Partial Understanding

Idealization and the Aims of Science

Angela Potochnik

Draft: Please don't cite

Note: My talk at MS6 ("Idealization and the Limits of Science") will be based on material from two chapters of a book manuscript that I am currently working on. I have included drafts of those two chapters here. My talk will draw from §2.2, §3.2, and §3.3.

These chapters are at an early stage of development, and I welcome any comments. If you have any feedback you are willing to provide directly, please email me at angela.potochnik@uc.edu.

Thank you!

Chapter 2

Complex Causality and Simplified Representation

In this chapter I establish the most basic consequence of science performed by humans in a complex world, namely, widespread idealization. As a first step, I make the case that a significant feature of science is the search for causal patterns. These patterns qualify as causal on a human-centric interventionist approach to causation. By and large, they hold over a limited range of circumstances and have exceptions even within the range over which they generally hold. Moreover, whether a causal pattern emerges depends on our representational choices. I argue that there is no need to posit a straightforward relationship between these causal patterns and any metaphysically basic causal processes. This discussion grounds a more specific characterization of the relevant sense in which the world is complex, which I term causal complexity. I argue that causal complexity is a pervasive feature of the world and that it significantly impacts scientific practice.

In the second half of the chapter, my attention turns from the scientific significance of causal patterns and causal complexity to widespread scientific strategies for discerning patterns in the face of complexity. I first discuss model-based science, a common scientific practice that has lately received increasing philosophical attention and that naturally leads to an emphasis on idealization. Then I further explore the role idealization plays in science. I argue that this role is much more expansive than has been appreciated. Idealizations are widespread even outside of model-based science, and there are many intertwined reasons to idealize. Moreover, I develop the view that idealizations actually play a direct representational role. As a result, the practice of idealization does not fit within expected bounds. Idealizations are both pervasive and ineliminable in science.

2.1 Causal Patterns in the Face of Complexity

2.1.1 Causal Patterns

In Chapter One I discussed the shift away from accounting for an eventual, perfected science or a prettified, more rational version of science toward instead accounting for today's actual scientific practice, messiness and all. This shift has been accompanied by a great deal of attention to ways in which science falls short of discovering laws, understood as exceptionless regularities that are universal in scope. One influential critic of universal laws is Nancy Cartwright (1983), who argues that physics' most fundamental laws are in fact false. In Cartwright's view, those laws are successful, for they are explanatorily powerful, but they do not accurately describe nearly any actual systems. Cartwright gives as an example Newton's law of universal gravitation, which states that any two bodies attract each other with a force directly proportional to the product of their masses and inversely proportional to the square of the distance between them:

$$F = G\frac{m_1 m_2}{r^2}$$

Here F is the force, G the gravitational constant, m_1 and m_2 the two masses, and r the distance between the centers of those masses. As Cartwright points out, the physicist Richard Feynman called this law "the greatest generalization achieved by the human mind" (1967, p.14). Yet the law of universal gravitation ignores the influence of charge on the two bodies in question, and so the stated relationship is false of any systems with charged bodies. It also assumes that the mass of bodies is concentrated in a single point, that is, it applies to "point masses." This assumption is, of course, false of every body. It does accurately reflect the behavior of spatially extended masses, but only for bodies that are spherically symmetrical. Cartwright argues that this is so not just of Newton's law of universal gravitation, but also of all our other best law-like generalizations. It is not that these scientific laws occasionally have exceptions, but that they do not hold true for nearly *any* real systems. They are, strictly speaking, false.

Considerations like those Cartwright introduces have resulted in laws, taken as universal, exceptionless regularities, becoming less central to philosophical investigations of science. Ronald Giere (1999) also argues that so-called scientific laws are not true. However, rather than maintain with Cartwright that the nature of these laws must be reinterpreted, Giere instead concludes that there simply are no laws of nature. Other philosophers investigate the status of ceteris paribus laws. "Ceteris paribus" literally means "all other things being equal." Ceteris paribus laws have been embraced especially in what are often called the "special sciences" or "inexact sciences," viz., those fields of science other than fundamental physics. Though Cartwright focuses on fundamental physics, the absence of universal laws is more broadly appreciated in the so-called special sciences. More generally, philosophical attention has shifted significantly from the search for scientific laws and their subsequent application toward other activities of science, including modeling, causal analyses, and mechanism sketches. The covering-law approach to explanation due to Hempel and Oppenheim (1948) has been replaced by a range of other approaches. Machamer et al. (2000), in a seminal paper in the mechanisms literature, note that laws of nature have little if any relevance to their fields of focus, which are neurobiology and molecular biology. All of this suggests that, if discovering bonafide, exceptionless laws of nature plays any role in science, it is a very small role at best. Instead, there is increasing an acknowledged diversity of scientific projects, none of which lead to the discovery of exceptionless laws.

There is, however, a different type of characterization that better applies to these diverse scientific projects. Instead of universal laws, science—by and large—depicts or otherwise capitalizes on *causal patterns*. Let me first discuss what I mean by "patterns" before moving on to address the sense in which these patterns are causal. I have spoken as if laws are supposed to be universal regularities in phenomena, about which we humans make generalizations. I suspect there is often ambiguity about which is the law—the regularity or the generalization made about it. Such ambiguity is of little consequence if science accurately depicts (or provides) universal laws of nature. It becomes problematic, though, when such a construal is questioned, and the truth or reality of laws along with it. If science is supposed to provide insight into laws of nature, then following Giere those laws are not real; if scientific laws are our generalizations about the world, then following Cartwright those laws are not true. I intend causal patterns to be a successor to laws of nature, taken as regularities in phenomena themselves. Patterns are thus not human representations, but are depicted by our representations. These regularities are patterns, and not laws, for two reasons: they hold only in a limited range of circumstances, and even within that range, there are deviations from them and exceptions to them.

These limitations are reflected in Giere's (1999) suggestion that what are traditionally understood as universal laws are instead "restricted generalizations." Giere points out that to productively employ Newton's law of universal gravitation, one must explicitly restrict it to certain kinds of systems, and even then, one must often employ approximation techniques before the systems appear to embody the posited relationship. Consider, first, that such generalizations have a restricted domain of application. This is because the patterns they depict hold only in a limited range of circumstances. This is also exemplified by the ideal gas law, $PV = nRT$, where P is the pressure of a gas, V its volume, n its number of molecules measured in moles, T its temperature, and R the ideal gas constant. The pattern represented in this generalization only holds in a limited set of systems; it does not obtain at very low temperatures and at very high pressures. Appending a ceteris paribus clause to a generalization is one way that is used to acknowledge this kind of limitation, especially when the exact circumstances in which a pattern obtains is unclear. Consider, second, that generalizations can also have limited accuracy even within their domain of application. This is due to deviations from the patterns they depict. The ideal gas law illustrates this too, for it ignores both molecular size and intermolecular attraction, and these factors can diminish its accuracy. It is possible to introduce correction terms, which yields the van der Waals equation, but the original form of the ideal gas law is often judged to be sufficient. When deviations from a pattern are great enough, this results in an exception to the pattern, viz., a system that we would expect to embody a pattern failing to do so.

What holds for Newton's universal gravitation and the ideal gas law holds for the patterns identified in other fields of science as well. Most patterns hold only in limited circumstances, and most also have deviations and exceptions even within those circumstances. The idea that scientific generalizations depict these limited patterns is a natural refinement of the idea that they depict universal laws of nature, once one takes seriously that our science is the product of limited human faculties and concerns, grappling with a world as complex as ours. Powerful generalizations such as Newton's law of gravitation that Feynman praised would be lost if the standards for laws of nature, including universality and exceptionlessness, were upheld. Cartwright is right that these "laws" are of great scientific significance. But Giere is right that they are not, after all, laws. Instead, they are members of a broader class of scientific products, a class that also includes generalizations with explicit ceteris paribus clauses, and many more (as we will see in the second half of this chapter). The members of this class depict patterns—limited patterns with exceptions. The simplicity and straightforwardness of a pattern at once increase its usefulness for limited humans and decrease its universality in a world that is neither simple nor straightforward. This is so for fundamental physics as much as biology and sociology. Physics is the product of limited human beings as much as any other field of science, and physical phenomena outside of the laboratory are just as complex.[1]

This conception of patterns can be further elucidated by considering what Daniel **?** calls "real patterns." Dennett's main focus is the ontological status of folk psychological states like belief, but his discussion of patterns yields considerable insight here. The discussion regards when a pattern exists in data, but it can be adapted to apply instead to when a pattern is exhibited by a *set of systems*. This is the proper sort of pattern to consider here because causal patterns, as a successor to laws of nature, are supposed to be regularities exhibited by phenomena themselves. Consider figure 2.1. This is Dennett's simple example of a pattern, which he dubs "bar code." He takes the objects to be data sets, but on my alternative construal, they together constitute a set of systems (or phenomena). Notice, first, that the pattern bar code might be present with different degrees of noise, or deviations from the pattern. In these six objects, the level of deviation varies from 1% to 50% of the dots comprising each object. Nonetheless, A through E (at least) all embody the pattern bar code.

One might exactly describe any one of these systems by first describing the pattern of 10 rows of

[1]For a defense of this last idea, please see (Cartwright, 1983; Giere, 1999; Kennedy, 2012).

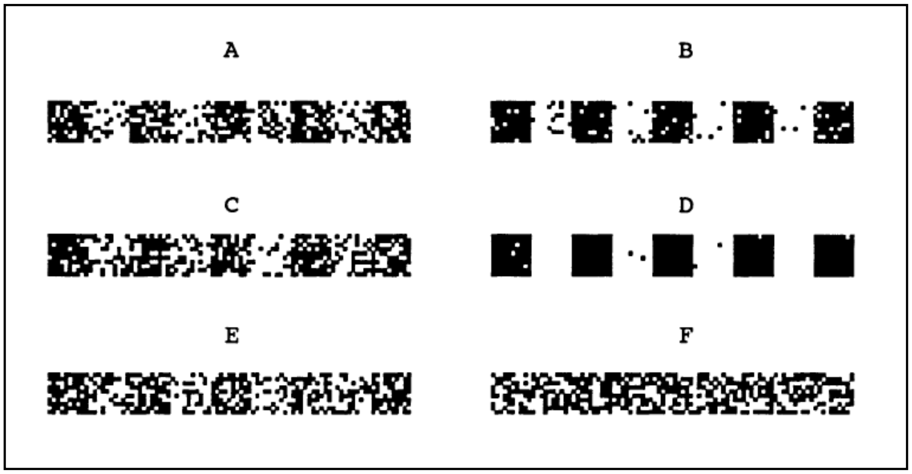

Figure 2.1: **?** uses these objects to articulate his conception of a pattern. There is a pattern that Dennett calls "bar code" apparent in A through E, along with a varying amount of "noise." For my use of Dennett's treatment of patterns, these objects should be taken to exemplify systems, that is, phenomena. If the noise is great enough, this leads to an exception to the pattern. This may be the case in F.

90 dots, each row a repetition of 10 black dots then 10 white dots, and then detailing the individual dots that vary from that algorithm. However, as Dennett says, "for some purposes, we need not list the exceptions to bar code, but only transmit the information that the pattern is bar code with n% noise" (p.34). That yields an identical description for A and C, for these each have 25% noise. I would go one step further. For some purposes, we need not even specify the *amount* of deviation. Sometimes all that matters to us is whether this is the sort of system that embodies the pattern in question. In such circumstances, A through E (at least) have an identical description. They can be described simply as bar code. This is the status of our scientific "laws": they depict a pattern that is instantiated by a range of systems generally without indicating how much noise, or deviation, we should expect. Oftentimes approximation techniques are required to accommodate a system's deviations from the pattern. Mistakenly construing these patterns as universal laws of nature occurs when, instead of ignoring the deviations, it is posited that there are in fact no deviations, or that any deviations are insignificant.

The status of F is less clear. Dennett says that, if F embodies the bar code pattern it is indiscernible, and the notion of an indiscernible pattern is self-contradictory. Yet F *is* patterned in the sense that it was produced by the same algorithm as A-E; it simply has the highest level of noise, at 50%. F thus belongs in this set of systems in virtue of its similarities with the other systems. The question is whether the degree to which it deviates from the pattern is great enough for it to constitute an exception to the pattern. This leads to another insight of Dennett's discussion. Dennett points out that two individuals may perceive different patterns in the same data, in virtue of the individuals' different interests and perspectives. So too might different individuals, with different interests and perspectives, judge a system to embody different patterns, or even to embody a pattern or be an exception to that pattern, without contradiction. Many systems confronted in science embody multiple patterns (with varying degrees of deviation). Which pattern is salient or focal, and how much deviation is tolerable, depends on our needs and interests. Whether the ideal gas law or the van der Waals equation is better depends not only on features of the gas in question, but also on the purpose to which we will put our description of the gas.

* * *

Laws of nature were supposed to be universal, exceptionless regularities. Though patterns are neither universal nor exceptionless, they are regularities. Indeed, the patterns sought in science are regularities in dependence relations. The law of universal gravitation shows how the force between two bodies depends on their mass and the distance between them. The ideal gas law shows the interdependence among a gas's pressure, volume, and temperature, as well as how these depend on the amount of gas.[2] This should come as no surprise, for our science is aimed at facilitating human action and human comprehension of our world, and uncovering predictable relationships of dependence is key to both of these goals. Further, when an account of causation is settled upon that is well-tailored to our scientific enterprise, it emerges that these patterns of dependence are in fact *causal* patterns.

Both Cartwright (1983, 1989) and James Woodward (2003) develop accounts of causation that put human action center stage. Cartwright emphasizes that causal regularities must be posited in order to enact strategies effective at producing desired results. For Woodward, the practical utility of causal knowledge in manipulation and control is a key motivation for his manipulability

[2]Though I have argued there are no laws, I will continue to use the word "law" where it is part of the conventional terminology for particular scientific results.

account of causation.[3] This strategy of grounding causal analysis on action, and specifically human action, is appropriate for a science that is the product of humans. As a result, we should expect the sort of causal information uncovered by science to be grounded in its application to objects and types of objects observed by humans, and to be useful in manipulations in the circumstances faced by humans. One of many differences between Cartwright's and Woodward's approaches is the role accorded to singular causation versus causal regularities, or as Woodward puts it, token-causal and type-causal claims. Cartwright's skepticism about laws leads her to posit singular causes as basic, from which causal capacities derive. In contrast, Woodward stresses the scientific significance—and significance to manipulation—of causal relationships between variables.[4] On his view, understanding of singular causal relationships derives from these. As indicated above, my attention to patterns leads naturally to a focus on sets of systems. For that reason, I follow Woodward in focusing on regularities across systems, and thus type-causal claims.

