Introduction

Levels of Explanation

Katie Robertson and Alastair Wilson

The world around us can be explained on many different levels. 'Why'-questions can have multiple distinct correct answers, with context playing a central role in determining what is asked and how it is answered. Different explanations of the same phenomenon can enhance, rather than exclude, one another; a complete understanding of the phenomenon requires grasping all these explanations. Understanding levels of explanation and how they relate is the main project of this volume.

Explanations at different levels, distinctively, complement rather than clash. That is to say that distinct candidate explanations at a single level tend to exclude one another in a way in which candidate explanations at different levels need not. The bushfire was very likely started either by a lightning strike or by a discarded match; it is very unlikely that any individual fire was started by both acting together. But explanations of the bushfire in terms of a lightning strike and in terms of the effects of climate change on extreme weather events need not exclude each other; the former may explain the fire at the level of weather, while the latter explains the very same fire at the level of climate. Each explanation plays a role in our overall understanding of the fire.

Levels of explanation, as we understand them here, are commonplace and may be grasped from an early age. Children are voracious consumers of explanations. When a curious child asks a question, 'why P?' and receives an answer, 'because Q', typically she will accept the answer and move on, or ask a follow-up question, 'why Q?'. But sometimes she will frown, and repeat her initial question. It turns out that further elaboration of the answer Q does not help; she understands it perfectly well already. What her frown is calling for is an explanation on a different level from the one she has been offered.

Our practice of multilevel explanation is familiar and everyday, but it also takes centre stage in some of the most sophisticated contemporary work in philosophy of science and metaphysics. In this volume we have collected together some of the best of this philosophical work to give an overview of explanatory levels[1] and of

[1] We use 'explanatory level' and 'level of explanation' interchangeably.

Katie Robertson and Alastair Wilson, *Introduction* In: *Levels of Explanation*. Edited by: Katie Robertson and Alastair Wilson, Oxford University Press. © Katie Robertson and Alastair Wilson 2024. DOI: 10.1093/oso/9780192862945.003.0001

their applications. Explanations at multiple levels can coexist without conflict or redundancy, can be employed for different epistemic or pragmatic purposes, and can be combined together to give richer explanatory models. These features give them broad potential application and so it is no surprise that talk of levels of explanation is commonplace across a wide range of disciplines. This book touches *inter alia* on explanatory level structures in cognitive science, sociology, molecular biology, materials science, particle physics, geometry, set theory, and the metaphysics of dispositions.

While many of the contributors to this volume are enthusiasts about levels of explanation—seeking to either contribute to our understanding of levels and their scientific usefulness, or to apply levels to solve philosophical problems raised by science, mathematics, or metaphysics—the concept has also come in for plenty of critical discussion. Explanatory level sceptics in one way or another argue that levels frameworks distort the scientific reasoning they are intended to illuminate. Eronen (2015), for example, argues that the work of explanatory levels is better assigned to specific notions of scale and composition, while Potochnik (2017) likewise argues that there is no useful generalized notion of explanatory level. Sceptical voices are represented in this volume: in particular, the chapters by Bechtel and Chirimuuta question straightforward realism about levels of explanation, while Potochnik argues we should reject the entire framework. Others, such as Franklin, address extant challenges for explanatory level frameworks. However, most of the contributors are positive about the usefulness of explanatory levels, seeking either to understand levels or to exploit them in their various projects. One of our aims with the volume is to give a rounded picture of what the concept of levels of explanation can do for philosophers, and thereby to give a clearer idea of the costs that would be involved in rejecting it as the sceptics urge.

As an initial step in systematizing our thinking about levels of explanation, we can model levels as classes of answers to explanatory questions—typically 'why?'-questions but potentially also 'how?'-questions. For some given explanatory question there will usually be many prospective answers, and each question-answer pair may then be grouped into classes unified by the equivalence relation 'is at the same level as'. Of the prospective question-answer pairs at a given level, one (or maybe more) may be correct; but there is no presumption that every explanatory question has a correct answer at every level. For further insight into the nature of levels of explanation, we then need to ask after the nature of the same-level-as relation (or, perhaps more saliently in some circumstances, the different-level-as relation): what is it for two candidate explanations of some phenomenon to be at the same level? Which features make for explanatory stratification? Different philosophical accounts of levels of explanation, including those explored in some chapters of this volume, fill in the details in different ways. In particular, the chapters by List, Chirimuuta, Kincaid, Crowther, Knox, and Strevens provide illustratively different accounts of the nature of explanatory levels which have this core structure in common.

Over recent decades, it has gradually become orthodoxy in the philosophy of explanation that the best explanations need not reside at the fundamental level. In Potochnik's chapter, this thesis is named *explanatory anti-reductionism*. The higher-level explanations which higher-level sciences provide are central to the case for the irreducibility of the respective higher-level theories, and consequently to the case for the reality of the emergent entities that these theories describe. Explanations in higher-level terms are said to be more proportionate (Yablo 1992), deeper (Hitchcock and Woodward 2003; Strevens 2008), more abstract (Knox 2016), and/or more computationally tractable (Weisberg 2007) than explanations given in fundamental terms.

Explanations which reside wholly at a single level—whether they are provided by (putatively fundamental) classical field theories or by (evidently non-fundamental) evolutionary biology—remain conceptually relatively straightforward. But as well as mapping individual explanations at higher levels, scientific practice is rich in explanations drawing on multiple levels at once. Many of the hardest philosophical puzzles of explanation derive from the interplay of explanatory factors across multiple levels of description, as when the cooling of a cup of coffee is explained in terms of collective molecular motion, or when a cognitive process causes a bodily movement. In some cases lower-level events might directly influence higher-level events, such as when an atomic decay causes an explosive nuclear chain reaction, and sometimes higher-level events might directly influence lower-level events, such as my desire to type causing my fingers to move. Any account of levels of explanation must be able to do justice to the links between levels.

How are higher-level explanations connected to and constrained by other levels of explanation? Strategies for answering this question have tended to vary across subdisciplines. Debates about multilevel causation have a long pedigree in philosophy of mind, but recent work on causal modelling has both shed new light on mental causation and tied the problem closely to the philosophy of science, where a causal modelling account of explanation is dominant; Weslake's chapter provides an up-to-date account. More generally, enthusiasm about levels of explanation has tended to go hand in hand with enthusiasm about emergence in some domain or other. This motivating connection between emergence and explanatory levels plays a central role in the chapters by Knox, Franklin, and Crowther.

In philosophy of physics, theory reduction has often been the focus—at least since the highly influential work of Nagel (1935, 1949, 1961, 1970) (later refined by Schaffner (1967, 1990)) in which examples from physics played a central role. Reductionists in this tradition, including Lewis (1994), Loewer (2001), Butterfield (2011), and many others, have tended to endorse the idea of 'in principle' reducibility of high-level physical phenomena to lower-level phenomena. Meanwhile, anti-reductionists such as Cartwright (1999) and Dupré (1993) have argued we have no evidence for and some evidence against the thesis of in-principle-reducibility. However, authors including Batterman (2013, 2021) and Wilson (2017) have argued that this focus on reduction distorts scientific practice, arguing for a more nuanced picture. For Batterman, the 'tyranny of scales'—the stark choice between top-down and bottom-up explanatory strategies—is a false dichotomy.

