti, Michele Paolini, and Francesco Orilia, eds. 2017, *Philosophical and scientific perspectives on downward causation*. New York, NY: Taylor & Francis.

Putnam, Hillary. 1980. "Philosophy and our mental life". In *Readings in Philosophy of Psychology*, ed. Ned Block, volume 1, pp. 134-143. Cambridge, MA: Harvard University Press, Cambridge.

Qu, Zhilin, Alan Garfinkel, James N. Weiss, and Melissa Nivala. 2011. "Multi-scale modeling in biology: how to bridge the gaps between scales?." *Progress in Biophysics and Molecular Biology* 107, no. 1: 21-31.

Salthe, Stanley S. 1985. *Evolving hierarchical systems. Their structure and representation*. New York, NY: Columbia University Press.

Serban, Maria, and Sara Green. 2020. "Biological robustness: Design, organization, and mechanisms." In *Philosophical Perspectives on the Engineering Approach in Biology: Living Machines?,*141-164, eds. Sune H. Holm & Maria Serban. New York, NY: Taylor and Francis.

Sober, Elliott. 1999. "The multiple realizability argument against reductionism." *Philosophy of Science* 66, no. 4: 542-564.

Soto, Ana M., Carlos Sonnenschein, and Paul-Antoine Miquel. 2008. "On physicalism and downward causation in developmental and cancer biology." *Acta Biotheoretica* 56, no. 4: 257-274.

Stylianopoulos, Triantafyllos. 2017. "The solid mechanics of cancer and strategies for improved therapy." *Journal of Biomechanical Engineering* 139, no. 2: 021004.

Torquato, Salvatore. 2002. "Random Heterogenous Materials: Microstructure and Macroscopic Properties." New York: Springer.

Tyson, John J., Katherine C. Chen, and Bela Novak. 2003. "Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell." *Current opinion in cell biology* 15, no. 2: 221-231.

Tyson, John J., and Béla Novák. 2010. "Functional motifs in biochemical reaction networks." *Annual review of physical chemistry,* no. 61: 219-240.

Ulanowicz, Robert E. 1986. *Growth and development. Ecosystems phenomenology*. New York, NY: Springer.

Ulanowicz, Robert E. 1997. *Ecology, the ascendent perspective*. New York, NY: Columbia University Press.

Wiener, Norbert. 1948. "Cybernetics or Control and Communication in the Animal and the Machine". MIT press.

Wimsatt, William C. 1976. "Reductionism, levels of organization, and the mind-body problem", In *Consciousness and the Brai*n, 199-267, ed. G.G. Globus, G. Maxwell, and I. Savodnik. New York: Plenum.

Wimsatt, William C. 2007. *Re-Engineering Philosophy for Limited Beings*. Cambridge, MA: Harvard University Press.

Woodward, James. Forthcoming. "Downward causation and levels." In *Biological Levels: Composition, Scale and Evolution in Complex Systems*, eds. Daniel S. Brooks, James DiFrisco & William C. Wimsatt (Eds.), MIT Press.

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, no. 9: 4649-4653.