Indeed, I adopt several of the insights of Woodward's manipulability account of causation. The two core concepts of that approach are intervention and invariance. Suppose you are ascertaining whether one variable, call it X, causally influences another variable Y. An *intervention* on X changes its value surgically, so to speak, so that any change induced in Y is due solely to the change to X, not to any side effects of the intervention. At various places Woodward likens his concept of intervention to an "ideal experimental manipulation." For Woodward, X is a cause of Y just in case some interventions on X change Y's value, in some background circumstances. Interventions, or manipulability relations, are thus the ultimate guide to causal relations. This is as it should be for a concept of causation employed in a science built from a human perspective to facilitate human action. Then, if X is a cause of Y, that causal relationship is *invariant* over some interventions and range of background circumstances. That is, the causal relationship—how interventions on X change Y—would continue to hold in those circumstances. Invariance comes in degrees. Causal relationships all require some amount of invariance, for how changes to X affect Y must at least be stable under *some* intervention in some circumstance. But a causal relationship may be stable

[3] Woodward's account is one recent and especially popular version of a manipulability or interventionist theory of causation.

[4] In this use, variables are "properties or magnitudes... capable of taking more than one value" (Woodward, 2003, p.39).

over larger or smaller ranges of interventions and background circumstances. Invariance is key to formulating generalizations about causal relationships, or put another way, to identifying causal patterns.

Science's reach extends beyond human powers of manipulation. Scientific practices, of course, involve much more than experimental manipulations: they also include simple observation, various means of investigating past events, and as we will see, many practices more distant from any form of data collection. It is thus appropriate that Woodward's account does not require an actual intervention be performed in order for a causal relationship to exist, or even be ascertained. The relationship between intervention and causation is instead conceptual (and, perhaps, metaphysical). Indeed, Woodward claims that there may be causal relationships even in cases when an intervention is not physically possible. He argues that this can be so when a surgical change to a variable's value can be assured by means other than the requirements he sets out for an intervention. Though rooted in facts about manipulation, the causal patterns ascertained in science go beyond the directly manipulable.

Woodward's manipulability approach to causation helps overcome one immediate stumbling block for the idea that apparent laws actually depict causal patterns. The stumbling block is that many of those patterns may not appear to be causal in nature. Consider again the ideal gas law, $PV = nRT$. This depicts a synchronic dependence among several features of a gas; it is not patently a depiction of causal dependence. Indeed, Wesley Salmon (1984) argues that the ideal gas law does not represent a causal relationship. Similarly, other equilibrium models, including optimality models and game theoretic models in evolutionary biology, depict structural features that together determine an equilibrium point. These have also been construed as non-causal (Sober, 1983; Rice, forthcoming). However, on Woodward's conception of causation, applications of both qualify as causal. In a variety of circumstances, an intervention on the volume of a container would change the pressure of the gas inside according to the relationship expressed by the ideal gas law. Intervening on the factors determining an equilibrium point disrupts the expected equilibrium by changing the value of the equilibrium point or eliminating it entirely. For example, Goss-Custard (1977) develops an optimality model to apply to the eating habits of the redshank sandpiper (*Tringa* *totanus*), a bird that feeds on worms in mudflats. This sandpiper exhibits a preference for eating large worms over small worms. The model demonstrates that, if large worms and small worms are both readily available, a redshank's energy intake is maximized when large worms are chosen. This selective advantage would lead to the evolution of the preference in question. But if, for example, large worms had been historically more difficult to find in the bird's environment as it evolved (an intervention), the preference—or at least the degree of preference—would be different.

A prominent alternative to a manipulability account of causation instead emphasizes physical causal processes. One example of this is Salmon's (1984) account of causation based on mark transmission. On that view, a process is causal if it is capable of carrying forward a modification or mark. A second example is Dowe's (2000) account, according to which a causal process is a world line possessing a conserved quantity. A world line is an object's trajectory through space-time, and conserved quantities are whatever our best scientific theories tell us is universally conserved, e.g., mass-energy. In each case, the concept of a causal process, together with causal interaction, is supposed to provide the basis for an account of causation.

These physical process views would judge the patterns depicted by the ideal gas law and other equilibrium models—and many others patterns uncovered in science—as distantly related to causal relationships, if related at all. This is why, as we saw above, Salmon judges the ideal gas law to be non-causal. But this obscures an insight that a manipulability approach to causation provides. That insight is that there are patterns of dependence in our world, dependences that we humans are sometimes in the position to exploit. These patterns of dependence are, accordingly, crucial to human action and insight. This is why a manipulability account of causation is scientifically useful. By obscuring the relationship between causal dependence and human powers of manipulation, physical process accounts of causation err in the same way as accounts of science that emphasize universal laws of nature. Both fail to accommodate how the features of science are shaped by its status as a human enterprise. Notice that I have not compared the merits of the manipulability account and physical process accounts as rival approaches to the metaphysics of causation. I simply laud the manipulability approach's scientific serviceability. I briefly address the relationship between this discussion and the metaphysics of causation just below. For now, any who are distracted by my adherence to a manipulability approach to causation are free to simply rename what I call causal patterns "manipulability patterns," and set the matter of causation aside.

Thus, the ideal gas law, universal gravitation, and the like do not depict universal laws of nature, but patterns instantiated by a limited range of systems to varying degrees, and with exceptions.[5] I need to refine this proposed successor for laws of nature in one more way before moving on. I said above that science depicts or otherwise capitalizes on causal patterns, but so far I have only discussed scientific products that depict patterns. It will emerge in the next chapter that science has a wide variety of aims. Accordingly, scientific projects make use of causal patterns in different ways. Many of these uses are naturally construed as depicting or representing patterns. This is the class of products I mentioned above, including apparent laws that are better understood as restricted generalizations, generalizations that explicitly employ ceteris paribus clauses, and a number of other kinds of scientific representations. But other scientific products may resist such an interpretation. The exploration of causal influence can instead involve careful causal diagnosis in one or a few specific phenomena of interest. Other projects are even more distant from causal patterns. The literature on mechanisms shows that other kinds of dependence, such as compositional or organizational dependence, can also be significant. Purely predictive models may have little to do with the representation of causal patterns. Yet despite all this variety, all capitalize in one way or another on causal patterns as I have articulated the concept. As Woodward's approach makes salient, causal patterns are key to action, to exerting influence over our world. This will become increasingly clear as we proceed, and particularly in Chapters Four and Five.

* * *

Causation is, of course, a metaphysically charged topic. There is a question of what (if anything) the causal relation really is. There are at least three positions one might have on the metaphysical status of a manipulability approach to causation. One might take manipulability to be a guide to causal metaphysics; perhaps this is what Woodward has in mind. At any rate, Woodward (2007) expresses doubt that causal claims are grounded in any fundamental

[5] Indeed, Woodward also intends his generalizations about invariant causal relationships to be a successor to laws of nature.

physical causal relationships. Instead, in his view, "macroscopic causal claims (like 'chances' in a deterministic world) reflect complicated truths about an (i) underlying microphysical reality and (ii) the relationship of macroscopic agents and objects to this world" (p.102). This kind of view seems to suggest that "macroscopic causal claims" themselves are the best candidates to provide insight into the metaphysics of causation. Alternatively, one might hold a physical relationship of causation, such as outlined by Salmon (1984) and Dowe (2000), to be metaphysically basic and conceive of facts about manipulability and difference-making as dependent on facts about fundamental causal relationships. Alyssa Ney (2009) endorses this latter view, which she calls "causal foundationalism." Finally, one might hold that neither of these accounts is metaphysically basic. **?**Cartwright (2007) advocates causal pluralism. On that view a variety of criteria are relevant to causal judgments, and how to weigh these criteria depends on the particularities of individual cases.

As far as the present project is concerned, any of these positions on the metaphysics of a manipulability account of causation may be correct. I signaled this above by suggesting that those with allegiances to other theories of causation might translate my "causal patterns" into "manipulability patterns." Note that Woodward also stresses the "unmetaphysical character" of his account in (**?**). This neutrality about the metaphysics of causation exemplifies a position about the relationship between science and metaphysics that will be made explicit in Chapter Three and then broadened in Chapter Four. It also has two significant implications, but before addressing those, let me first discuss a crucial regard in which a manipulability construal of causation must be basic.

Facts about manipulability may not be metaphysically basic, but they are *epistemically* basic. These facts ground our science, providing the basis for the discovery and use of causal patterns. This is reflected in Woodward likening his concept of an intervention to an ideal experimental manipulation. It is the possibility of such manipulations that provides the foundation for our causal reasoning. In contrast, any account of physical causation is the product of our best theories of fundamental physics. These theories, no matter how epistemically secure, require successful scientific reasoning already in place. This is apparent in Ney's discussion of causal foundationalism, for she gives the proviso that "to the extent that today's fundamental physics is true, it provides us with facts about causal relations that obtain at our world" (p.746). This epistemic limitation of any physical account of causation is significant, for as Ney also points out, today's fundamental physics is unlikely to be true, in light of the fact that it is inconsistent. And, in any case, our best fundamental physical theories are scientific products like all the rest. Any epistemic security they might possess is grounded in their epistemic basis in scientific reasoning, viz., causal reasoning based on facts about manipulability.[6]

The first implication of distancing a manipulability conception of causation from causal metaphysics is that this obviates a potentially unsettling metaphysical commitment. I have argued that the causal patterns uncovered in science hold only approximately and in limited circumstances. Allowing for the possibility of distance between these causal patterns and metaphysically basic causal relationships enables one to avoid committing to a world in which metaphysically basic causal relationships hold only approximately and in limited circumstances. Our world may indeed be law-governed; these laws may or may not be causal in nature; and they may or may not govern all features of every phenomenon. But if there are such laws, science is by and large not in the business of uncovering them—as we have seen, not even fundamental physics. Whatever the ultimate nature of causation, the sense of causal relationships employed in scientific reasoning is one applicable to everyday human experience, including especially how we in fact and in principle exert our own influence on the world. This is what results in a science useful to and comprehendible by limited human agents.

A second implication of this distance from causal metaphysics is that it grants my conception of causal patterns immunity from a set of concerns that would otherwise need to be addressed. As a set, these can be summarized as concerns about how causal relationships are structured by levels of description, viz., when causes and effects are characterized more specifically or more generally. One of these concerns regards causal exclusion. Put in general form, causal exclusion is the idea that if a higher-level, abstract property is realized in any given case by some specific, physical property, then it is this physical property that is doing all of the causal work. Accordingly, the

[6] I suspect that, for similar reasons, a manipulability approach to causation is also conceptually basic.

higher-level property is itself not causing anything. For an influential version of this idea, see (**?**). If this idea is right, then it is misguided to ascribe causal powers to any properties that are multiply realized, that is, occur in a variety of physical systems. This seems like it may bar many of what I want to call causal patterns. However, the causal exclusion argument seems to presume a causal foundationalism, where physical causation is metaphysically basic. Setting aside the question of the metaphysics of causation and focusing instead on the epistemically basic manipulability approach to causation renders the causal exclusion argument irrelevant. On a manipulability approach to causation, any property (or, variable in Woodward's sense) on which we might possibly intervene can turn out to be a cause. Other claims have been made about the limitations levels of description place on causal relationships. Some are concerned about "top-down" causation, that is, positing higher-level properties as the causes of lower-level properties. Others posit that lower-level causal relationships constrain higher-level causal relationships. None of these limitations are reflected in the causal ascriptions warranted by a manipulability account of causation. This may lead one to reject manipulability as causal metaphysics. But if the use of a manipulability approach is distanced from metaphysical implications, there is room for commitments to ideas about the structuring of causal metaphysics by levels *and* causal patterns based on manipulability relations.[7]

2.1.2 Causal Complexity

One of the two bases for this project identified in Chapter One is the recognition of pervasive complexity. The above argument for the centrality of the search for causal patterns in science has provided the groundwork for a more careful characterization of the relevant sense of complexity. Phenomena that are the target of scientific investigations almost always result from a wide range of diverse causal influences that interact in complicated ways. Consider the range of causal influences that come to mind for, say, the trajectory of a forest fire, animals' traits, and climate change, to name just a few examples. I term this *causal complexity.* Causal complexity may not be overtly controversial, but it is certainly underemphasized in philosophy, and its implications are accordingly underappreciated.

[7] I am, nonetheless, suspicious of all such metaphysical conclusions based on levels of organization. I take up that issue in Chapter Four.

Let's look more carefully at the example of the causal influences on animals' traits. This example has some themes in common with the earlier discussion of genomics projects, where I briefly touched on the broad range of influences on human traits (summarized in table 1.1). Any evolved trait—say, the feather color(s) of some species of bird—is influenced by numerous types of causes. Feather color can result from pigmentation, how light is refracted, or a combination of these. Pigmentation and light refraction are, in turn, each subject to a number of causal influences. Genetic and epigenetic factors together result in the heritability of these traits; this is why offspring tend to have similar coloring to their parents. Numerous developmental factors also influence the expression of these traits, including the factors influencing the development of feathers themselves. Feather color in a variety of species has also been subject to population-level causal processes such as natural selection and drift. It is also subject to within-lifespan influences, including direct environmental influences. Carotenoids, one of the main types of pigments influencing the color of feathers, is acquired through birds' diets. Some of this list of causal influences occur simultaneously, the occurrence of others partially overlap, and still others take place at wholly different points in the trait's causal history. Many of these causal influences interact with one another. Genetic and developmental factors, for example, can only exert their influence in combination with one another. Similarly, parrots' green plumage results from feathers that reflect blue light and that contain yellow pigment.

Untangling such causal complexity is made more difficult by the limitations of our representations. Depicting the full gamut of causal influences in a single representation is impossible. First of all, causal histories stretch indefinitely far back in time. Even in a given period of time, there is often a wide range of influences, as well as interactions among them. For one generation of a single population of birds, feather color is influenced by genes (probably many, separately or in combination); developmental processes; environmental factors; and influences like predation. At least some of these influences exert effects on one another: genes influence development, as does the environment; predation may influence representation of feather colors, but feather color also influences survival. Moreover, a delimited set of causal influences can be represented in multiple, incompatible ways. In other words, the causal space can be parsed in different ways. Genes can be represented in all their molecular glory, or represented more abstractly as Mendelian units of inheritance, or more abstractly still as simply trait heritability.