In philosophy of biology, mechanisms have often been seen as key to understanding complex causal explanations. The right level at which to understand a biological system, in mechanistic terms, is one at which the system's behaviour can be understood as resulting from the interaction of distinct subsystems with distinct characteristic functions (Craver (2007), Bechtel (2008)). Another key concept has been levels of organization; see Wimsatt (1976) for an early statement, and Brooks, DiFrisco, and Wimsatt (2021) for the state of the art. Levels of organization bear some similarities to our general concept of levels of explanation, but also differ in certain respects. We discuss both mechanistic levels and levels of organization below, as potential rivals to the explanatory levels approach.

In metaphysics, levels of explanation have often been discussed in the context of Jaegwon Kim's causal exclusion argument (Kim 1983); levels offer a possible route to a robust explanatory role for higher-level properties in mental causation, free action, and related phenomena. Metaphysical relations including grounding and realization have been offered as candidates for the level-connection relation between scientific levels; grounding is often supposed to connect levels 'vertically', with causation connecting events 'horizontally' (Fine (2012), Bennett (2017)). But we can also identify cases of levels of explanation which are distinctively metaphysical: in many cases, theorizing in metaphysics aims to provide explanations at some level even more fundamental than the level of fundamental physics. When metaphysicians aim, for example, to account for the instantiation of quantitative properties in terms of relations to universals (Mundy (1987)) or in terms of mereology (Perry (forthcoming)), or when they aim to reduce spacetime to a causal structure between events (Baron and Le Bihan (forthcoming)), they can be regarded as operating at a distinctively metaphysical level of explanation.

The levels of explanation framework as we understand it here is general, in that it is applicable to many kinds of investigative context and to many varieties of explanation—not just to causal explanation, for example. This makes the framework flexible and open-ended, such that one can see different chapters of this volume as developing the general approach in very different ways. But the explanatory levels approach is not without distinctive content; not just any old description counts as an explanatory level. To provide a clearer sense of the levels concept we have in mind, we next contrast it with other approaches to levels which have appeared in the philosophy of science literature.

Rival one: descriptive levels. A descriptive level is a kind of imprecision in description: we describe our target system up to a certain level of detail, but not beyond. As authors including Lewis (1988a, 1988b) and List (2019) have emphasized, descriptive levels can be readily characterized using only simple modal resources: the basic idea is that the less-detailed/coarser-grained description supervenes on the more detailed/finer grained description, but not vice versa. However, for our purposes descriptive levels are too thin.[2] For a start, they are cheap; any arbitrary choice of details to exclude or include, however gerrymandered, characterizes a descriptive level. More tellingly, descriptive levels bear no direct epistemic significance; we learn nothing of any substance about a system when we are told that it can be described at a certain level of detail. Dropping decimal places in a quantitative description would, on this approach, automatically give a new level—without this being at all enlightening. Knox emphasizes this point, arguing that levels which are useful in physics need to be linked to *useful* variable changes rather than to any variable change whatsoever. Identifying distinct levels of explanation of some phenomenon is a non-trivial epistemic achievement which requires us to be able to identify—if not yet answer—distinct classes of explanatory questions concerning the phenomenon.

Rival two: levels of scale. Here talk of levels is understood as relating entirely to phenomena analysed at different physical scales—for example, at a length scale corresponding to metres or at an energy scale corresponding to gigaelectron-volts. This representation of levels is often found in introductory textbooks in biology or in introductory physics courses, where one presents the subject as spanning/investigating from the scale of subatomic particles to the scale of galaxy clusters. But there are also more sophisticated accounts of inter-theoretic relations which focus on scale: Ladyman and Ross (2007) incorporate the notion into their distinctive thesis of scale-relativity of ontology. Thinking of levels of explanation in terms of scale doesn't work across the board, however: variation in scale is too inflexible to accommodate all of the different ways in which explanatory levels might interrelate. For example, the explanatory levels structures characterized by List in Chapter 1 are typically partial orderings, while the ordering of levels on any given scale is a total ordering; Potochnik's chapter also discusses this point. Kincaid's chapter argues that scale is insufficient for understanding levels of explanation in social science. And Knox discusses the example of the 'level of Newtonian physics': this way of thinking is useful for some explanatory purposes, but it does not correspond neatly to any physical scale. In some cases perhaps—in particular the literature on effective field theories (Castellani (2000), Franklin (2018), Wallace (2019))—scale and explanatory levels do correlate well. But the link between scale and explanatory levels is not fully general.

[2] This is not to say that descriptive levels are useless, of course; they have been employed as modelling devices within metaphysics and philosophy of language (Lewis (1988a, 1988b), Yablo (2017)). Indeed, descriptive levels may be employed as part of a substantive theory of levels, as in List's approach in Chapter 1; our point here is that some further ingredient is needed.

Rival three: compositional levels. Here we have in mind the tradition following the classic paper of Oppenheim and Putnam (1958), who identify entity-based levels such that the smaller things at one level generically compose the larger things at the next level up. Other more sophisticated versions of this approach include Schaffer (2009), who uses grounding as an interlevel link as part of his priority monism. Compositional levels find few supporters these days, however, because of the way they tend to generate an exclusive focus on the compositional and mereological relations between the entities they structure. The 'Lego' view of science no longer seems tenable: interaction effects between the constituents of composites are endemic in real-world level structures. An atom isn't just a simple aggregate of its constituent nucleons and electrons; instead it derives nearly all of its interesting physical and chemical properties from the balance of interactions between these constituents. Likewise, in the classic application of explanatory levels to cognitive science by Marr (1981) discussed in Chirimuuta's chapter, Marr raises the worry that the whole can't be effectively studied by studying only the parts.

Rival four: levels of organization. Taking account of the limitations of compositional accounts of levels tends to take us towards a different class of rival approaches, which retain certain compositional aspects but emphasize organization of the elements. This category most prominently includes the kind of levels of organization deriving from the work of Wimsatt (1976), which are prominent in the philosophy of biology. Crowther's chapter also touches upon the usefulness of levels of organization in physics, and they can be modelled using List's notion of an ontological level. For our purposes, however, levels of organization are unsuitable: they lack a sufficiently direct link to explanatory value. The world may be organized in all sorts of different ways at different scales, but only those arrangements which support non-trivial explanations are capable of giving rise to explanatory levels in our sense. But to build an explanatory criterion into the levels of organization approach then turns it into a version of our own framework (although one which may be more restricted in terms of which kinds of explanations can be involved). To put things slightly differently: insofar as levels of organization give rise to useful explanations, they may either be identified with, or correlated with, levels of explanation in our sense.

Rival five: mechanistic levels. Mechanistic levels are perhaps the closest existing approach to explanatory levels as we envisage them: they have a constitutive link to explanations of the mechanistic sort (Craver (2007), Andersen (2014a, 2014b)). However, we don't want to restrict the scope of our account only to levels of mechanistic explanation. Mechanistic approaches to explanation in physics (Felline (2022)) remain somewhat underdeveloped; mechanistic approaches get no traction at all in mathematics or metaphysics, or more generally where the relations between levels do not have any compositional character.[3] Our framework for levels

[3] Some chapters of this volume explore levels associated with clearly non-mechanistic explanation: Taylor considers metaphysical explanation, Antos and Colyvan consider mathematical explanations.

of explanation can be, but need not be, combined with an imperialist thesis that all explanations are mechanistic explanations; so explanatory levels are more general than mechanistic levels.

One crucial point which our preferred framework of levels of explanation leaves open (whereas many of the accounts already discussed settle it in some way) is the status of the explanations involved as more or less objective (or mind-independent, or interest-nonrelative). The framework can be combined with robust scientific and metaphysical realisms (as in the chapters by List, Franklin, and Kincaid) or with some variety of pragmatism about explanation (as in the chapters by Chirimuuta, Bechtel, and Hicks). The discussion in Chirimuuta's chapter about the connection between epistemic and ontological levels is relevant here, as is the tripartite distinction between types of levels—descriptive, ontological, explanatory—in List's chapter.