Consider, as an example both of causal complexity and the investigative and representational quagmires it produces, the Foresight Tackling Obesities project (**?**).[8] One of the main products of that project is an "Obesity System Atlas." This atlas attempts to represent in a single "map" all of the primary causal influences on human obesity, including positive and negative causal influences (figure 2.2). Many of these causal influences are actually feedback loops. As communicated in the introduction to the atlas, the project aims are to confirm the "inescapably systemic and messy nature" of the phenomenon of obesity; to help in the identification of relevant variables; and to help determine where in the system to intervene effectively. This emphasis on intervention accords nicely with a manipulability approach to causation, as well as the focus on human action that approach reflects.

Despite this overt focus on complexity, this representation of the influences on human obesity is limited. Consider the following three limitations. First, this map only treats causal influences quite close to the phenomenon of obesity. It would be possible to further expand the map by identifying other causal influences that act upon the represented causes. Second, the causal diagram is meant to represent the general phenomenon of human obesity. It is thus best interpreted as representing causal patterns, from which we should expect some degree of variation in any individual case of obesity. Third, this map does not represent significant features of each of the focal causal relationships, including the magnitude of causal influence and its degree of invariance. By magnitude of causal influence, I mean the strength of a given causal relationship, or how much some cause positively or negatively contributes to the production of obesity or some intermediate variable. The concept of invariance was introduced in the discussion of Woodward's manipulability account of causation; this is the range of conditions over which a specific causal relationship continues to obtain. Without knowledge of the ranges of invariance for any of these causal relationships, we cannot say whether a shift in the value of some variable will alter or eliminate its role in the causal system of obesity. For example, this map represents degree of physical education as positively influencing level of recreational activity. It may be, though, that there is some threshold of physical

[8]Robert Skipper recommended this example.

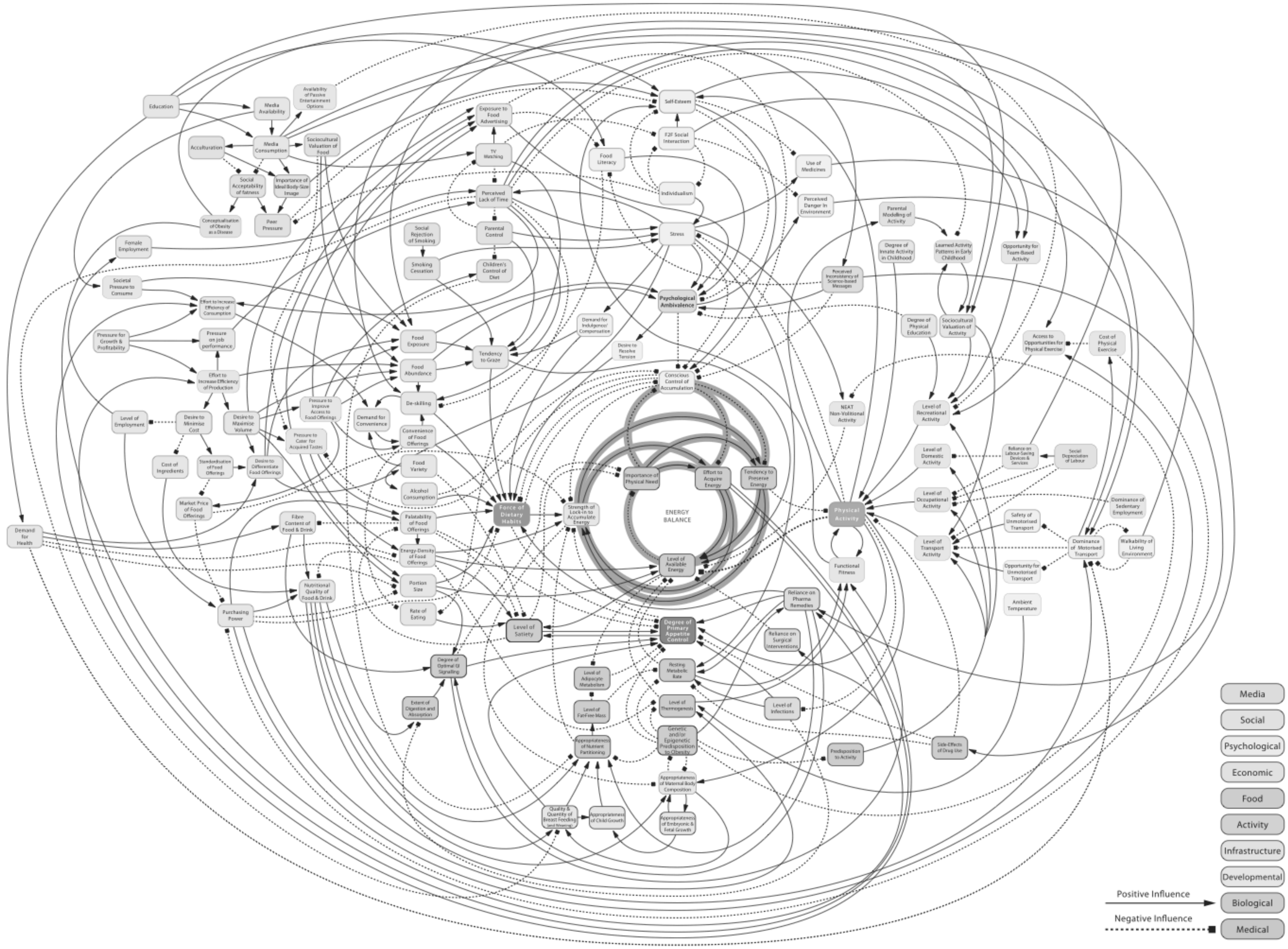

Figure 2.2: This map from the Obesity System Atlas shows all causal influences they identified for human obesity. Solid lines represent positive causal influences and dotted lines negative causal influences.

education that, if exceeded, does not result in further increases to recreational activity. This shows that scope of invariance is just as significant to effective action as are manipulation relationships. To return to my general point, I point out these limitations of the primary obesity system map not to criticize the representation, but in order to show that causal complexity far outstrips even this nuanced representation.

The atlas also includes a number of other maps that depict various elements or layers. These provide additional insight into some of the difficulties in investigating and representing these diverse influences. One series of maps divides the identified causal influences into several "thematic clusters," such as social psychology, individual psychology, physiology, etc. The influences in these clusters are largely investigated in different fields of science. This is another dimension of causal complexity—*types* of causal influences on a phenomenon do not fall neatly into one or even a few fields. Another series of maps details the causal connections among influences falling within different thematic clusters. These causal connections that transcend field boundaries can pose special difficulties, as fields of science tend to differ in their methodology, viz., their tools of investigation and representation. Finally, a series of maps explores which causal networks and feedback loops potential interventions would affect. This conforms to the project's stated goal of effective intervention on the system. Notice, though, that effective intervention requires additional investigation into, at least, the invariance of the causal relationships in question.

The scope of invariance for a causal relationship is just one of many complications in how causes exert their effects. We have also discussed a few types of causal interdependencies, including feedback loops, causal influences on some effect also influencing one another, and causes that require the presence of other factors in order to exert their influence. Mitchell (2012b) argues that causal influence often fails to be modular, that is, that causal contributions are in general not additive and may thus be inseparable. She appeals to an example related to my main illustration here, namely, a gene's causal contribution to the production of a phenotypic trait. She points out that one gene in a network may causally contribute to the production of some phenotypic trait under normal conditions but, when a different gene is disrupted, have a very different effect. This makes employing anything like Woodward's "ideal experimental manipulation" very difficult, if not impossible. An intervention requires that the values of all variables not on the causal path in question be held fixed, but in this case, intervening on the focal gene leads to changes in the causal contributions of other genes as well.[9]

So far I have illustrated causal complexity and indicated some of its dimensions. One might yet wonder how common causal complexity, especially in its more extreme forms, really is. I can offer two types of justification for a belief in widespread causal complexity. First, what we know about and are increasingly learning about the world corroborates this view. There is a proliferation of complex systems approaches in science, applied to a wide range of fields. I surveyed a few examples of these approaches in Chapter One: dynamic systems theory, developmental systems theory, systems dynamics, chaos theory, and systems biology. Furthermore, in many fields of science, there is an ever-broadening conception of important causal factors. This is exemplified by the trajectory described in Chapter One from the Human Genome Project to the Thousand Genomes Project to the Human Microbiome Project, as well as by the intricacies of the Obesity System Atlas. Any number of other examples of causal complexity can be generated with the following exercise: choose any phenomenon investigated in science, then consider the types of causal influences on that phenomenon. Remember to include background conditions, causes at earlier and later time periods, and causal influences on the causes you have already identified. When your list grows long, begin to consider how those influences overlap and influence one another. This is causal complexity.

My second justification for a belief in widespread causal complexity is the scientific significance of causal patterns—in particular, of limited patterns with exceptions. The lawlessness of science, science's continuing failure to identify laws of nature, is itself a corroboration of causal complexity. With multitudinous interacting causal influences, universal laws are either nonexistent or incomprehensible by limited human intellects. We are left with the search for causal patterns as described above. There are other views that could account for the lawlessness of science. Cartwright (1983, 1999) supposes that phenomena in our world are simply unruly. She defends the plausibility

[9]Mitchell couches this discussion as a criticism of Woodward's account of causation. However, she bases her remarks on (?). Woodward (2003) is explicit that modularity is a desirable property of representations of complex systems, not a property of causal systems themselves, and not a requirement for causal attributions.

of the idea that nature is “constrained by some specific laws and by a handful of general principles, but it is not determined in detail, even statistically (Cartwright, 1983, p.49). The downside to such a position is that it cannot account for the scientific successes we do find, namely the causal patterns that have been discovered for a broad swath of phenomena of interest to us humans. It is thus the combination of lawlessness and, yet, widespread success with discovering causal patterns, that suggests a world rife with complex and variable causal interactions.

2.2 Simplified Representation

So far in this chapter I have made the case that science is profitably understood as a search for causal patterns in the face of causal complexity. In this section I explore a widespread strategy for accommodating this situation, namely, with representations that idealize away much of the complexity. I begin by surveying some existing treatments of idealized representation, a topic that relates closely to the practice of scientific modeling. I then defend a stronger view of the significance of idealization in science. I argue that there are many intertwined reasons to idealize. I then make the case that idealizations play a positive representational role, which distinguishes them from abstractions, as well as other differences between a representation and the system(s) it represents. Finally, I use this account of the practice of idealization to motivate the idea that idealizations are both rampant and unchecked in science. By *rampant* I mean that idealizations are found throughout our best scientific products, and they stand in for even crucial causal influences. By *unchecked* I mean that little effort is put toward eliminating or even controlling these idealizations.

2.2.1 Model-Based Science and Reasons to Idealize

The important role that models play in science has, in the past decades, been increasingly appreciated by philosophers. Mary Hesse (1966) articulated a view of scientific models as analogies, and she argues that this analogical role is essential to science. A different understanding of models rose to prominence in philosophy of science with the semantic theory of science (Suppe, 1977), according to which models were understood to be mathematic structures that serve as interpretations of axiomatic scientific theories. This is consistent with the logician’s sense of models.

Models in this sense were also central to Bas van Fraassen's (1980) constructive empiricism. Giere (1988) contributed to the prominence of something more like Hesse's view of models as analogies. Like Hesse, Giere was struck by the overt idealizations prominent in science textbooks, such as frictionless pendulums and bodies subject to no external forces. He notes the overlap with logicians' terminology but is critical of the idea that models should be isomorphic to real-world systems. Instead, on Giere's view, successful models are related to the world via their similarity. Similarity is a weaker requirement than isomorphism, and it also requires the specification, at least implicitly, of the respect and degree of the similarity.

Giere (1988) points out that observations of science as it is actually practiced shows that models—in his sense—occupy center stage. His view of models has partly inspired a literature on scientific modeling that emphasizes accounting for the role of models in actual scientific practice. This accords well with the commitment expressed in Chapter One to account for current science as it is actually practiced. Additional early inspiration was drawn from Richard Levins (1966), who addresses population biology in particular, as well as Wimsatt (1987). Distinctive features of this approach include its focus on models' incorporation of abstractions and idealizations, and thus only partial representation of real-world systems, as well as the recognition that models can be employed independently of theory or without the aim of immediately representing a real-world system.

Peter Godfrey-Smith (2006) introduced the term *model-based science* to characterize this approach to science, based on the construction and analysis of abstract and idealized models. Godfrey-Smith (2006) and Michael Weisberg (2007b, 2013) both emphasize that this type of modeling is not used throughout all of science, but is instead a distinctive approach. In an alternative approach that Weisberg terms abstract direct representation, the aim is simply to describe an actual system in order to investigate it directly. In contrast, the aim of modeling is to indirectly represent a real-world system by describing a simpler, hypothetical system and investigating that simpler system, in order to draw conclusions about the actual system of interest. This is similar to the emphasis that Margaret Morrison and Mary Morgan (**?**) place on models' autonomy from both theory and the world. In virtue of this strategy of indirect representation, models are explicitly intended to represent their target systems only partially. Models and the target systems they are used to represent bear some features in common, while other features are neglected or falsified. This is accomplished via the use of abstractions and idealizations.

Attention to model-based science is, thus, related to an investigation into the nature of idealization. As Wimsatt (1987, 2007) says, "Any model implicitly or explicitly makes simplifications, ignores variables, and simplifies or ignores interactions among the variables in the models and among possibly relevant variables not included in the model (p.96)."[10] These are all idealizations. Most broadly, *idealizations* are features of representations that misconstrue the represented systems. Many people are familiar with the common assumption in physics of frictionless planes, and with the common assumption in economics that humans are perfectly rational agents. These are both idealizations: every plane has friction, and no human is a perfectly rational actor. We saw above that Newton's law of universal gravitation also idealizes, for it assumes that each massive body occupies a single point.

Although model-based science cannot proceed without idealizations, idealizations need not be unique to models. In what follows, I thus refer in general to "representations" that idealize. I mean this as a neutral term to include models as well as any other representational structures that may be utilized in science (e.g., Weisberg's abstract direct representations, theories, and ceteris paribus laws). I also wish to avoid the questions of what exactly is being represented, as well as the nature of the representation relation. I adopt Weisberg's (2013) terminology and speak in terms of representations of target system(s), but for my purposes this is a placeholder, not a substantive assertion about what is represented.

An initial puzzle about idealizations is why, when the aim is to represent one or more systems, one would intentionally introduce an assumption that is false of those systems. It turns out that there are several answers to that question. Many different motivations have been suggested for the incorporation of idealizations. For instance, Cartwright (1983) claims that idealizations make for more illuminating, explanatory models. **?** discusses how the evaluation of several models with different idealizations can lead to the discovery of "robust" results that do not rely on any particular

[10] (Wimsatt, 1987) is republished as Chapter 6 in (Wimsatt, 2007). All page numbers in my citations refer to the latter publication.

false assumption. Ernan McMullin (1985) discusses how idealizations facilitate mathematical or computational tractability. Robert Batterman (2002) emphasizes idealizations' contribution to accounting for "stable phenomenologies," or repeated general behavior.

Weisberg (2007a, 2013) assimilates several of these views about the nature of idealization in order to distinguish among three distinct types of idealization. The first of these is Galilean idealization. These are simplifications needed in order to secure computational tractability, to be eliminated—and the model "de-idealized" (McMullin, 1985)—if and when it proves possible. The second is minimalist idealization, which is the elimination of all but the most significant causal influences on a phenomenon. The third type of idealization Weisberg identifies is multiple-models idealization, which is the use of several distinct, simplified models that together shed light on a phenomenon. Weisberg appeals to different "representational ideals," or what elements of a system researchers desire to represent, in order to distinguish among these three types of idealization. In his view, Galilean idealizations are employed (and eliminated whenever possible) to the end of complete representation; minimalist idealizations facilitate representation of crucial causes; and—significantly for my purposes—multiple-models idealizations can facilitate a range of different representational ideals. Rohwer and Rice (2013), in turn, argue that there is a further type of idealization, overlooked by Weisberg, which they call hypothetical pattern idealization. On their view, this results in a model that represents no actual systems but that is helpful in theory-development.

These accounts of idealization indicate that there are a variety of motivations behind the incorporation of idealizations into scientific representations. Weisberg's view that no single motivation accounts for all idealizations must be right. As Weisberg demonstrates, idealizations serve a wide range of purposes in representations, and the uses of scientific representations vary greatly as well. Both of these ideas will be additionally corroborated by the considerations and case studies introduced in this chapter and the next. Rohwer and Rice's point that Weisberg's taxonomy is not adequate to capture all of the purposes to which idealizations are put must also be right. Weisberg only allows for idealizations as temporary expedients, or to facilitate the representation of the core causal factors, or when employed in combination with other models, incorporating different idealizations. As Rohwer and Rice note, this overlooks at least one important circumstance in which idealizations are found: idealized models that obviously neglect important causal factors and yet are not employed in combination with other models. This circumstance of idealization is very common and, it will emerge, crucial to an accurate interpretation of the role of idealization in science. This style of idealized model is often required in order to represent the causal patterns embodied by causally complex systems.

* * *

Nonetheless, there is an important shortcoming of both Weisberg's and Rohwer and Rice's attempts to delimit types of idealizations. As a first step toward motivating the problem, notice that almost all of the circumstances in which idealizations occur fall within Weisberg's description of multiple-models idealization. On Weisberg's analysis, multiple-models idealization occurs when "our cognitive limitations, the complexity of the world, and constraints imposed by logic, mathematics, and the nature of representation, conspire against simultaneously achieving all of our scientific desiderata" (Weisberg, 2013, p.104). I have argued that the effects of causal complexity and cognitive limitations on science are quite general. Weisberg should agree that this category of idealization is by far the most significant, for what he characterizes as multiple-models idealization defines model-based science, which is his main focus. Furthermore, we saw above that multiple-models idealization is also the type of idealization that on Weisberg's view is not defined by a single representational ideal, even though Weisberg proposes that these representational ideals are what distinguish the different types of idealization. It follows that the most significant type of idealization cannot be clearly delineated on Weisberg's taxonomy.

Weisberg might instead delineate this type of idealization by the fact that multiple models are in use, but this introduces a further difficulty. There are two very different ways in which multiple models might be in use: within a single research program, or across the scientific enterprise as a whole. The former is a narrow sense of employing multiple models. Employing multiple models with conflicting assumptions within a single research program is a distinctive approach to science that facilitates comparisons among models' assumptions and findings, often used as the basis for robustness analysis (**?**Weisberg, 2006). In contrast, the latter is a broad sense of employing multiple models. The scientific enterprise as a whole employing multiple models with conflicting assumptions for a single phenomenon is a quite common state of affairs, for this is a straightforward consequence of causal complexity and human cognitive limitations. A research focus on one or another type of causal influence will lead to a distinctive set of idealizations to facilitate the representational goal immediately at hand. It is unclear whether Weisberg intends the narrow or broad sense of employing multiple models. On the one hand, he appeals to the United States National Weather Service's practice of employing multiple incompatible models, which is an instance of the narrow reading. On the other hand, he also cites Levins' view that "communities of scientists" construct multiple models that "collectively can satisfy our scientific needs" (p.104), which suggests the broad reading.

This ambiguity may in part account for Rohwer and Rice's proposed amendment to Weisberg's taxonomy. Rohwer and Rice (2013) focus on idealized models that neglect significant causal influences (so do not qualify as minimalist idealizations) and that are used singly, not in combination with models incorporating conflicting assumptions (so do not seem to qualify as multiple-model idealizations). This type of idealized representation is exceedingly common in light of the combination of causal complexity and focused research interests. Perhaps Weisberg would consider this use of idealization to qualify as multiple-models idealization in the broader sense that others in the scientific community employ models with different idealizations. Weisberg account leaves us, then, with two options. Either idealized representations like those upon which Rohwer and Rice focus—a very important type of idealization—are not accommodated by Weisberg's taxonomy, or else the category of multiple-models idealization is so broad that it is a sort of dustbin category, uninformative about the features of idealizations of that type. Neither of these options is satisfying. Moreover, Rohwer and Rice's strategy of simply introducing a fourth category is not the right approach. This overlooks the key insight to be gained from the limitations of Weisberg's taxonomy, namely, that taxonomizing types of idealizations in this way yields little insight into *why* idealizations are employed.

For Weisberg, representational ideals are supposed to determine the type of idealization by indicating the value of the idealization. Is it a temporary expedient to the end of complete representation; a way of accomplishing the representation of core causes; or (in the case of multiple-models idealization) does it serve one of many possible purposes in service of one of many possible aims? This approach is of limited success due to the large number of reasons for employing idealizations—reasons that occur in combination. The types of idealization that Weisberg isolates are supposed to be motivated by computational tractability; the relative unimportance of some causal influences; and tradeoffs in achieving scientific desiderata due to cognitive limits, complexity, and constraints of logic, math and representation (cf. Levins, 1966). But surely these reasons to idealize are often present simultaneously. I have suggested that causal complexity and cognitive limits are quite general features of science. So too is an interest in expedience, including computational tractability. Weisberg overlooks other ways in which idealizations can be expedient, such enabling the reapplication of modeling techniques in which a researcher is well-trained to unrelated phenomena. Rohwer and Rice focus on the identification of general patterns as a motivation for idealization, and if what I have said in the first half of this chapter about the identification of causal patterns is right, then this motivation too applies quite broadly.

All of these reasons to idealize and more recur in various combinations throughout the scientific enterprise. Accordingly, instead of distinct types of idealization, there are many *intertwined reasons to idealize.*[11] I address how these reasons to idealize positively contribute to science's aims in Chapter Three. I suggest there that the list of reasons to idealize is open-ended. It is, nonetheless, possible to categorize these reasons. First, some of them justify the incorporation of an idealization merely temporarily, and others justify permanent idealization. This decides whether de-idealization will ever be warranted, a difference that is also reflected in Weisberg's taxonomy. Second, some of these reasons are due primarily to features of our complex world, and others are due primarily to features of scientists themselves or of their audience. These distinctions are independent, giving rise to a categorization like what is depicted in table 2.1.

One might wonder why these observations should be taken to supplant Weisberg's taxonomy of types of idealization, instead of just supplement that scheme or something like it. There are two reasons. First, this account of intertwined reasons to idealize accomplishes the tasks for

[11] A version of this idea initially emerged from a discussion with Anthony Chemero, Thomas Polger, and Robert Skipper.

	Due primarily to the world	Due primarily to scientists' features
Temporary	technique happens to get traction	limits of computational power familiar technique preparatory for different approach
Permanent	computational limits captures core causal influences	cognitive limits limited research focus pedagogical value

Table 2.1: Some of the intertwined reasons to idealize

which Weisberg's taxonomy and representational aims were intended. It answers the question of why idealizations are employed (and, like Weisberg's taxonomy, give a pluralistic answer). The categorization of reasons to idealize also indicates whether idealizations are temporarily or permanently of value and, thus, when de-idealization is warranted. Second, because reasons to idealize occur in combination, the distinctions among the idealization-types that Weisberg develops break down. Different idealizations in a single representation may be present for different reasons, or one idealized assumption in a representation may have different justifications over time or in different circumstances, all without a change to the representational aim. For example, consider evolutionary game theory models, including the Hawk-Dove game which serves as Rohwer and Rice's main example. One of the many idealizations common in evolutionary game theory is the assumption that a population of organisms is infinite in size. This idealization enables the neglect of the possible influence of genetic drift, that is, changes in frequency due to random sampling. When a game theory model is applied to a very large population, this idealization does not interfere with the accurate representation of the core causal influences, for in that case, random sampling is not a significant influence. If the model is instead used to represent a smaller population, this idealization *does* interfere with the representation of a significant causal influence. What justifies the idealization in this latter instance is not a feature of the world, but instead a research focus that does not include the causal role of drift. In table 2.1, these respective motivations for the assumption of an infinite population are shown as "captures a causal pattern" and "limited research focus."

Figure 2.3: This version of the BART map depicts the stops within San Francisco as equidistant and in a straight line. The official BART map also depicts the Richmond-Millbrae line as red in color. The former is an idealization, while the latter is not. Credit: San Francisco Bay Area Rapid Transit District / www.bart.gov / CC BY 3.0.

2.2.2 Idealizations' Representational Role

Above I considered a variety of motivations that have been suggested for the incorporation of idealizations into scientific representations. I argued that these motivations cannot be used to neatly distinguish different types of idealizations, and that they are due both to features of the world and to features of scientists themselves. Here I shift my attention from the motivations for idealization to the nature of idealizations themselves. In particular, I suggest that idealizations play a positive representational role.

To begin, notice that not every difference between a representation and the target system(s) qualifies as an idealization. Consider, as a toy example, the official map of the Bay Area Rapid

Transit (BART) system in the San Francisco Bay Area. A simplified version of this map appears in figure 2.3. Notice two features of this map. First, in the original map, the Richmond-Millbrae line is depicted as red in color. Second, all stops within San Francisco are depicted as equidistant and in a straight line. Both are differences between the map and the BART system, but only the second counts as an idealization. One might articulate the relevant difference between these inaccurate features of the map by pointing out that the map represents the BART lines as if their stops within San Francisco were equidistant and in a straight line. In contrast, the map does not represent the Richmond line as if it were red. This same way of distinguishing between idealizations and other differences between a representation and system(s) is natural for scientific representations as well. Newton's law of universal gravitational represents massive bodies as if they are point masses when they are not; this is a difference between the law and any system to which it applies that qualifies as an idealization. Newton's law can also be written in a mathematical formalism, whereas no system to which it applies can be. This is a difference between the law and the system, but it is not an idealization.

This distinction between idealizations and other differences between a representation and what is represented is that idealizations represent a system as if it were some way that it is not. I propose, then, that idealizations actually play a representational role. An idealization does represent the target system(s), for it represents a system as if it were some way that it is not. I call this *representing as-if*. This idea fits naturally with the paradigmatic type of idealization, namely, representing a system as if it were ideal in some regard. A surface may be represented as if it were a frictionless plane; an individual as if she were a perfectly rational agent; a population of organisms as if it were infinite in size. Not all idealizations are naturally construed as representing a system as if it were ideal; an ideal physical body may not be a point mass. The general insight is that an idealization represents a system as if it possessed some feature(s) that it does not. Idealizations thus must be accorded a positive representational role.

Many distinguish idealization from abstraction, where the former is the misconstrual of a system's features and the latter is neglect of a system's features. As Cartwright (1989) says, in regard to the idealization that a plane is frictionless, "The model may leave out some features altogether which do not matter to the motion, like the color of the ball. But it must say something, albeit something idealizing, about all the factors which are relevant" (p.187). Weisberg (2013) lumps both practices together into the category of idealization. The preceding considerations about idealizations' positive representational role indicate an important difference between abstraction and idealization. Whereas with idealization, the target system(s) is represented as if it had some features it does not, with abstraction some features of the system(s) are simply omitted, that is, are ignored. Idealizations are, in this sense, fictions.[12]

What to say, then, about the nature of idealizations' representational role? Accounts of representation appeal to commonalities between a representation and that which is represented, such as their isomorphism (van Fraassen, 1980), analogy (Hesse, 1966), or similarity (Giere, 1988; Weisberg, 2013). In light of this, it is puzzling how something that is patently false of a system can represent that system. To begin to address this puzzle, consider effective population size, or N_e, which is a common variable in population genetics. A popular population genetics textbook (Hartl and Clark, 1997) defines effective population size as "the number of individuals in a theoretically ideal population having the same magnitude of random genetic drift as the actual population." N_e is thus an idealization. It represents the population of organisms that is of interest as if it were a size that it is not. Indeed, N_e is defined as the property of a fictional entity—an ideal population. It represents the real population in virtue of the similarity in the behavior of that population and the ideal population characterized. Effective population size represents a feature of real populations not in virtue of a similarity between the population and the described ideal population, but in virtue of a similarity in their behavior, in particular, the similarity in how they exhibit genetic drift. To generalize from this observation, idealizations represent systems as if they had some feature(s) they do not, and idealizations qualify as representational in virtue of some resulting behavioral similarity.

This is fitting. I have argued that much of science is productively understood as the search for causal patterns, and that this is furthered by idealized representation. Now we see that idealizations

[12]That I call idealizations fictions should not be confused with the view that models are fictional systems. I do not address the status of models, nor the question of in virtue of what models represent the world. My attention here is solely on the idealized parts of models and other representations, and *their* representational relationship to the world.

represent systems as if they have some feature they do not, a fiction that is warranted on the basis of behavioral similarities, that is, similarities in the causal patterns they exhibit. Idealizations enable the portrayal of causal patterns in virtue of this representing as-if. This results in the elevation of fictional properties and entities—like ideal populations and frictionless planes—to central roles in scientific investigations. In physics there are references to ideal gases, defined as theoretical gases composed of non-interacting point particles in random motion, and perfect gases, which are fictional gases additionally simplified with other stipulations, including that they are not chemically reactive. A nice example of the central role of such constructs is provided by the idealization of effective population size. Consider how Lande and Barrowclough (1987), in a heavily cited chapter in a collection on conservation and population biology, employ the concept. These biologists say that the purpose of their chapter is to "show how the effective size of a population, the pattern of natural selection, and rates of mutation interact to determine the amount and kinds of genetic variation maintained" (p.87). This indicates that they are investigating, in part, the influence of a property of a fictional, *ideal* population. The influence of effective population size is surely of interest to these researchers in virtue of its similarity to the influence of the size of actual populations.