We regard it as a key advantage of explanatory levels over other types of levels frameworks that it leaves the question of explanatory realism open, along with other central metaphysical and epistemological questions. Thinking in terms of levels of explanation does not commit us to any particular approach to explanation; rather, most questions about the status and epistemic role of explanatory levels are outsourced to the broader account of explanation that is combined with the levels framework. As an example, accounts of explanation which make explanation highly context-dependent and localized (and hence not governed by any universal generalizations, as in the classical deductive-nomological (D-N) approach) will typically give rise to highly localized levels structures, as described by Knox, Franklin, and Bechtel. The resulting levels are not characterized neatly by the indiscriminate application of some individual theory or theories: to use another example of Knox's, it makes no sense to talk of 'the sun at the level of statistical mechanics'. Instead, each level is constituted by a patchwork of different theories operating together in a specific modelling setting.

Flexibility about the nature of explanation allows explanatory levels to be a more neutral starting point for levels discussion. This neutrality does, however, come with risks of its own: there may be a temptation to equivocate on the notion of levels, for example in a slide from the modelling of a scientific practice in terms of epistemic levels to the attribution of a levelled structure to reality. Chirimuuta warns of the danger of this slide, while other authors including Kincaid, Knox, and McKenzie seek to place restrictions of various kinds on circumstances in which the inference from epistemic stratification to real explanatory levels is well founded.

The value, or the function, of talk of explanatory levels is a first key theme linking chapters across the volume. Why are levels of explanation useful? We have already discussed some advantages of explanatory levels over other accounts of levels—neutrality on matters of background metaphysics and epistemology, combined with the robust epistemic implications which flow from explanatory levels' constitutive link with explanation. Explanatory levels, on our account, are identified in terms of questions asked and answered, which does not tie them to any specific entity-based or scale-based metaphysics. So explanatory levels may be metaphysically thin, but they are epistemically thick: they have robust links to understanding, prediction, and manipulation.

In our terms, acquiring knowledge of explanations of some phenomenon at multiple different levels *enriches* our explanatory knowledge of the phenomenon. This sense in which multiple explanations at different levels confer richness of explanatory knowledge may be contrasted with the depth of an individual explanation, where (roughly) deeper explanations are those which generalize better (see Hitchcock and Woodward (2003), Weslake (2010)). If we assume a simple link between explanation and understanding—that understanding is a matter of possessing suitable explanatory knowledge—then it immediately follows from what we have said that possessing richer explanatory knowledge gives rise to richer understanding. On more complex accounts of the explanation-understanding link, we may look to account in some different way for the role of explanations at different levels in increasing understanding, but it is plausible that there is a close relationship: better multilevel explanations provide better—in our terms, richer—understanding.

Of course, levels of explanation have their detractors. Some critiques target the coherence of explanatory level frameworks, but more often it is argued that in some sense or other levels of explanation are too narrow a framework in which to fit the complexities of actual scientific inquiry. A clear example of this kind of critique is Kim (2002), who rejects the Oppenheim-Putnam model for requiring a total ordering of levels and instead advocates a more flexible partial ordering, according to which entities can be incommensurable with respect to level, with neither higher level than the other: Kim offers the example of plants and animals. It is also very natural to think that psychology and geology operate at incommensurable explanatory levels. Another style of critique—emphasized by authors as different as Jackson and Pettitt (1990) on programme explanations, Gillett (2016) on machretic explanation, and Potochnik and Yates in their chapters in this volume—is that structural and contextual conditions at different levels may be equally or more crucial to understanding a system's behaviour than external causal agents at the same level as the system.

A common factor in these critiques is that an adequate framework for levels of explanation should not be too rigid; placing metaphysical constraints on what belongs to which level, or on how explanations align with features of some other metaphysical hierarchy, is liable to render a levels concept insufficiently flexible for the full range of applications needed. In light of this sort of consideration, the chapters in this volume which make positive use of levels of explanation tend to employ a flexible concept of level, which invokes partial rather than total ordering (List, Knox, Kincaid) and which allows for multilevel explanation rather than restricting explanations to hold only within individual levels (Franklin, Weslake, Hoffmann-Kolss, Yates).

Adopting a highly flexible conception of levels may avoid various objections, but it risks generating a new kind of over-flexibility problem. If we allow for explanations to range across levels, do we risk losing distinctions between the levels altogether, such that science and metaphysics collectively characterize one single 'wide' level? This is the problem of *too much interaction across levels*: we regard it as an important challenge for detailed spelling-out of theories of explanatory levels. Different approaches will tend to solve this problem in different ways, by identifying some substantive criterion other than mere potential explanatory relevance by which to stratify the different levels. For example, Knox's chapter individuates levels via changes of variable that are physically useful, Strevens's chapter individuates them in terms of probabilistic 'semi-detachment', and the cognitive-science levels of explanation of Marr (1981) (the subject of Chirimuuta's chapter) are individuated in functional terms.[4] Relatedly, some characterizations of explanatory levels have appealed to explanatory proportionality (in the sense of Yablo (1992)) to help individuate levels: even if explanations can hold across levels, the most proportionate explanations might still be between variables at the same level. Wimsatt (1972, 2007) employs a strategy of this kind in terms of 'local' explanations; this strategy is the target of criticism in Potochnik's chapter.

A second major theme of the volume, then, is the opportunities and pitfalls of using models which span different explanatory levels. The existence and usefulness of multilevel explanations refute simplistic approaches to levels according to which explanation always holds within and only with any given level. Typical compositional level approaches tend to sharply distinguish explanation at a given level[5] (typically causation) from explanation across levels (typically grounding or constitution). We take it that one lesson of the literature on levels of explanation is that these types of explanation are not straightforward to disentangle. In different ways, this entanglement is addressed in the chapters by Potochnik, Franklin, Weslake, Hoffmann-Kolss, Yates, and Bechtel.

Suppose that the coherence and applicability of explanatory levels is granted; why think that they are indispensable? One might instead attempt to understand the phenomena we have been aiming to capture simply in terms of the concurrent use of multiple models. Why, it might be asked, do we need to posit anything like levels at all in order to legitimize the use of multiple explanatory models for different theoretical and practical purposes? What we will call the multiple-model approach dispenses with level talk altogether in favour of exclusive talk of the use of multiple models.

[4] The relevant functions here: solving a specific computational problem, connecting specific input to specific outputs, or implementing specific mechanical operations.

[5] Potochnik's chapter refers to this as 'local' explanation.

The difference between explanatory levels and multiple-model approaches ought not to be overstated. Both approaches involve denying that any single explanatory model can tell the whole story about systems of interest; both approaches involve putting together distinct explanatory models in certain ways to give a richer understanding of the target systems; in each approach, we need to take care not to mix levels or models in ways which lead them to break down. Generally, though, levels-based approaches are associated with a stronger commitment to the unity of science, and correspondingly there is some pressure on defenders of these approaches to exhibit in detail the relation between levels. By contrast, multiple-model approaches leave open whether the models are connected in any interesting manner. In particular, levels frameworks typically assume some kind of 'level connector'; a relation between levels, typically transitive, which makes sense of how a higher level can harmlessly coexist with a lower level.

The explanatory level framework in general need not be committed to any specific level connection; indeed, the level connection might vary from case to case. And of course, how to understand this connection is highly controversial amongst level enthusiasts. A range of options for level-connectors have been considered in the literature: supervenience (e.g., List (2019), List's chapter in this volume, Ladyman and Ross (2007), Strevens (2008), Woodward (2021)), reduction (Oppenheim and Putnam (1958), Rosaler (2015, 2019), Crowther in this volume), grounding (Bliss and Trogdon (2021), Bryant (2018)), diachronic emergence (Humphreys (2016), Guay and Sartenaer (2016)), and other more epistemic and pragmatic notions (Dennett (1991), (2009), Chirimuuta in this volume). List's framework for explanatory levels in Chapter 1 is specified in terms of supervenience, and he explicitly argues that this doesn't entail a reductive relation between levels. However, most discussion of levels in physics foregrounds reduction (as in Franklin's and Knox's chapters), whereas contemporary discussions in the metaphysics of science often focus on grounding (as in McKenzie's and Crowther's chapters). What isn't controversial within the explanatory levels approach is that there is *some* link between the levels. This highlights again what we see as one of the underlying functions of levels talk—that it illuminates the different levels of explanation of a phenomenon as aspects of a larger unified explanatory whole, rather than as a collection of dissonant and disconnected fragments of explanation.

Our final theme is the applicability of the levels of explanation framework beyond science to explanations within philosophy—including to those explanations which we give while thinking about explanation itself.[6] Throughout this volume

[6] This kind of turning of theories of explanation on themselves has precedent in the recent literature: Emmerson (forthcoming) compares different metaphysical accounts of explanation in terms of their explanatory depth. Taylor's chapter in this volume also considers the 'explanatory distance' associated with certain explanations in the theory of explanation.

we find candidate explanations being offered for some feature or other of our practice of explaining things at different levels. In several cases, we can think of these candidate explanations as operating at different levels of explanation.

The question of how levels of explanation are possible is closely linked to the broader question of how higher-level explanations are possible at all. Why don't we have to depend on quantum field theory for every explanation we give? Why don't we always use physics for every explanation—why is there *independence* from physics at the higher levels? Fodor (1998) declared himself unable to answer this question;[7] Loewer (2008) makes an attempt at it. We see the chapters in the final section of the volume as collectively contributing to this explanatory project. We face a fraught balancing act: higher levels are usually thought to depend in some way on lower levels (cf. List's discussion of supervenience or Knox's discussion of reduction)—but this dependence mustn't crowd out the independence of the higher level. The conundrum is how to have the right amount of independence (such that there are levels at all) and the right amount of dependence (such that we can make systematic sense of the connections between levels). The fingerprints of this conundrum are found in a variety of related debates: in the reduction-emergence debate, emergence emphasizes higher-level independence, whilst reduction emphasizes the instances of interlevel dependence. Similarly, much of the philosophical interest in effective field theories, and renormalization group methods, as discussed by McKenzie, comes from the possibility that these mathematical methods give us insight into how levels are possible.

To illustrate this theme, the final section, VI, contains three chapters offering different kinds of explanation of how levels of explanation are possible—Bhogal in terms of a pragmatic account of naturalness which makes levels seem more or less inevitable, Strevens in terms of objective probabilistic relations between facts at different scales, and Hicks in terms of features of the underlying Humean laws of nature. These accounts do not directly compete—indeed all three could in principle be combined into a consistent neo-Humean account of explanatory levels and their epistemic role. Accordingly, the three chapters can be seen as complementary explanations, at different levels, of how explanatory levels are possible.

The theme of explaining how levels of explanation are possible connects Section II, which is about the scope and limits of causal modelling at different levels, with Section V, which is about the metaphysical preconditions for level structures. We can factor the preconditions for multilevel explanations into two sorts. We must be able to keep track of the different variables at the different levels and how they are incorporated into a single model; this is the focus of the chapters by Weslake, Hoffmann-Kolss, and Yates. The relevant variables themselves must in addition

[7] 'Well, I admit that I don't know why. I don't even know how to think about why. I expect to figure out why there is anything except physics the day before I figure out why there is anything at all, another (and presumably related) metaphysical conundrum that I find perplexing' (Fodor 1998, p. 161).

actually stand in the right relations for the multilevel explanatory practice to gain traction; this requirement is scrutinized in the chapters by Knox, McKenzie, Emery, Bhogal, Strevens, and Hicks.

Having laid out the main focus and themes of the volume, we turn next to summarizing the individual chapters and identifying their relations to these larger themes. The contributions are grouped into six sections: I) on the foundations of level frameworks, II) on levels of causal explanation, III) on explanatory levels in higher-level sciences, IV) on explanatory levels in physics, V) on levels of explanation in mathematics and in metaphysics, and VI) on the metaphysical conditions which make explanatory levels possible.

Section I of the volume presents and compares some different frameworks for thinking about explanatory levels in the sciences and relating them to other level-based hierarchies. Capturing the relationships between different levels of explanation is a central challenge, but one which is too often answered only schematically or metaphorically. Christian List has in recent work offered a unified formal framework for modelling different types of levels and the relationships between them (List, 2009). In his chapter, List extends this influential framework in several new ways. Typically the relationships between levels have been modelled as supervenience mappings. List's chapter considers and compares multiple different interpretations of the mappings between levels, including supervenience, grounding, and reduction. In particular, he discusses how formal features of grounding such as irreflexity and the potential lack of transitivity fit less easily into his formal framework. List then goes on to explore the conditions under which supervenience entails reducibility. Since Oppenheim and Putnam (1958), level hierarchies have predominantly been 'entity based'; whether compositional or not, the levels are populated by individuals. By contrast, List endorses a 'fact-based' rather than an 'entity-based' conception of levels, and in his chapter he defends this key foundational move. Armed with this precise and formal account of levels, List then goes on to demonstrate the ways in which such precise levels talk can be useful. By being careful about which level a concept belongs to, or at which level a question is asked, new light can be shed on questions ranging from free will to chance and determinism.

Whilst List's chapter considers the technical foundations of explanatory levels, Potochnik's chapter challenges the core motivations for the levels paradigm. A key original motivation for thinking in terms of levels of explanation stemmed from the explanatory anti-reductionist view that explanations in non-fundamental terms are ineliminable, and indeed often provide the best explanation of a given phenomenon. But Potochnik argues that the focus on a hierarchical structure of levels places artificial constraints on our theorizing about scientific explanation and especially about its pragmatic and interest-relative aspects. In particular, the interesting explanations that different sciences generate often don't have any clear relations of metaphysical determination between them; instead different sciences such as neuroscience and biology investigate different causal factors contributing to a particular phenomenon. Sometimes these causal factors will stem from structural, or non-local, factors. Potochnik argues that the plausible thesis of scientific anti-reductionism is not best explicated in terms of levels, since the defining features of different explanations are not readily characterized as defining levels. Instead, Potochnik suggests, there is a great variety of explanations that bear no special relationship to one another but simply feature different influences.