The consensus in the literature on model-based science is that idealized models can represent despite their false assumptions. A stronger position still is that the idealized parts of those models—the false assumptions themselves—can represent, despite their falsity. This is a significant departure from how idealizations have been treated in philosophy of science, but I see no reason to balk at the move. Idealizations represent features of systems, for they represent those systems as if they possessed features that they do not. This qualifies as representation in virtue of similarities in the behavior of the fictionalized representation and of the represented system(s), and as such, it contributes to the search for causal patterns. Idealizations, and the fictional entities and properties they posit, are thus integral to scientific practice.

2.2.3 Rampant and Unchecked Idealization

We have seen that idealizations are a central feature of many representations, and that there are many different reasons to idealize that, in various combinations, motivate the incorporation of idealizations into representations. I have also argued that idealizations actually play a positive representational role, representing systems as if they had features they do not. Together these points about idealization lay the groundwork for the view that idealizations are absolutely central to science. Here I develop that view, suggesting that idealizations are both rampant and unchecked in science.

Consider, first, the idea that idealizations are *rampant* in science. By this I mean that idealizations exist throughout our best scientific representations. Most basically, this is because causal complexity renders idealizations not only helpful, but even requisite in the pursuit of causal patterns. Any choice of causal influences to represent requires some degree of misrepresentation elsewhere. This is particularly so for representations tailored to human needs and, thus, human limitations, as our scientific representations are. These circumstances give rise to the many intertwined reasons to idealize, both temporary and permanent, some due primarily to features of scientists and others primarily to features of the world. The extensive idealizations that result exceed what many philosophers expect. Weisberg's view accommodates continuing idealization only when the idealizations do not stand in for central causal influences or when multiple models are employed, with other models representing the idealized factors. In contrast to that view, it seems that significant causal influences are regularly idealized without recourse to multiple models. This is so for highly idealized models that are not applied to any particular systems, as Rohwer and Rice (2013) emphasize. It also is the case for predictive models that sacrifice realistic causal representation, and many other purposes of idealized representation. There is another regard in which the centrality of idealizations exceeds expectations. Fictional properties and entities, such as effective population size and perfect gases, that constitute idealizations can take on a scientific life of their own. As we have seen in the case of effective population size, research can be devoted to exploring these constructs in much the same manner as their real analogs.

Misrepresentation by idealizations, and not simply the omission of causal influences via abstraction, is required for several reasons. First, representing causal patterns requires stage-setting. We have seen that causal patterns hold only within a limited range of circumstances, and have deviations and exceptions even within that range. Focal systems must be represented in ways that reflect the behavioral regularity—the causal pattern—and this requires making idealizing assumptions. Accordingly, idealizations represent systems as if they had properties they do not. For instance, the idealization of an infinite population size is required in order to represent patterns in the influence of natural selection across populations that have experienced different degrees of genetic drift. Second, when there are many interacting causes, focusing on a subset of causes requires somehow accommodating the interaction with other, non-focal causal influences. The idealization of effective population size reflects the influence of a number of a population's features, including its sex ratio and mating structure, without depicting those features. This facilitates the representation of patterns across populations that vary in those features but respond to drift similarly. Third, false assumptions can facilitate computational tractability in a way that simple omission cannot. Newton's law of universal gravitation would not be nearly so simple without the assumption of point masses. Misrepresentation instead of omission is not just required, but is in some regards beneficial. Representing as-if is of practical benefit, for it enables the reapplication of existing approaches to disparate systems. It is also of epistemic benefit, for it demonstrates similar behaviors among very different kinds of systems. These considerations are reflected in the many intertwined reasons to idealize set out above.

The second part of my view about the centrality of idealizations is that these rampant idealizations are also *unchecked*. By this I mean that there is little focus on eliminating idealizations, or even on controlling their influence. Consider first the claim that there is little focus on eliminating idealizations, or on what has been called de-idealization. Weisberg (2007a, 2013) and others rightly observe that much idealization is unaccompanied by the goal of eventually reducing or removing the idealizations to increase the accuracy of representations. Indeed, several of the intertwined reasons to idealize identified above are due to permanent features of human scientists and their limitations or of the complex world we inhabit. Idealizations can best facilitate the representation of core causal influences or can enable limiting the representational focus to causes of primary research interest. They can also be of permanent value in light of absolute computational limits or our human cognitive limits.

The idea that there is little focus on controlling the influence of idealizations is stronger still. Idealizations that persist tend to be justified in one of two ways. One might think that idealizations are warranted in virtue of features of the world, and in particular facts about causal significance and insignificance. This is the case for Weisberg's minimalist type of idealization and Strevens' (?) view of idealization. For both, idealizations are supposed to stand in only for insignificant causal influences. Naturally, this does not accommodate the persistent idealization of significant causal influences. Weisberg's multiple-models type of idealization does sanction this use of idealizations, but as we have seen, only when the requirement is met that alternative models, employing alternative idealizations, are also developed. Rohwer and Rice (2013) make the case that there are idealizations that do not satisfy either of these requirements (and that are not subject to de-idealization either). Yet, they in turn emphasize that representations employing this version of idealization cannot possibly provide explanations, in virtue of their unchecked idealizations. All of these are ways of "checking," or controlling, the influence of idealizations. In contrast, I have developed a view of idealization according to which there are many reasons to idealize, and on a permanent basis, even when central, focal causal influences are idealized. I make the case that this does not result in substandard representations (e.g. non-explanations) in the next chapter.

Unchecked does not mean unprincipled. Intertwined reasons to idealize, like those identified above, motivate the inclusion of idealizations and determine their nature. Idealizations play a positive representational role by representing as-if, and the nature of that representation must be appropriate for the actual causal patterns in focal phenomena, for the aims of the research, and for the model or other representation employed. An infinite population is not a helpful idealization when a recent bottleneck in population size swamps all other evolutionary influences; or when the research focus is the role of genetic drift, however major or minor; or when one is employing the logistic model of population growth, which requires reference to the actual (finite) population size. An idealization that is inappropriate in any of these regards will be eliminated or replaced with a different idealization whenever possible. The same is true if the research focus or methods change in a way that warrants a different approach to idealization.

Chapter 3

Science Isn't After the Truth

A fairly unified picture of scientific practice emerged in Chapter 2. There I argued that depicting and otherwise capitalizing on causal patterns is a central feature of science. More specifically, scientific research often involves a search for causal patterns in the face of causal complexity. These are patterns and not laws because they hold in a limited range of circumstances, and phenomena deviate from the pattern to some degree—sometimes to such a great degree that they are exceptions to the pattern. Phenomena can embody multiple, distinct causal patterns, and in our causally complex world, many phenomena do. Accordingly, scientific projects involve continuing idealization; that is, idealization is both rampant and unchecked.

This fairly unified description belies an astounding diversity of scientific projects. This chapter begins with a small survey of that diversity. Even in the few examples I consider, there is remarkable difference in the aims of the research, the relationship to data, and the type of connection to other research. The analysis of these research projects and the tremendous diversity among them is the topic of 3.1 below. Despite the many differences among these projects, though, some generalizations can be made about them. In particular, I will argue in 3.2 that their diverse features have a common cause. That cause is the influence of unchecked, rampant idealization on the aims of science. I thus develop a revised version of the basic aims of science in light of this reconceptualization. In my view, science simply isn't after the truth. Instead, there are a variety of aims that are in tension with one another, and the ultimate epistemic aim of science is not truth but understanding. In

3.3 I explore a related issue, namely, the relationship between science and metaphysics. It emerges that this conception of the aims of science drives a wedge between scientific practice and any direct metaphysical implications, which somewhat constrains the possibilities for a naturalized metaphysics.

3.1 The Diversity of Scientific Projects

Attending to even a few contemporary scientific research programs reveals a surprising diversity among the projects. Here I consider investigations of cooperation in behavioral ecology, a variety of approaches to investigating human aggression, computational models of climate change, fluid dynamics, and astrophysics. Differences include the aims of the research, the relationship to data, and how the project connects to other scientific research. Attending to some of the philosophical issues surrounding these research programs further illuminates their differences. And yet, despite this diversity, all employ idealized representations. Furthermore, all resist the interpretation that they are aiming for truth, or even accurate representations in order to generate predictions and explanations.

3.1.1 Modeling Cooperation

One main research focus in behavioral ecology is the evolution of cooperative behavior, especially among animals. Historically there have been three main approaches to accounting for cooperation: a variety of group selection models, kin selection models, and models of reciprocal altruism. Perhaps the best-known research into the evolution of cooperation is modeling reciprocal altruism using the prisoner's dilemma, a game theory model that was first applied to the evolution of cooperative behavior by Axelrod and Hamilton (1981). The prisoner's dilemma represents a way in which cooperation can emerge among self-interested individuals faced with particular circumstances; in its evolutionary interpretation, self-interest simply means maximizing one's individual fitness. The prisoner's dilemma model applies to encounters between organisms in which one strategy immediately benefits one individual's fitness at an immediate cost to the other individual's fitness, whereas a different strategy is less immediately advantageous but maximizes both individuals' fitness in the longterm, provided there are many such interactions.[1]

This model of the evolution of cooperation is, of course, highly idealized. Assumptions that are strictly speaking false include the assumption that reproduction is asexual; that the population is infinitely large; that individuals have serial, pairwise interactions; that payoffs are constant across individuals and iterations of the game; and that selective advantage is the sole influence on evolutionary outcomes (Maynard Smith, 1982; Potochnik, 2012; Rohwer and Rice, 2013). Most of these idealized assumptions are general across all behavioral ecology approaches to modeling the evolution of cooperation. Those modeling approaches differ in their setups and thus in how they represent the evolution of cooperative behavior. But all aim to account for the role of natural selection in producing cooperative behavior, and this focus determines what idealizations are useful. For example, the assumptions of an infinite population that reproduces asexually allow any influence of genetic drift or recombination to be ignored. These assumptions also greatly simplify the representation, thereby facilitating the use of modeling techniques that would otherwise be unavailable.

The prisoner's dilemma, then, is a highly idealized model of the evolution of cooperation. As suggested by the fact that kin selection models and group selection models share many of the same false assumptions, they too are highly idealized representations of cooperation. Kin selection models represent how cooperation can emerge among related individuals in virtue of the beneficial effects to inclusive fitness, viz., the fitness benefits one's relatives derive from one's cooperative actions. Group selection models, in turn, show how cooperation can evolve as a result of how it improves the fitness of a *group* of organisms, taken as a whole.

It is instructive to consider the directions in which this behavioral ecology research into cooperation is progressing. There is an ever-expanding variety of heavily idealized models of circumstances that generate cooperative behavior, such as other game theory models, like the stag hunt and snowdrift games. There is also work on the relationships among these alternative models, like Kerr and Godfrey-Smith (2002) on the mathematical relationship between game theory models and group selection models. There is exploration of the variability of the model

[1]Reciprocal altruism was first articulated by Trivers (1971). This is a very brief overview of reciprocal altruism and the prisoner's dilemma, but there is an expansive literature on each.

structure itself, such as Worden and Levin (2007) on the conditions under which populations might evolve away from the payoff structure of the prisoner's dilemma. There are also attempts to simultaneously represent evolution and within-lifespan influences on cooperation, like Akçay et al. (2009) on the evolution of mutual regard. All of these are further developments and explorations of highly idealized, mathematical representations. Many regard the mathematical machinery itself, without any attention to the representational relationship between the model and actual cooperative behavior.

Indeed, there are few examples of research that would lead to a more accurate representation of actual instances of the evolution of cooperation. When specific populations are discussed, they are almost always used as exemplar cases to motivate a general, idealized model. And quite often, exemplar cases are not actual phenomena, but imagined, simple scenarios that are evocative of real phenomena. Similarly, little work has been invested in de-idealization. This would involve incorporating into a behavioral ecology model more realistic assumptions about the features a particular population in which a cooperative trait of some kind has evolved, the genetics governing the inheritance of a cooperative trait, or other central causal influences that existing models wholly idealize. There thus seems to be little emphasis on applying these models to accurately represent specific instances of cooperative behavior.

How, then, is this research program related to empirical investigation? Above I claimed that these models are aimed at accounting for the role of natural selection in producing cooperative behavior. In (Potochnik, 2009) I call this the weak use of such models, to distinguish it from the aim of accounting for all significant causal influences. Selection is just one of several causal influences on actual evolutionary outcomes. These models are, thus, attempts to capture partial causal regularities—in particular, to show ways in which selection can causally contribute to the emergence of cooperation. Because these models ignore so many other causal influences, they have a limited range of application and limited accuracy of most any evolved trait. In Chapter Two I suggested representations of partial causal regularities generally have both these limitations. Research such as Eshel and Feldman's (2001) analysis of the conditions under which evolution can be expected to lead to game theory equilibria is profitably understood as demonstrating the range of these models' application and its limits. I suspect that the limited accuracy of these models of cooperation is the reason there is little emphasis on application to specific evolutionary phenomena. If this is right, then the relationship to empirical investigation is indirect, established by the partial causal regularities demonstrated in these models.

There are generally two forms of empirical corroboration that might be pursued for these models of cooperation (cf. Lloyd, 1988). First, researchers sometimes empirically corroborate the satisfactoriness of a model's assumptions for a particular population. This corroboration does not in general involve demonstrating the accuracy of assumptions. As we have seen, many assumptions of these models are idealizations. Instead, what is corroborated is that the population dynamics can be represented *as if* they possessed some feature they do not—asexual reproduction, infinite size, etc. Second, researchers sometimes empirically corroborate a model's prediction. In general what is sought is a qualitative fit. This demonstrates that the representation of natural selection's causal role is satisfactory. On the interpretation I am developing here, satisfactory representation requires both that the posited causal influence is present and that there are not unaccounted-for causal influences that obliterate the anticipated causal regularity. Beyond the empirical corroboration of idealized assumptions and qualitative prediction, the bulk of the attention in this research program is focused on exploration of the nature and extent of the focal partial causal regularity.