In recent work, Alex Franklin has argued that the value of higher-level explanation needn't rest on a failure of reduction. In his chapter for this volume, Franklin turns to the multiscale modelling practice recently discussed by authors including Robert Batterman, Julia Bursten, and Mark Wilson (Batterman (2013, 2021), Bursten (2018), Wilson (2017)). One theme of this recent work is that the relationships between levels are more complicated than mere averaging of variables, and often unrealistic assumptions are involved, such as assuming the environment to be uniform. The way that models involve different scales is taken by some to imply that the world cannot be neatly divided into levels that are then connected by reduction. This multiscale argument is incompatible with methodological reductionism—which Franklin joins Bursten, Batterman, and Wilson in rejecting—but he asks: is it compatible with a more sophisticated form of reductionism? Franklin answers yes, provided that we are suitably nuanced in our account of reduction. In agreement with proponents of the multiscale argument, Franklin argues that we should pay attention to the complexities of modelling. Doing so will involve local applications of collections of techniques—rather than whole theories—which signals a move away from a Nagelian view of reduction, and a step towards local, and contextual, levels. But there is still room for reductive explanations which do not explain the phenomena of the multiscale model; rather, they explain the effectiveness of that model for its target phenomena. Why we use particular mesoscopic and macroscopic variables should be explicable from the bottom up, if this more nuanced reductionist position is to succeed.

Section II of the volume focuses on the complexities of multilevel causal explanation. The most familiar form of explanation is causal, and in recent years our understanding of causal explanation has been revolutionized by the formal framework of causal modelling and the interventionist interpretation of this framework, associated with Judea Pearl et al. (2000) and James Woodward (2003). Causation is most familiar to us in the higher-level domain of macroscopic objects, stable over time, and interacting with nearby objects; the philosopher's paradigm case is a billiard-ball impact. But causal explanation extends throughout the higher-level sciences, from evolutionary biology to cognitive psychology to astrogeology, and it can hold across levels, as when a cosmic ray causes a cell mutation, or when a human decision causes a click of a mouse. Whilst the recent developments in scientific causal modelling techniques are hugely promising and successful, they are tested to their conceptual limits by application to multilevel phenomena, and an influential objection by Michael Baumgartner casts doubt on the applicability of causal models to such phenomena (Baumgartner (2009)).

In his chapter, which has already seen considerable circulation and discussion in manuscript form, Brad Weslake argues that the popular interventionist account of causation offers a distinctive built-in escape route from Jaegwon Kim's notorious causal exclusion argument (Kim, 1993). Weslake defangs Baumgartner's objection by imposing a condition—the metaphysical possibility of the independent manipulability of the variables in a model—which enables mental and physical variables to be coherently combined into models. This permits a vindication of the possibility of causation of physical effects by mental causes (and vice versa) on both internalist and externalist understandings of mental content. But interventionist accounts of multilevel explanation are not yet out of the woods; Weslake draws attention to a problem raised by Rescorla (2014) which threatens to make high-level interventionist explanations too abundant. In conclusion, Weslake identifies a need to address Rescorla's problem through improving our understanding of what makes a causal model an apt representation of a given causal system.

Aptness conditions are also at the core of Vera Hoffmann-Kolss's searching examination of the prospects for interventionist accounts of multilevel explanation. In her chapter, Hoffmann-Kolss identifies a new class of problem for these accounts, a problem which cannot be addressed by existing interventionist responses to Baumgartner's objection. The trouble is that these accounts tend to trivialize the constraint that models represent all relevant confounding factors. The broader lesson drawn by Hoffmann-Kolss is that in order for the causal modelling framework to succeed in application to multilevel explanation, metaphysically 'thick' notions such as grounding, naturalness, and nomological modality are required—notions which the interventionist has traditionally shunned. Hoffmann-Kolss argues however that some such metaphysical commitments are needed to aptly model multilevel causal structures.

Section II is rounded out by David Yates's chapter, which argues that causal closure of the physical (the idea that all causes are physical causes) can come apart from the causal-explanatory closure of fundamental physics (the idea that all causal explanations reside at the level of fundamental physics). Yates takes vector composition as his core example, arguing that any principle of causal closure that is strong enough to entail the causal-explanatory closure of fundamental physics is falsified by vector composition. Conversely, any weaker principle of causal closure which is compatible with vector composition fails to rule out downward causation. This, Yates argues, makes room for a potential causal role for higher-level properties more generally. The chapter concludes by comparing Yates's approach to Marc Lange's hierarchy of necessity for laws and metalaws, which is built on a network of primitive 'subjunctive facts'.

The third section of the volume turns to higher-level sciences and the explanatory levels that they involve. The chapters of Section III deal with three progressively higher-level sciences—biology, cognitive science, and social science. These sciences are typically taken by realists to stand in some relation of functional realization: cognitive agents are typically realized biologically, while social systems typically involve the interaction of multiple cognitive agents.

The chapter by William Bechtel focuses on the life sciences, and in particular on the notion of a control mechanism as it appears in biological explanations within different domains and at different scales. Bechtel argues that there is no simple over-arching hierarchy of levels of control, but that a level-like notion can nonetheless be recovered in particular cases through consideration of specific self-controlled systems and their components. Bechtel is accordingly cautiously positive about the usefulness of a suitably flexible notion of level in biology: levels of mechanisms and levels of control are important for the explanatory practices of biologists and for philosophers understanding those practices, including both 'top-down' and 'bottom-up' causal claims. He emphasizes however that these levels must remain local and contextual, and that they do not lead to any global stratification of 'biological entities'.

Within the cognitive sciences, levels talk has centred around David Marr's three levels: computation, representation, and implementation (Marr (1982)). Mazviita Chirimuuta's chapter considers how analogies with machines and artefacts (most famously the analogy between the brain and a computer) motivate Marr's three levels, and how this motivation fits into a wider anti-reductionist view of explanation in brain and behavioural sciences. Chirimuuta raises some challenges for the analogy between designed artifacts like a radio and evolved systems like the central nervous system. Chirimuuta emphasizes the explanatory importance of these levels: oftentimes a problem is intractable if we start with all the details, but just as analogies might lead us astray by framing problems/phenomena in an overly simplistic way, the levels structure might be a heuristic that is helpful only to a limited extent. Then we run the risk of ending up attributing levels not just to our representations, but to reality itself—thus projecting methodological levels into metaphysical levels.

The final chapter of Section III, by Harold Kincaid, concerns explanatory levels in the social and behavioural sciences. The complexity and messiness of theorizing in social science may suggest that there will be no neat application of any explanatory levels framework; but Kincaid argues that levels can nonetheless be distinguished and used successfully by social scientists. Kincaid emphasizes taking a naturalist 'science-first' approach, one which emphasizes how explanatory levels are a contextual matter. Kincaid argues that macrosociological entities are not epiphenomena by emphasizing their causal roles in a range of different case studies, including some which he argues are benign examples of downwards causation. The result is again a flexible approach to explanatory levels which is motivated by the need to make sense of the full range of scientific practice rather than by any metaphysical precepts.

Section IV turns to explanatory levels in physics. As Eleanor Knox emphasizes in her chapter, connections between levels are often relatively well understood in physics in comparison to other sciences. The tools of mathematics allow us to wield precise control over transitions between physical explanatory levels, as in the important case of renormalization techniques in quantum field theory and condensed matter physics. Oftentimes we know how to construct one physical explanatory level from another in a way that would seem fanciful in the mind/brain case—even to the point where we can claim full-scale reduction (Robertson (2022)).[8] Still, explanatory levels in physics otherwise have much in common with explanatory levels in other sciences. Since hardly any of physics is fundamental physics, as Knox emphasizes, most theorizing in physics takes place at an effective level; many of the same issues then arise for explanatory levels in physics as for levels in biology and other higher-level sciences. This helps cast light on how apparently conflicting physical theories at different levels can be reconciled: McKenzie uses the example of how symmetries which hold at one physical level may be broken at another.