This research into the evolution of cooperation has featured in two philosophical debates that yield additional insight into their nature. First, there is a significant literature in philosophy of biology on the various approaches to accounting for cooperative and, especially, altruistic behavior of animals. This is the main type of evolutionary phenomenon addressed in the levels of selection debate. Sober and Wilson (1998) are influential advocates of the group selection approach as a "unified evolutionary theory of social behavior." They argue that cooperative behavior modeled using evolutionary game theory or kin selection should be acknowledged to involve group selection. Kerr and Godfrey-Smith (2002), in turn, show that an individualist perspective, such as provided by evolutionary game theory, is mathematically equivalent to a group perspective. These authors argue that each approach has its own "heuristic advantage," so it is useful to move between these perspectives. This debate draws attention to the question of whether there is in fact a uniquely right approach to modeling the evolution of cooperation, even for one specific instance of an evolved cooperative behavior.

Second, there is also a substantial philosophical literature on criticisms of evolutionary biology methodology that Gould and Lewontin (1979) disparagingly dubbed "adaptationism." As applied to behavioral ecology models like those considered here, the concern is that it is simply assumed that traits—including cooperative traits modeled using these approaches—are the direct products of natural selection. The possibility of a host of other influences, such as genetic and developmental constraints, is wholly neglected. In part, this criticism is accurate. I have motivated the idea that these models do focus on the influence of selection. But this is fine so long as there are other research programs addressing the other significant evolutionary influences. This debate thus draws attention to the partiality of behavioral ecology models of cooperation.

In summary, idealizations pervade behavioral ecology models of cooperation. Little attention is paid to the accurate representation of any one system, i.e., any single cooperative trait in a particular population. The aim instead seems to be to account for the causal role(s) of natural selection in producing cooperative behaviors. Accordingly, the focus is not on generating actual predictions, nor accurate representation, but partial, skewed representations. Finally, multiple partial causal regularities have provided insight into the evolution of cooperative behavior, as demonstrated by the different behavioral ecology approaches to modeling the phenomenon. It may be that multiple partial causal regularities provide insight even into individual instances of evolved cooperation.

3.1.2 Studies of Human Aggression

A reasonable response to my analysis in the previous section is to blame the heavy use of idealizations on the features of behavioral ecology in particular. The focus—the evolution of cooperative behavior—is an exceedingly broad phenomenon, encompassing a number of specific behavioral traits across a wide variety of species. Further, the modeling techniques brought to bear are in large part analytic, mathematical models. These features conspire to provide highly idealized treatments of general, partial regularities that are not applied in detail to any specific systems. So let us now shift our attention to a very different type of scientific project: investigations of specific human behaviors. Scientific approaches to human aggression and sexuality are the focus of (Longino, 2013). Here I discuss research on human aggression in particular, and my discussion relies heavily on Longino's study. This discussion will have a somewhat different character from above, since instead of one research program, I will discuss a variety of approaches to studying aggression.

Human aggression is a research focus in a range of behavioral sciences. Longino (2013) discusses several of these approaches in depth. Here I will focus on the four that deal most directly with specific human behaviors, including aggression. For the most part, this discussion follows Longino's analysis. Quantitative behavioral genetics employs twin studies to compare the relative influence of genetic variation and environmental variation on behavioral variation in a population. When used to study aggression, a proxy for aggression such as antisocial behavior is compared in twins raised together and raised apart. This is used to estimate the level of heritability for aggression. What Longino terms social-environmental approaches focus on establishing specific environmental influences on behavior, including aggressive behaviors. A common strategy is to seek correlations between specific features of individuals' social environments and a proxy for aggression. Molecular behavioral genetics, in turn, attempts to establish associations between specific genes and behavior. In this research, much attention has been directed to the role of genetic influences on serotonin metabolism for variation in levels of aggression. Finally, neurobiological approaches focus on the neural correlates of behavior. Here too serotonin metabolism is a focus, as well as the localization of brain activity concurrent with aggression.

Longino's analysis of the relationship among these different approaches to studying human aggression begins with a list of the broad range of possible causes investigated by one or more of these approaches. Possible causal influences on aggression include individual alleles, as well as the genome taken as a whole; intrauterine environment; anatomical and physiological characteristics; an individual's environment, the family's environment, and socioeconomic status (Longino, 2013, p.127). We can see from this that human aggression is likely to be a causally complex phenomenon, with diverse types of causal influences operating at different temporal scales, types of influences that are investigated by different fields of biological and social sciences. Indeed, Longino differentiates among the different approaches to studying human aggression according to which of these possible causal influences they can potentially measure, and which go unmeasured. It seems that unmeasured causal influences face a variety of fates: some are simply ignored, while some are assumed not to be causally relevant, despite indications from other investigations that they are. In either case, these causal influences are set aside in the pursuit of generalizations about other types of influences.

Longino's analysis of how these different approaches to human aggression differently parse the causal space thus demonstrates that each approach neglects a host of other causal influences. That neglect is achieved with the use of abstractions and, often, idealizations. Twin studies implicitly assume that the only causal history shared between twins reared apart is genetic, but this is an idealization, for they also share epigenetic influences and intrauterine environment, at least. Many neurobiologists assume that the neural substrate is genetically determined, but this is an idealization, for it has been shown that many neural structures are actually quite malleable. In the attempt to isolate the causal influence of one or more social influences on aggressive behavior, social-environmental approaches introduce the idealization that there are no significant genetic or physiological differences that interact with social differences. In light of all of this idealization, it appears that these approaches also aim to establish partial causal regularities, viz., information about how a particular set of causal factors influence the phenomenon under investigation. They abstract and idealize away from many significant features of the causally complex phenomenon of human aggression in service of this aim. This is so, even though these research programs all have the more concrete focus on human aggression in particular.

For this research, it is also interesting to consider the roles played by empirical investigation. Data is much more relevant to much of this work. All of the approaches to studying human aggression we have surveyed employ experimentation or other data collection techniques in order to corroborate or disconfirm specific hypotheses about causal influences on aggressive behavior. This is almost certainly due to the goal of accounting for the more specific phenomenon of human aggression, in contrast to behavioral ecology research into cooperative behavior of any kind, in any species. Yet, because of the idealizing assumptions used to limit the causal factors under consideration, the type of data considered by any one of these approaches is highly constrained. Accordingly, available data is limited to determining the nature of the causal regularity in question. Significantly less is done to corroborate idealized assumptions than for the behavioral ecology models of cooperation. Perhaps this is because those assumptions are implicitly acknowledged to be less well-founded and, accordingly, it is acknowledged that any of this research captures at best a partial truth about the influences on aggression.

This use of data suggests that, despite the focus on a more specific, limited phenomenon, little or no attention is directed toward increasing the overall accuracy of the representation of causal influences. Instead, these different strands of research are pursued largely in isolation from one another, each with its own body of idealized assumptions. This accords with the form of pluralism that Longino articulates, which might be called a methodological pluralism (Potochnik, 2013b). On that view, different approaches to a single phenomenon are irreconcilable, and each offers partial knowledge—knowledge that is incommensurable because of the different parsings of causal space. For this reason, the research is of limited predictive value. The accuracy of all of these approaches is severely constrained by their extensive use of idealizations. These idealizations, by and large, represent as inactive causal factors that are recognized to be not only relevant, but often even *interacting* with the focal causal factors. What this style of idealization does enable, though, is the promise of intervention. The representational aim of a partial regularity involving one or a few focal causes increases the potential of this research to facilitate influence on levels of aggression via control of one of its causes. Different partial causal regularities, and thus different fields of research, hold the potential for different kinds of interventions, be they neurobiological, hormonal, or social in nature.

In summary, idealizations are pervasive in these approaches to researching human aggression as well. Here the focus is a rather specific phenomenon: levels of aggression in human populations. But even so, these research programs limit themselves to establishing the causal role of certain types of influences. This is accomplished with various ways of generating empirical confirmation for specific hypotheses. Yet the extensive idealizations persist. Little attention is devoted to confirming the adequacy of these idealizations, and they often do a very poor job of reflecting the causal influence of neglected factors. No attempts are made to replace these idealizations with more realistic assumptions. There is a greater focus on accurate representation of a specific phenomenon than with behavioral ecology models of cooperation, but because the focus is limited to a subset of causal factors, the result is still partial, skewed representation, with quite limited predictive value.

3.1.3 Physical Sciences

We have now surveyed a couple of research programs in the biological and social sciences and seen that, despite the variety of their agendas, all are productively understood as the pursuit of partial causal regularities, facilitated by significant idealizations. It would be reasonable to attribute those characteristics to the fields of science from which my examples are drawn. After all, some have argued that these sorts of characteristics distinguish these fields from physics. Rosenberg (1994), for example, defends a view according to which physics and chemistry aim to describe reality, whereas the life sciences and social sciences merely aim to furnish us with tools for better controlling phenomena in these domains. I thus now turn my attention briefly to a few example research programs from fields of the physical sciences, in particular, climate change science, fluid mechanics, and astrophysics. Here I rely heavily on others' work, as my expertise with these fields is limited.

Climate change science is perhaps not naturally construed as an example of physical science. Nonetheless, for two reasons I want to start with a discussion of this research. Climate change research provides a nice contrast to the research discussed so far, insofar as it relies heavily on computational models and apparently does aim for accurate predictions, as I have argued behavioral ecology research into cooperation and studies of human aggression do not. It is well established that anthropogenic climate change has occurred and is continuing. A prominent focus of climate change research is now to ascertain the extent of future warming and the nature and extent of climate effects this will produce. Here the focus is not only limited to a single *type* of phenomenon, as is research on human aggression, it is limited even to a single phenomenon, that is, the changing climate on Earth. This system is expansive, and it is extremely causally complex. Winsberg (2010) demonstrates the value of computer simulations in grappling with such complex phenomena in general, as well as with climate change in particular. As indicated above, a prominent research focus is generating accurate predictions of future states of this system.

Multiple computer models are often developed to tackle a single predictive question. These models may differ in their assumptions, parameter values, and the type of causal processes they take into account. A primary strategy for assessing the predictions generated by multiple, different models is to ascertain what predictions are robust across a range of models with different assumptions and types of representations. This technique is called robustness analysis. The idea seems to be that demonstrating that a result holds across a range of different plausible assumptions and representations of the system indicates that it is not sensitive to any of those specific assumptions or representational choices, and is thus confirmed as a plausible prediction. The use of robustness analysis is a topic of philosophical controversy. Orzack and Sober (1993) criticize the idea that robustness is in general a guide to truth; Weisberg (2006, 2013) responds to this criticism by outlining how robustness analysis can provide a form of confirmation without direct empirical support. His analysis, though, is focused on the use of robustness analysis to ascertain what I would call partial causal regularities—how one or a few factors influence a recurring phenomenon of interest. Parker (2011) addresses the use of robustness analysis in climate science in particular, so her treatment is tailored specifically to the aim of accurate prediction. Parker is critical of robustness analysis's current ability to support predictive hypotheses. As she notes, it does not follow from this that climate policy decisions should be postponed until the science is better. Instead, if Parker is right about the limitations of predictive robustness analysis, climate science must be taken to have at least two main aims: first, to generate accurate predictions about future states of the Earth's climate, and second, to guide policy even in the absence of accurate predictions.

As this consideration of robustness analysis indicates, in this research as well, there is often an indirect relationship to empirical confirmation. Individual assumptions and causal hypotheses can sometimes be confirmed, but not thoroughly enough to identify one or even a few best models. When it comes to the main predictions, other methods of corroboration are required, which accounts for the significance of robustness analysis. The prominent use of robustness analysis suggests that, in this research program, predictions or accurate policy guidance are sought at the expense of accurate causal representation or explanation. Nonetheless, a specific set of partial causal regularities is sought here as well. The primary causal interest regards the relationship between human influences on climate—a particularly significant example is carbon emissions—and predicted future scenarios. A full account of the causally complex system, or even of the variety of relevant factors, is less important than an appreciation for the causal role of the factors over which we humans might exert control.

I now briefly consider the nature of two types of research in the physical sciences that others have used to exemplify the significance of idealizations. First, Batterman (2009) makes the case that idealizations are essential to a "full understanding" of some physical phenomena. One of his examples regards the hydrodynamical discontinuity that occurs when water drips from a faucet, and a single mass breaks into two or more droplets. So the attention here is on a highly general but precisely defined phenomenon. It is plausible that the same types of causal influences are active in determining the shape of the fluid at the point where it breaks (the singularity), namely the velocity of the fluid and curvature at the breaking point. Yet, as Batterman shows, this is "a difficult and complex moving boundary value problem" (p.434). What enables a solution are simplifying idealizations that can be made as the breaking point is approached, but not otherwise: first, that the water is a vertical line; second, that there is no acceleration due to gravity; and third, that axial and radial length are arbitrarily small. Batterman demonstrates how these idealizations can be justified by properties of the phenomenon in question. According to him, they serve a methodological role, for they enable exact solutions to the relevant equations, as well as an explanatory role, for they demonstrate why different fluids, dripping from nozzles of different shapes and sizes, have the same shape at the breaking point.

This is an essential use of idealizations in physics. This research is similar to behavioral ecology research into cooperation insofar as both attempt a highly general account of a broad phenomenon. In the process, both incorporate idealizations that enable the neglect of many influences on the phenomenon. Both also involve some attention toward justifying those idealizations. At least superficially, the considerations Batterman surveys that enable the simplifying assumptions are of a different sort than in the behavioral ecology case. They regard geometrical relationships, such as between the axial and radial extension of a water drop, and claims about causal significance motivated by considering the models themselves, viz., the structural equations, such as the claim that surface tension, viscous forces, and inertial forces are of equal importance. In contrast, we saw that idealizations in models of cooperation are often justified by direct empirical corroboration, or else left unjustified. A further similarity is that each type of research results in a generalization about a partial causal regularity. It seems limitations and exceptions are relevant in this case from fluid mechanics also. Batterman says, of the simplified treatment that enables application to a wide range of scales, "to a large extent and for a wide range of fluids, this turns out to be the case" (p.435).

Kennedy (2012) also argues in favor of the centrality of idealizations to research in the physical sciences, and in particular, astrophysics. She bases her argument on two examples, one of which is a model designed to provide the probability of observing a cometary maser (microwave amplification by stimulation emission of radiation) from the ground. As Kennedy describes it, this model was constructed in response to researchers coming up empty handed when they searched for maser emission from the comet Hale-Bopp. The model employed several idealizations, including the false assumptions that jets are the only cometary source of masers and that jets are located at either the pole or nucleus of a comet. The model identified geometric reasons for the absence of maser observations for Hale-Bopp. As Kennedy describes,

> The jet must be located along the line of sight of the observer in order to be detectable, the angle between the jets axis and the Sun-comet line must be small in order for the Sun to efficiently heat the hot spot and cause gas sublimation, and the angle between the normal and the Earth-comet line must be small in order to allow for detection of a maser. This sort of favorable geometrical setup only very rarely occurs during a random search (p.329).