We have already raised a concern about whether explanatory levels might be too easily achieved. In her chapter, Eleanor Knox asks: what makes for a thicker notion of levels? (Or, as she puts it: 'levels worth having'.) Knox emphasizes the contrast between mere levels of description, and the more interesting levels that we come across in physics. In typical cases of levels in physics, the level connection is well understood, but might crucially involve changes of variable to reveal new dependencies. Here her position contrasts with the related approach of Franklin, according to whom explanatory levels in physics are cheap and plenitudinous. Emergence has often been given a two-part treatment: in terms of robustness/autonomy and then in terms of novelty. Much of the discussion in the philosophy of physics has focused on the former—and Knox suggests that to understand the contrast between thin and thick levels, we should turn our attention to novelty. Building on her previous work, Knox proposes that in contrast to the subjective/psychological view of Butterfield (who analyses 'novelty' as 'surprise'), we should understand novelty as novel explanatory value.

Discussions of levels often presuppose the level of fundamental physics as given unproblematically; the truth is far more complicated. Current physics neither offers us a plausible candidate fundamental theory, nor any clear guidance as to what such a theory would look like. Moreover, metaphysics is sometimes prone to usurp the claims of physics to fundamentality: as already discussed, metaphysically

[8] The improved prospects for reduction in physics have led a number of philosophers of physics to the conclusion that reduction and emergence can be compatible (Wilson (2010), Butterfield (2011b), Crowther (2015)). However, to accommodate this compatibility, the relevant notion of emergence will tend to be a weaker one (see, e.g., the distinction between strong and weak emergence in Chalmers (2006)). The explanatory levels paradigm thus ties in with sophisticated recent discussions of emergence in physics such as Franklin and Knox (2018). Still, if the levels involved are too cheap or abundant, then they might not be strong enough to support even a weak notion of emergence.

orientated philosophers of science have explored the prospect of giving metaphysical reductions of physical concepts like 'length' and 'mass'. This raises the question as to whether the levels of explanation of metaphysics and physics can be brought into line. Do the levels of fundamentality acknowledged in recent metaphysics line up with the levels of explanation given by different physical theories? The two chapters by Karen Crowther and Kerry McKenzie tackle this problem head-on—and give opposing answers.

Karen Crowther sees explanatory levels in metaphysics and physics as more in harmony than McKenzie does. Crowther aims to bring metaphysics and physics closer in line, by employing a new definition of 'relative fundamentality' according to which theoretical descriptions with broader scope can count as more fundamental than theoretical descriptions which can be derived from them in specific domains. This new definition aims to shift the emphasis in the metaphysics of physics from questions about the mereological structure of levels to questions about the relations of reduction and constructability that hold between different theoretical descriptions. Crowther argues that reduction relations between theories operating at different scales are often thought to have ontological import, in particular she holds that derivability can be understood as natural, or, indeed, ontological dependence. But the radical part of her argument is that we should also think of reduction relations between old and new theories as similarly having ontological import. This is a version of an idea which has gained traction in recent years, sometimes under the guise of 'effective realism' (Williams (2019), Egg (2021), Saatsi (2022), Robertson and Wilson (forthcoming)).

McKenzie also focuses on the nature of interlevel relationships in theoretical physics, with a sustained critique of applications of the metaphysics of grounding to the topic. McKenzie identifies a tension between the approximation inherent in interlevel relations in physics and the exactness presupposed by standard metaphysical accounts of interlevel grounding. This, she argues, undermines any attempt to use the levelled structure of scientific explanations to support a metaphysically heavy-duty account of levels of explanation in grounding terms. The key problem is that the approximations involved in understanding inter-theoretic relations end up being interest relative, and hence not completely objective in the way that grounding is meant to be. McKenzie's lead example is proton decay—but it can be extended to applications of the effective field theory framework more generally. The tension McKenzie identifies is exacerbated since effective field theories are sometimes regarded as helping to explain *why* we have different levels in the first place; this latter type of question is revisited in Section VI.

Section V broadens discussions of explanatory levels to new questions and domains within mathematics and metaphysics. As McKenzie's chapter already highlights, there is a kind of motivation present here that is at least partly separate from the hierarchy of different sciences. The centrality of fundamentality in metaphysics and of axiomaticity in mathematics build the idea of one thing being at a more basic level than another right into the self-conceptions of the relevant fields. The chapters in this section contrast the multiplicity of kinds of explanations found in mathematics and in metaphysics with the classical model of explanatory levels.

In their chapter, Carolin Antos and Mark Colyvan explore cases where different styles of mathematically explanatory proof exist for a given result. Taking Fermat's Little Theorem and Descriptive Set Theory as examples, Antos and Colyvan show that different styles of proof of a single result can exhibit different balances of theoretical virtues, with mathematicians' intuitions about explanatory power favouring more general styles of proof in some cases and more local styles of proof in others. This verdict reinforces the volume's theme that explanatory levels can be useful even when applied piecemeal and in the absence of any global ordering.

The next chapter turns to explanatory levels within metaphysics. Elanor Taylor raises a worrying objection to recent accounts of dispositions as grounded in lower-level metaphysical facts. Taylor argues that these explanations look very similar to 'dormitive-virtue' style explanations, the archetypal example of unsuccessful explanation. But what exactly is wrong with dormitive-virtue explanations? Taylor teases apart the good from the bad in these explanations, and argues that the dispositionalist's explanations can be partially vindicated by appeal to a 'backing' model of explanation—a model which appeals to different dependence relations in different explanatory contexts within metaphysics.

Completing Section V is Nina Emery's chapter, which criticizes existing accounts of higher-level chances as unable to do justice to their explanatory power. Emery then offers a new account of the nature of higher-level chances with better prospects of vindicating the explanatory role which leads us to postulate those chances in the first place. The proposal does however present a challenge for Humean accounts of the role of chance in a deterministic world: the specific way in which Humeans take chances to depend on the mosaic of fundamental facts compromises the chances' ability to play the explanatory role they are assigned.

The volume thus far has discussed what explanatory levels might look like, and what use they might be in a variety of different settings. Even the most sceptical about levels tend to agree that the different sciences can investigate the world independently of one another, to some extent at least. Yet it seems like a possibility that the world could have been different in this respect. It apparently could have been the case that geological questions couldn't after all be answered without in-depth reference to quantum physics, or that no useful social science generalization could be formulated without delving into the details of individual psychology. So: why are there distinct levels of explanation in the first place?

The final section, VI, takes a step back and asks about the pre-conditions for levels of explanation. This returns us to the volume's third general theme: levels of explanation for levels of explanation. A foundational challenge, yet one which is rarely confronted head on, is to explain why there are any high-level regularities at all. As we discussed above, Jerry Fodor (1974, 1997) posed this question, but there is no consensus about what the answer should be. In the first chapter of this section, Harjit Bhogal offers a distinctive new solution to Fodor's challenge which draws in part on our pragmatic interests and projects as scientific reasoners. Bhogal points out that Fodor's challenge presupposes an identification of certain high-level properties as natural properties, and then shows that the existence of higher-level regularities and of the corresponding explanations falls straight out of some plausible contemporary accounts of natural properties—in particular, out of Bhogal's own proposal and out of an influential recent account by Barry Loewer (2020). Bhogal concludes by contrasting two complementary levels of explanation of high-level regularities: a 'bottom-up' scientific strategy in terms of the specific features of the low-level realizers, and a 'top-down' metaphysical strategy in terms of the nature of properties.

In the following chapter, Michael Strevens offers a general version of Bhogal's bottom-up strategy, a version which highlights the core role of stochastic independence between variables at different levels in underwriting high-level explanations. Strevens outlines a schematic theory of the possibility of high-level explanation in terms of *semi-detachment* of variables at different levels, a theory which is compatible with the various ways in which these variables are causally entangled. In emphasizing independence between variables at different levels, Strevens's discussion dovetails with Knox's discussion of levels in physics. But why is semi-detachment so common? Strevens next turns to the possibility of providing a top-down explanation of semi-detachment itself, in terms of the propensity of the details of lower-level systems to cancel or balance one another out. He argues that this propensity to cancel out explains why high-level explanations in the sciences are so widespread.

A full account of the explanatory power of higher-level sciences not only needs to address the properties and regularities as Bhogal does, and to demonstrate their relationship to different levels as Strevens does; it also needs to encompass the explanatory role of the related higher-level laws—and, as we saw in Emery's chapter in the previous section, higher-level chances. Although best-system accounts of laws in the tradition of Mill, Ramsey, and Lewis are deservedly popular, they tend to face difficulties in accounting for laws other than those which hold at the fundamental level, and in making sense of the explanatory role of chance. Higher-level laws are a central concern of explanatory levels; for those who hold that explanation hinges on laws, including proponents of the D-N model and its descendants,[9] without higher-level laws there would be no distinctive higher-level explanations. The law-like status of regularities in the higher-level sciences is central to the Fodor-Oppenheim-Putnam debate and the genesis of explanatory levels; one of

9 See Woodward and Ross (2021) for a survey.

Fodor's key objections to the reducibility of the higher level stems from whether the lower level can explain/account for lawhood at the higher level.[10]

This theme is taken up in the final chapter of the volume, in which Michael Townsen Hicks offers a new and improved version of the popular 'better best system' analysis of laws at higher levels. Craig Callender and Jonathan Cohen (2009), and separately Markus Schrenk (2014), have attempted to characterize special-science laws by modifying the best-system approach so as to apply at multiple levels. Hicks improves on these proposals by accommodating systematic interlevel relationships without undermining the explanatory power of higher-level laws: the key is to assess a candidate law's informativeness in terms of facts at multiple different levels of interest. In this way, Hicks offers a *big* better best system that he terms the 'democratic view' of laws. This democratic view treads a fine line between those that hold that the higher-level laws are independent (anarchists) and those that hold that the more fundamental laws are responsible not only for the regularities of higher-level laws but also for their lawhood status (imperialists). Imperialists require too strong a connection between higher and lower levels, one which leaves features like the counterfactual robustness of higher-level laws unexplained: they have no account of how the economic law of supply and demand would still have held even with a different periodic table of the elements. On the other hand, anarchists have no good account of how we achieve detailed understanding of interlevel relations in many cases; the need to account for this epistemic success was particularly emphasized throughout Section V, as well as in Franklin's discussion of reductive explanations.

Taken together, the chapters of Section VI give us insight into what underlying aspects of our world allow the different levels of explanation to be possible in the first place. They thereby offer explanations at various different levels for the multifaceted phenomenon of levels of explanation.

References

Andersen, Holly (2014a). "A Field Guide to Mechanisms: Part I," *Philosophy Compass* 4: 274–283.

Andersen, Holly (2014b). "A Field Guide to Mechanisms: Part II," *Philosophy Compass* 4: 283–297.

Batterman, Robert W. (2013). "The Tyranny of Scales." In *The Oxford Handbook of Philosophy of Physics*, ed. Robert W. Batterman. Oxford University Press, pp. 256–286.

Batterman, Robert W. (2020). "Multiscale Modeling: Explanation and Emergence." In *Methodological Prospects for Scientific Research*, ed. Wenceslao J. Gonzalez. Springer, pp. 53–65.

Batterman, Robert W. (2021). *A Middle Way: A Non-Fundamental Approach to Many-Body Physics*. Oxford: Oxford University Press.

Baumgartner, Michael (2009). "Interventionist Causal Exclusion and Non-Reductive Physicalism," *International Studies in the Philosophy of Science* 23(2): 161–178.

[10] See Sober (1999) for a response.

Bechtel, William (1994). “Levels of Description and Explanation in Cognitive Science,” *Minds and Machines* 4(1): 1–25.

Bechtel, William (2008). *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge.

Bennett, Karen (2017). *Making Things Up*. New York: Oxford University Press.

Bliss, Ricki, and Kelly Trogdon (Winter 2021 Edition). “Metaphysical Grounding.” In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University, available at https://plato.stanford.edu/archives/win2021/entries/grounding/.

Brooks, Daniel Stephen, James DiFrisco, and William C. Wimsatt (eds.) (2021). *Levels of Organization in the Biological Sciences*. Cambridge: MIT Press.

Bryant, A. (2018). “Naturalizing Grounding: How Theories of Ground Can Engage Science,” *Philosophy Compass* 13(5): e12489.

Bursten, Julia R. (2018). “Conceptual Strategies and Inter-theory Relations: The Case of Nanoscale Cracks,” Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics 62: 158–165.

Butterfield, Jeremy (2011a). “Emergence, Reduction and Supervenience: A Varied Landscape,” *Foundations of Physics* 41(6): 920–959.

Butterfield, Jeremy (2011b). “Less Is Different: Emergence and Reduction Reconciled,” *Foundations of Physics* 41(6): 1065–1135.

Butterfield, Jeremy (2014). “Reduction, Emergence, and Renormalization,” *Journal of Philosophy* 111(1): 5–49.

Callender, Craig, and Jonathan Cohen (2009). “A Better Best System Account of Lawhood,” *Philosophical Studies* 145(1): 1–34.

Cartwright, Nancy (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Castellani, Elena (2000). “Reductionism, Emergence, and Effective Field Theories,” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 33(2): 251–267.

Chalmers, David J. (2006). “Strong and Weak Emergence.” In *The Re-Emergence of Emergence: The Emergentist Hypothesis From Science to Religion*, eds. P. Davies and P. Clayton. Oxford University Press, pp. 244–254.

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

Craver, Carl. F. (2015). “Levels.” In *Open MIND*, eds. T. Metzinger and J. M. Windt. Frankfurt am Main, Germany: MIND Group, pp. 1–26.

Crowther, Karen (2015). “Decoupling Emergence and Reduction in Physics,” *European Journal for Philosophy of Science* 5(3): 419–445.

Dennett, Daniel C. (1991). “Real Patterns,” *Journal of Philosophy* 88(1): 27–51.

Dennett, Daniel C. (2009). “Intentional Systems Theory.” In *The Oxford Handbook of Philosophy of Mind*, eds. Ansgar Beckermann and Brian P. McLaughlin. Oxford University Press, pp. 339–350.

Dupré, John (1993). *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*. Cambridge: Harvard University Press.

Egg, Matthias (2021). “Quantum Ontology Without Speculation,” *European Journal for Philosophy of Science* 11(32).

Emmerson, Nicholas (2022). “Plumbing Metaphysical Explanatory Depth,” Philosophical Studies: 1–22. https://doi.org/10.1007/s11098-022-01886-3

Eronen, Markus I. (2015). “Levels of Organization: A Deflationary Account,” *Biology and Philosophy* 30(1): 39–58.

Felline, Laura (2021). “Mechanistic Explanation in Physics”. In *The Routledge Companion to the Philosophy of Physics*, eds. Eleanor Knox and Alastair Wilson. Routledge, pp. 476–486.

Fine, Kit (2012). “Guide to Ground.” In *Metaphysical Grounding*, eds. Fabrice Correia and Benjamin Schneider. Cambridge University Press, pp. 37–80.