In this case, idealizations are used to secure the model's computational tractability, and there seems to be little focus on justifying their well-foundedness. Interestingly, the aim of this research is explicitly construed as accounting for scientists' observations—or lack thereof—and to provide an account of the conditions required for desired future observations (of cometary masers). The result seems to be largely of theoretical significance, with little attention to direct empirical confirmation. It is clear that the geometric considerations are intended to apply quite generally and, it would seems, suggest a revision of research strategies in astrophysics along the lines of abandoning random searches for masers from the ground.

In this section I have surveyed three types of research in the physical sciences, and the conclusions I made about my earlier case studies apply to this research as well. Even in the small sampling of scientific research I have described so far in this chapter, there is an astounding diversity of aims and methods. The goals of these research projects vary from accounting for the causal role of one factor in producing a very general, heterogenous phenomenon in the case of the evolution of cooperation, to accounting for the action of certain types of factors in the more specific phenomenon of human aggression, to issuing specific predictions and guiding policy decisions in the case of climate science. A large variety of roles are played by data, but direct, empirical confirmation of a prediction is not a frequent occurrence. Many do make significant contact with the goal of illuminating partial causal regularities, although this too occurs in different ways. Variations include what types of causal factors are focal, as well as the principles governing the selection of those focal factors; whether the aim is to illuminate a general causal role, or causal action in more specific, set circumstances, or to treat causal influences so as to maximize predictive accuracy.

I have come nowhere near to cataloging the full diversity of the aims and methods of science. The purpose of this exercise was, first, to take a step back from the across-the-board generalizations I made in Chapter Two about the aim of illuminating partial causal regularities. Partial causal regularities are certainly more central to science than are laws of nature, but which are significant and the nature of that significance can vary. Second, it is remarkable that even with all of the differences among the varieties of scientific research I have discussed, all prominently employ idealizations. Lastly, all of this diversity begins to suggest that science does not proceed in lockstep toward truth, in the hopes of generating accurate predictions and fulfilling explanations. Something very different—and much messier—is going on.

3.2 Redefining the Aims of Science

In Chapter Two, I established that there are several intertwined reasons to idealize; that idealizations play a positive representational role; and that there is rampant, unchecked idealization throughout the scientific enterprise. In 3.1 I showed that those idealizations are present for different reasons, accomplish different ends, and different steps are taken to accommodate them. Yet they are not eliminated, nor are they minimized. One common view is that all of this idealization may be necessary, and that it might be here to stay, but that it results in representations that are lacking in various ways. Accordingly, the view goes, we must look for a subsequent step, a way to connect these idealized representations to the successful pursuit of the aims of science, whether those are prediction, empirical confirmation, explanation, true representation, etc. The textbook version of this view would hold that science aims for truth, and so idealized representations must be de-idealized in order to be useful. It seems that Odenbaugh and Alexandrova (2011) assumes something like this view, for they argue that without the removal of all idealizations—complete de-idealization—we have "no ground, beyond that of our background knowledge that informed the model, for claiming that the model specifies a causal relation" (p.765). Odenbaugh and Alexandrova conclude that even the use of multiple models, with different idealizations, cannot yield the description of a causal mechanism. Thus, they claim, this does not allow for the confirmation of models, nor can it generate explanations.

Other versions of this view do not hold de-idealization to be necessary, but still anticipate the need to bridge the gap between idealized models and the traditional aims of science. Wimsatt (2007), for instance, argues that idealized, "false" models can be used to produce "truer" theories without recourse to de-idealization. Similar to Odenbaugh and Alexandrova's concern with causal description and explanation, Rohwer and Rice (2013) argue that at least one purpose of idealizations, namely the investigation of general patterns across heterogenous systems, prevents the accurate description of causal factors and, thus, the formulation of explanations (though they hold that resultant model may still be explanatory). These views all endorse the continuing practice of idealization, but they also hold idealized models to be somewhat distant from the traditional aims of science. They accordingly explicitly or implicitly commit themselves to an intermediary step of some kind between idealized representation and achieving the aims of science. On this strategy, even though idealized representations are of scientific value, they are not sufficient to provide adequate explanations, trustworthy predictions, causal representation, etc.—at least not by themselves.

One might instead take a very different approach to reconciling idealizations and the aims of science. The observation of rampant and unchecked idealization in science, and the distance between idealized representations and traditional articulations of the aims of science, might be seen as grounds for concluding that those traditional articulations of the aims of science are incorrect. On this alternative approach, nothing has gone wrong with or is lacking from idealized representations, and no intermediary step is needed for idealized representations to achieve the aims of science. Those aims just stand in need of clarification. This is the tack I take here. I defend the view that, in an important sense, science is not after the truth. That is, much of science resists interpretation as successive approximation or increasingly accurate representation. A result of rampant, unchecked idealization is that the immediate products of science are not things we believe to be true.

3.2.1 The Epistemic Value of Understanding

Wimsatt (2007) rightly points out, regarding idealized models, that "unless they could help us do something in the task of investigating natural phenomena, there would be no reason for choosing model building over astrology or mystic revelation as a source of knowledge of the natural world" (p.101). This must be right; models, even with all their falsity, must get us somewhere that, for instance, mystic revelation does not. At issue is how to understand what idealized models are helping us accomplish. In my view, false models are not a means to truer theories, as Wimsatt believes, but themselves accomplish the end goals of much of science.

Above I discussed how behavioral ecology, molecular genetics, and climate science all persist in the use of idealizations. Even in these few examples, the idealizations are present for different reasons and different steps are taken to accommodate them. But in each case, there is apparently little interest in attempting de-idealization. The goals of using idealized models vary, but in none of these examples is the goal to attain a more accurate representation of a specific target system.

I argued that the goal of the behavioral ecology research into cooperation is to represent general causal dependencies without representing or predicting any features of particular target systems. In the molecular genetics research, the goal is instead to represent specific causal relationships—between individual genes and phenotypic traits—at the expense of generalizations about these factors' causal roles as well as representation of all the other causal factors as work. Finally, for certain computational models of climate change, the goal is accurate prediction without any care for accurately representing the causal structure. These examples and, I believe, much of scientific research resists interpretation as aiming for the truth, successive approximation, or increasingly accurate representation.

A first step toward a more promising conception of the aims of science is provided by Elgin (2004). Elgin is also impressed by how many scientific laws, models, and theories diverge from the truth in various ways. Her aim is thus to show how these scientific products can be epistemically acceptable without being true. In her view, this is because they produce understanding. So, according to Elgin, a focus on truth should be replaced by a focus on understanding. She says,

> I take it that science provides an understanding of the natural order. By this I do not mean merely that an ideal science *would* provide such an understanding or that in the end of inquiry science *will* provide one, but that much actual science has done so and continues to do so (p.114, emphasis in the original).

Elgin's position results from accepting today's actual science as a successful venture, then looking to see what this science accomplishes. Rather than make excuses for the myriad ways in which our science falls short of truth, Elgin reconsiders the nature of epistemic success. This is reminiscent of the approach to this book outlined in Chapter One.

Elgin thus conceives of the epistemic success of science as understanding, which she distances from truth. Truth does not disappear from this picture; instead, it functions as a threshold concept. According to Elgin, a claim must be "true enough," and this requires that any divergence from the truth be negligible, that is, "safely neglected." Whether this is so depends on the function a claim plays in an argument, explanation, or theory—or, one might add, in a model or other representation. One of Elgin's examples is Snell's law, which governs the angle of refraction of light when it passes from one medium to another. Elgin notes that Snell's law is only true of optically isotropic media, but it is true enough of media that are nearly isotropic, which includes a wide range of media in which physicists are interested.

This suggests an amendment that must be made to Elgin's view. Whether a claim is true enough depends not only on its function in, say, a representation, as Elgin notes. It also depends on the purpose to which that representation is put. Elgin points out that Snell's law applied to anisotropic media is of limited use "if we are interested only in the path of a particular light ray," but is useful "if we are interested in optical refraction in general" (p.118). This divergence in whether or not Snell's law is true enough cannot be accounted for solely by referencing the role played in the law by the assumption of isotropic media; it requires reference to the researchers' different aims, viz., what they intend to get out of an application of the law. This is a minor expansion of the recognized influences on a claims' epistemic acceptability, but it hints at a much more significant shortcoming of Elgin's view that will be addressed below.

For now, notice that Elgin's proposal of distancing the epistemic success of science from truth is not as radical as it might at first appear. Elgin clarifies her position by adding, "I do not then claim that it is epistemically acceptable to *believe* what is false... Rather, I suggest that epistemic acceptance is not restricted to belief." She continues, "understanding is often couched in and conveyed by symbols that are not, and do not purport to be, true" (p.116, emphasis in original). Epistemic acceptance, then, is supposed to be a broader category than belief, and understanding is supposed to be sometimes "couched in" symbols that are not true. Elliott (2013) argues that scientists adopt a number of different cognitive attitudes toward the products of science, and he follows Cohen (1992) in distinguishing between accepting and believing a body of content, such as a hypothesis, theory, model, or other representation. It seems Elgin takes a similar approach, and then yokes understanding to acceptance—or, as she puts it, *epistemic* acceptance—rather than to belief.

Distancing scientific success from truth in this way is a natural first step toward redefining the aims of science in light of rampant and unchecked idealization. Elgin argues that such "felicitous falsehoods figure in cognitive discourse not as mistaken or inaccurate statements of fact, but as fictions" (p.123), and that these fictions facilitate understanding, for they "impose an order on things, highlight certain aspects of the phenomena, reveal connections, patterns and discrepancies, and make possible insights that we could not otherwise obtain" (p.127). She gives the example of drawing a smooth curve and treating the data's deviation from the curve as error or noise. This view nicely accommodates the positive representational role that I have defended for idealizations. In Chapter Two I suggested that idealizations are fictions in the sense that they represent system(s) as if they have some feature they do not, to the end of capturing a partial causal regularity. This accords with Elgin's account of how felicitous falsehoods facilitate understanding. My answer to Wimsatt's challenge of articulating the epistemic value of idealized models above, say, mystic revelation is that these provide understanding. Idealizations are thus not preparatory to truer theories, but contribute directly to understanding.

*   *   *

Elgin's focus on understanding is a move in the right direction, namely away from a truth-centric conception of science and toward a conception of science's aims that better accords with rampant and unchecked idealization. However, the resulting view is still too conservative of traditional conceptions of the aims of science. Elgin says, of the ideal gas law:

> The model is illuminating though, because we understand the properties of real gases in terms of their deviation from the ideal. In such cases, understanding involves a pattern of schema and correction. We represent the phenomena with a schematic model, and introduce corrections as needed to closer accord with the facts (p.127).

This sounds awfully like the traditional view that I outlined above, where idealizations are seen as distortions, to be overcome or circumvented in the pursuit of truth, or for Elgin, to "accord with the facts."

Part of the difficulty stems from the fact that Elgin speaks as if all of science generates claims, claims that figure into arguments, explanations, or theories. We saw in Chapter Two and in 3.1 above that this is wrong. Model-based science may proceed largely independently from theory, and there is a diverse array of scientific products, many of which have little or no relationship to theory or explanation. This broader conception of scientific products requires that Elgin's definition of "true enough" be significantly revamped. Her proposed standard of negligible divergence from truth, taking into account a claim's role in an argument, explanation, or theory, is still too truth-conservative. This can be seen from Elgin's treatment of the ideal gas law. For many scientific projects, a schematic model will suffice. Such projects would be hindered, not furthered, by closer accordance with the facts. Elgin later says, "if, for example, evidence shows that friction plays a major role in collisions between gas molecules, then unless compensating adjustments are made elsewhere, theories that model collisions as perfectly elastic spheres will be discredited" (p.129). But this too is wrong. There may be perfectly good reasons—including epistemic reasons—to continue to model gas-molecule collisions as if they were collisions among perfectly elastic spheres, even if the *theory* that they are similar to perfectly elastic spheres is discredited. In these cases too, partial causal regularities may be revealed, that is, regularities in the causal contribution of factors other than friction. We have seen that idealizations play a much broader role than simply standing in for factors that are not difference-makers. There are quite many reasons to represent as-if.

There is a second regard in which Elgin's account is too conservative of traditional conceptions of the aim of science. Elgin focuses exclusively on science's purely epistemic role, viz. the production of understanding, but science plays a wide range of roles. Some are epistemic, such as understanding; others non-epistemic, such as action within a short timespan; and still others seem to involve both epistemic and non-epistemic elements, such as accurate prediction (cf. Elliott, 2013). Instead of a single successor aim for truth, we should thus expect a variety of scientific aims, which suit science to the range of roles it plays. The list of science's roles must be open-ended, for science is a continually creative process. The procedures, as well as the products, are always in development. Nonetheless, some generalizations can be made about these aims. First, many relate in one way or another to establishing and employing partial causal regularities, as I set out in Chapter Two. Second, all of these aims further cognition, action, or both; indeed, significantly, they further *human* cognition and action, as a science by and for human beings should.

These criticisms of Elgin's view demonstrate the features that a reconstrual of the aims of science must possess in order to succeed. First, it must be acknowledged that science has a variety of aims. Understanding replaces truth as the ultimate purely epistemic aim of science. However, it is not the ultimate aim of science simpliciter. There are important non-epistemic aims of science, and likely aims that involve both epistemic and non-epistemic elements as well. Second, it must be acknowledged that what best facilitates understanding is not determined solely by the relationship between a representation and the world. What best facilitates understanding—and thus, in Elgin's terminology, what qualifies as "true enough"—depends also on a range of considerations about scientists themselves. Those considerations include, prominently, the scientists' particular research interests, but also their cognitive faculties and psychological characteristics, their temporal and spatial location, etc. This move resolves the difficulties identified in Elgin's interpretation of the role of the ideal gas law and, it stands to reason, of other heavily idealized representations. Idealizations, no matter how little they resemble the systems they represent, may nonetheless facilitate an understanding of those systems in the context of a science that is, ultimately, a human creation.