Fodor, Jerry (1974). “Special Sciences (or: The Disunity of Science as a Working Hypothesis),” *Synthese* 28(2): 97–115.

Fodor, Jerry (1997). "Special Sciences: Still Autonomous after all These Years," *Philosophical Perspectives* 11: 149–63.
Franklin, Alexander (2018). "Whence the Effectiveness of Effective Field Theories?," *British Journal for the Philosophy of Science* 71(4): 1235–1259.
Franklin, Alexander, and Eleanor Knox (2018). "Emergence Without Limits: The Case of Phonons," *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 64: 68–78.
Halpern, Joshua Y., and Judea Pearl (2005). "Causes and Explanations: A Structural-Model Approach. Part I: Causes," *British Journal for the Philosophy of Science* 56(4): 843–887.
Guay, Alexandre, and Olivier Sartenaer (2016). "A New Look at Emergence. Or When After Is Different," *European Journal for Philosophy of Science* 6(2): 297–322.
Hempel, Carl Gustav (1965). *Aspects of Scientific Explanation* (Vol. 1). New York: Free Press.
Hitchcock, Christopher, and James Woodward (2003). "Explanatory Generalizations, Part II: Plumbing Explanatory Depth," *Noûs* 37(2): 181–199.
Humphreys, Paul (2016). *Emergence: A Philosophical Account*. New York: Oxford University Press.
Kim, Jaegwon (1993). "The Nonreductivist's Troubles with Mental Causation." In *Supervenience and Mind: Selected Philosophical Essays*, ed. Jaegwon Kim. Cambridge: Cambridge University Press, pp. 336–357.
Kim, Jaegwon (2002). "The Layered Model,"
Knox, Eleanor (2016). "Abstraction and its Limits: Finding Space for Novel Explanation," *Noûs* 50(1): 41–6.
List, Christian (2019). "Levels: Descriptive, Explanatory, and Ontological," *Noûs* 53(4): 852–883.
Lewis, David (1988a). "Statements Partly About Observation," *Philosophical Papers* 17: 1–31.
Lewis, David (1988b). "Relevant Implication," *Theoria* 54(3): 161–174.
Loewer, Barry (2001). "From Physics to Physicalism". In *Physicalism and its Discontents*, eds. Carl Gillett and Barry Loewer. Cambridge: Cambridge University Press, pp. 37–56.
Loewer, Barry (2008). "Why There Is Anything Other Than Physics." In *Being Reduced: New Essays on Reduction, Explanation, and Causation*, eds. J. Hohwy and J. Kallestrup. Oxford: Oxford University Press, pp. 149–163.
Loewer, Barry (2020). "The Package Deal Account of Laws and Properties," *Synthese* 199(1–2): 1065–1089.
Nagel, Ernest (1935). "The Logic of Reduction in the Sciences," *Erkenntnis* 5: 46–52.
Nagel, Ernest (1949). "The Meaning of Reduction in the Natural Sciences." In *Science and Civilization*, ed. R. C. Stouffer. Madison: University of Wisconsin Press, pp. 99–135.
Nagel, Ernest (1961). *The Structure of Science. Problems in the Logic of Explanation*. New York: Harcourt, Brace & World, Inc.
Nagel, Ernest (1970). "Issues in the Logic of Reductive Explanations." In *Mind, Science, and History*, eds. H. E. Kiefer and K. M. Munitz. Albany: SUNY Press, pp. 117–137.
Nickles, Thomas (1973). "Two Concepts of Intertheoretic Reduction," *Journal of Philosophy* 70(7): 181–201.
Marr, David (1982). *Vision*. W. H. Freeman: San Francisco.
Mundy, Brett (1987). "The Metaphysics of Quantity," *Philosophical Studies* 51(1): 29–54.
Oppenheim, Paul, and Hilary Putnam (1958). "Unity of Science as a Working Hypothesis," *Minnesota Studies in the Philosophy of Science* 2: 3–36.
Owens, David (1989). "Levels of Explanation," *Mind* 98(389): 59–79.
Pearl, Judea (2000). *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University.
Perry, Zee (forthcoming). "On Mereology and Metricality," to appear in *Philosophers' Imprint*.
Robertson, Katie (2022). "In Search of the Holy Grail: How to Reduce the Second Law of Thermodynamics," *British Journal for the Philosophy of Science* 73(4): 987–1020.
Robertson, Katie, and Alastair Wilson (forthcoming). "Theoretical Relicts: Progress, Reduction and Autonomy," *British Journal for Philosophy of Science* 77, available at https://doi.org/10.1086/724445.

Rosaler, Joshua (2015). "Local Reduction in Physics," *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 50: 54–69.
Rosaler, Joshua (2019). "Reduction As an A Posteriori Relation," *British Journal for the Philosophy of Science* 70(1): 269–299.
Saatsi, Juha (2022). "(In)effective Realism?," *European Journal for Philosophy of Science* 12: 30.
Schaffner, Kenneth (1967). "Approaches to Reduction," *Philosophy of Science* 34: 137–147.
Schrenk, Markus (2014). "Better Best Systems and the Issue of CP-Laws," *Erkenntnis* 79(10): 1787–1799.
Sober, Elliott (1999). "The Multiple Realizability Argument Against Reductionism," *Philosophy of Science* 66(4): 542–564.
Strevens, Michael (2008). *Depth: An Account of Scientific Explanation*. Cambridge: Harvard University Press.
Weslake, Brad (2010). "Explanatory Depth," *Philosophy of Science* 77(2): 273–294.
Wallace, David (2019). "Naturalness and Emergence," *The Monist* 102(4): 499–524.
Williams, Porter. (2019). "Scientific Realism Made Effective," *British Journal for the Philosophy of Science* 70(1): 209–237.
Wilson, Alastair (2018). "Metaphysical Causation," *Noûs* 52(4): 723–751.
Wilson, Jessica (2010). "Non-Reductive Physicalism and Degrees of Freedom," *British Journal for the Philosophy of Science* 61(2): 279–311.
Wilson, Mark (2010). "Mixed-Level Explanation," *Philosophy of Science* 77(5): 933–946.
Wilson, Mark (2017). *Physics Avoidance: Essays in Conceptual Strategy*. Oxford University Press.
Wimsatt, William C. (1972). "Complexity and Organization." In *PSA: Proceedings of the Biennial meeting of the Philosophy of Science Association*, eds. K. F. Schaffner and R. S. Cohen. Dordrecht Reidel, pp. 67–86.
Wimsatt, William C. (1976). "Reductionism, Levels of Organization, and the Mind-Body Problem." In *Consciousness and the Brain*, ed. Gordon G. Globus. Plenum Press, pp. 205–267.
Wimsatt, William C. (2007). *Re-Engineering Philosophy for Limited Beings*. Cambridge MA: Harvard University Press.
Woodward, James (2003). *Making Things Happen: A Theory of Causal Explanation*. New York: Oxford University Press.
Woodward, James (2021). "Explanatory Autonomy: The Role of Proportionality, Stability, and Conditional Irrelevance," *Synthese* 198: 237–265.
Woodward, James, and Lauren Ross (Summer 2021 Edition). "Scientific Explanation." In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Stanford, CA: Stanford University. Available at: https://plato.stanford.edu/archives/sum2021/entries/scientific-explanation/
Yablo, Stephen (1992a). "Mental Causation," *Philosophical Review* 101: 245–280.
Yablo, Stephen (2014). *Aboutness*. Princeton: Princeton University Press.