Each of the intertwined reasons to idealize identified in Chapter Two (in particular 2.2.1), it now emerges, is valuable for its ability to facilitate understanding or some other aim(s) of science. Here I continue to focus on idealizations' role in the production of understanding, the narrowly epistemic aim of science; in the next section, I widen my focus to also consider the relation to other aims of science. As indicated in Table **??**, some reasons to idealize are temporary and others are permanent. Which is the case is determined by whether an idealization facilitates understanding for a reason only relevant to the current stage of scientific inquiry or of enduring relevance. For instance, an idealized model may better facilitate understanding because of a limitation in our current computational powers or because of an absolute computational limit. Other temporary reasons to idealize include what modeling approaches happen to be close at hand, or familiar to the researcher(s); what technique happens to get initial traction with a system of interest; and what modeling technique best positions the researcher(s) to transition to a less idealized, more successful approach down the road. But many reasons to idealize are due to idealizations' enduring ability to facilitate understanding. Reasons that justify permanent idealizations include cognitive limitations of the researchers or their audience such that an idealized model best leads to understanding; the depiction of a partial causal regularity; and a research program with a limited focus furthered by the idealization in question.

A few elements of how these intertwined reasons to idealized facilitate understanding bear mentioning. First, notice that some reasons to idealize—both temporary and permanent—are due to features of scientists or, sometimes, the audience for their science, including other scientists, policy-makers, or the lay public. This is at is should be, for we have seen that what best facilitates understanding depends in part on considerations about scientists themselves. Second, we should expect there to be an open-ended list of reasons to idealize. Idealizations can contribute to the production of understanding in many ways, depending on the features of phenomena of interest, as well as the background concerns and characteristics of the scientists. The wide range of reasons to idealize results, ultimately, from limited human cognizers grappling to understand our causally complex world.

Finally, it is significant that an idealization may be radically untrue, that is, quite different from the true state of affairs, but nonetheless facilitate understanding. Elgin (2004) points this out. She discusses the example of assuming mutually disinterested agents, an assumption that appears in a range of social science contexts. Another example is the common assumption in population genetics that a population is of infinite size. An even stronger point is that a radically untrue idealization may facilitate understanding *in virtue of* its distance from the truth. As Strevens (2009) argues, a radically untrue idealization can facilitate understanding when the idealized factors are irrelevant by advertising their unimportance. But in my view this does not exhaust this style of idealization's contribution to understanding. Above I established that idealizations can facilitate understanding not only in virtue of their relationship to the world, but also in virtue of their relationship to those seeking to understand the world. This is so for radically untrue idealizations as well. Assumptions like mutually disinterested agents and infinite population size can also serve to advertise not that a factor is *causally* irrelevant, but that a factor is irrelevant to *the current research focus*. This reason to idealize is particularly useful in the pursuit of partial causal regularities.

At first glance, this account of the epistemic aim of science might seem to offer a terribly subjective standard for success. There is one subjective element of the account. The view that understanding is the ultimate epistemic aim of science posits a subjective standard for epistemic success, for the satisfaction of that standard is dependent on the features of the practitioners of science. This is because the features of scientists help determine what best facilitates understanding. But this element of subjectivity is unproblematic. Indeed, it is to be expected for an account of the scientific enterprise that has been developed by limited and historically located human beings. What best meets human requirements—including our requirements of knowledge—depends in part on the features of humans. Other forms of subjectivity that would be problematic do not apply. One need not rely on the subjective experience of an "aha!" moment in order to judge whether a representation facilitates understanding. Instead, objective reasons guide that judgment. A representation must be accurate in the anticipated respects (cf. Elgin, 2004), and the respects in which it is inaccurate must each be justifiable by an increase in understanding. The relative judgment of whether an inaccuracy—an abstraction or, especially, an idealization—increases understanding is based on objective features of the world and of scientists themselves. This is recognizable of each the reasons to idealize I have outlined. Limited computational power and absolute computational limits; the limitations of human powers of cognition and the limitations in one researcher's training; the existence of a partial causal regularity and the existence of specific research interests: these are all objective features of the world under investigation and the scientists leading that investigation that together determine what best generates human understanding.

* * *

There is a startling consequence to this conception of the aims of science. Namely, it unseats truth as the ultimate epistemic aim of science and, thus, introduces the possibility that the products of science are not things we believe to be true. There is, of course, *something* true, or rather "true enough," about the products of scientific research. They generate understanding and are accordingly on epistemically firm footing, whereas mystic revelation is not. But the threshold for "true enough," on my reconstrual of this requirement, is quite low and is also relative to the features of scientists and their specific projects. In my view, science simply is not after the truth.

Although this may be a startling consequence, it is not problematic when properly understood.

To begin with, the idea that science does not aim at truth, but rather understanding, should not be confused with a version of antirealism. The issue of realism versus antirealism is nearly orthogonal to the present discussion. In its most abstract formulation, that debate regards whether our best science has achieved epistemic success, or alternatively, whether science aims to achieve epistemic success (Chakravartty, 2013). I have not challenged the idea that there is an epistemic aim to science; my criticisms are of the idea that this aim is best articulated as truth. The considerations raised here do, then, draw into question any version of realism that articulates that epistemic success as truth or even approximate truth. But other articulations of realism are readily available. In turn, that I have posited an epistemic aim for science is opposed to purely instrumentalist versions of antirealism. In the first section of this chapter, we saw that science is in part after predictions, but many of its activities belie that description. Any accurate construal of science must account for both its epistemic and non-epistemic aims. To state the point generally, the position developed here constrains what form of a realism or antirealism one might take, but it does not pretend to settle that issue.

Yet my position that science does not aim for truth does seem to be susceptible to some concerns similar to ones that have been raised against antirealist positions. One such concern is that denying science truth leads to an untenable distinction between ordinary everyday truths and scientific claims of the same style that I must deny are true. For example, it is true that, at the time of writing this sentence, I live in the city of Cincinnati. Is it not also true that, say, humans and chimpanzees share a common ancestor? As this illustrates, many downstream products of science are indeed true—or at least as true as analogous commonsense claims. But in my view, this extends only to the rather simple, concrete claims that sometimes result from successful scientific ventures. Quite many other claims and representations generated in science are not straightforwardly true. Generalizations that capture partial causal regularities have exceptions, and an even broader range of representations incorporate untruths in the form of idealized assumptions. These are vehicles of scientific understanding, but they are not straightforwardly true. Truth in science is limited to the particular and to the partial, viz., the true *enough*.

This position is consistent with Longino's (2001) suggestion that the measure of scientific success should be broadened from truth to what she calls "conformation."[2] Longino proposes this as an umbrella term for the epistemic success of any scientific content, which comes in degrees and in different respects. On her view, conformation is more precisely defined locally, in different arenas of science. Truth is a form of conformation, as is isomorphism, homomorphism, similarity, etc. Considered in its relationship to this idea, my proposal is that conformation, viz. epistemic success in science, is not often defined as truth, nor even accurate representation (or isomorphism, homomorphism, etc.). Instead, the epistemic success of science most often consists in partial similarity in some very limited respect.

One may wonder how different this position really is from existing treatments of idealization. After all, two of the three practices of idealization that Weisberg (2007a) discusses need not involve de-idealization. Rohwer and Rice (2013) go further by also endorsing idealizations that omit important causal influences. Yet the position I endorse here additionally elevates the role of continual idealization, and it thereby requires revamping the very aims of science. Weisberg's minimalist idealization and multiple-models idealization may not involve a promise of de-idealization, but each does provide an alternative way to keep idealizations in check. Minimalist idealization does not idealize any central causal factors, while multiple-models idealization involves the use of different models that more accurately represent what the others idealize. Weisberg thus allows for continual idealization but not unchecked idealization. In contrast, I have suggested that it is commonplace to have unchecked idealization even of important causal influences. Rohwer and Rice's hypothetical-pattern idealization looks very much like the practice I identify in behavioral ecology models of cooperation. However, in contrast to that view, I suggest that these idealized models are full-fledged, successful products of science. Rohwer and Rice deem these models explanatory, but they argue that they fall short of providing explanations. They thus account for idealization, yet keep it at a remove from full scientific achievement. I take the opposite approach by reconsidering the nature of scientific success, viz. the aims of science, in light of the pervasive practice of unchecked idealization.

[2] Though this word is only one letter distant from "confirmation," a term also of scientific significance, there is no relation between the concepts.

3.2.2 Separate Pursuit of Science's Aims

If the position developed above is right, the epistemic aim of science is not truth, but understanding. Yet, as I have already acknowledged, there are a variety of aims of science, both epistemic and non-epistemic. Traditionally appreciated aims include (at least) accurate prediction, explanation, and representation. Other aims of science have recently received increasing attention. These include providing information to guide policymaking (Douglas, 2009b); action within a short timespan (Elliott, 2013); and facilitating the public uptake of scientific knowledge (Elliott, 2011). There are surely many other aims and, indeed, many other existing articulations of these and other aims. My project here is not to delineate the range of scientific aims; above I suggested that we should expect an open-ended list of aims. My goal is instead to examine the relationship among the various aims of science. In particular, I suggest that it is the norm for the pursuit of one aim to occur at the expense of others. Successful pursuit of one among the various aims of science generally inhibits success with other aims. Accurate prediction is achieved by tools poorly suited to explain; the aim of quick action is at odds with full causal representation; etc. At root, this is because the different aims of science are furthered by different means. This too traces back to the centrality of idealizations in science.

Notice first that the diversity of scientific aims is linked also to a diversity of cognitive attitudes toward the products of science. Recall from above that embracing understanding, instead of truth, as the epistemic aim of science requires shifting from a focus solely on belief to the broader concept of acceptance. This is because many aids to understanding are not things we believe to be true. I followed Elgin in phrasing the alternative, broader concept "epistemic acceptance." But acceptance, like the aims of science themselves, comes in many varieties. Elliott and Willmes (2013) define acceptance as follows: "S accepts that h, iff S presupposes h for specific reasons in her deliberation" (p.5). This definition can yield the subspecies of epistemic acceptance by restricting the relevant specific reasons to the furtherance of understanding. In this way, one can generate other varieties of acceptance by focusing on other, specific aims of science. For instance, acceptance as predictively useful amounts to presupposing for purposes of predictive value (alone). It is clear that one form of acceptance need not entail another. The value of an assumption in the production of understanding, for example, suggests nothing about its predictive usefulness.

Just as one form of acceptance need not entail another, the pursuit of one aim of science need not contribute to other aims. Indeed, something stronger is true: success with one aim often inhibits success with other aims. Science as a whole employes a variety of tools to achieve, e.g., predictions, explanations, causal representations, and the basis for action. What suits a tool to further one of these aims does not well suit it for the other aims. For example, one method used to generate predictions is the analysis of a variety of models with competing assumptions, called robustness analysis. This is a common method in climate change modeling (Parker, 2011). But none of those models are expected to accurately represent the causal influences on climate, nor to explain climate change. The tool of robustness analysis helps achieve one aim, but there are many other aims to which it does not contribute.

This division in the pursuit of different aims of science is due in part to the widespread use of idealizations and the variety of purposes for which idealizations are used. Recall from above that whether an idealization furthers understanding depends on the specific goals of the research. This is why, e.g., Snell's law is appropriate to apply to anisotropic media when researchers want to understand optical refraction in general, but not when they want to understand the path of a specific light ray. More broadly, whether an idealization furthers *any* given aim of science depends on the specifics of that aim. Snell's law may help us understand optical refraction, but it is too idealized to give precise predictions of light's path of travel in anisotropic media. For that we need a different tool. This stands in contrast to the example just above of making climate predictions with the use of several idealized models, models that do not help us understand climate change.

The idea that different aims of science must be pursued separately is evocative of Cartwright's (1983) study of how laws are either false and explanatory or else predictive (but not both). Elsewhere I have also defended the view that the aims of explanation and prediction push in different directions [cit. omitted]. The separate pursuit of scientific aims is also related to the view that there are tradeoffs among the desirable features of models, such as their generality, precision, and realism or accuracy (Levins, 1966; Odenbaugh, 2003; Matthewson and Weisberg, 2009). This is the idea that increasing a model's generality, for example, is achieved by decreasing its precision or accuracy. Different features of models, such as greater generality or greater precision, will differently position those models to contribute to particular representational or predictive aims. In my view, the nature of the selected tradeoff reflects the purpose to which a model or other scientific product is put. Isaac (2013) similarly argues that models have specific functions, such as prediction and use in policymaking, and that many of these functions are best satisfied by abandoning the goal of realistic representation.

This conception of the aims of science in conflict requires one amendment. Recall from above that idealizations can facilitate understanding of a phenomenon by demonstrating that some of its features—including causally important features—are nonetheless unimportant to the current research focus. Which features should be understood is, thus, relative to a specific research focus. This means that an understanding of some features of a phenomenon may be purchased at the cost of misunderstanding, i.e. misrepresenting, other features. So, for example, an evolutionary game theory model may demonstrate the role of natural selection in producing cooperative behavior, while occluding the role of non-selective and non-evolutionary influences. These might include, for example, specific genetic influences and alternative, non-evolutionary influences like learning. In the context of other research programs, such as those with a population genetic focus or an epigenetic focus, understanding these other features of cooperative behavior will move to center stage. This shows how the epistemic aim of understanding can, in itself, motivate different scientific products that are appropriate for different research focuses. It seems the same is true for other scientific aims. For example, the prediction of different features of a phenomenon often must be accomplished by different means. Accordingly, not only do the aims of science conflict, but so too do deployments of a single aim, including even the narrowly epistemic aim of understanding.

There are two main reasons for the tension among different aims of science and different deployments of a single aim—reasons, also, for the variety of idealizations' contributions to the aims of science. These are the complexity of phenomena of scientific interest coupled with the limited powers of human cognition and action. The complexity of phenomena investigated in science is by now well appreciated; see, for instance, (Dupré, 1993; Cartwright, 1999; Mitchell, 2003; Strevens, 2006; Wimsatt, 2007). All of the scientific aims discussed in this paper are valuable for their ability to further human comprehension and control of this complex world. Their furtherance is, thus, relative to the limitations of human cognition and action.

Guiding policymaking and action are obvious examples, but prediction is similarly shaped by human objectives. Faced with the complex phenomena that are the norm in scientific investigation, scientists must choose which features of phenomena to focus on successfully predicting. Scientific explanation is also influenced by human limitations. Explanation is in furtherance of human comprehension, and its features are crucially shaped by this goal [cit. omitted]. Finally, if I am right that the epistemic aim of science is understanding, then this too is in service to, and shaped by, the particularities of human powers of comprehension. The limited powers of human beings and especially of our cognition, when faced with incredibly complex phenomena, require the focus on one particular scientific aim (at a time) to the exclusion of others. In order to successfully predict, or to represent a certain element of the causal structure, or to provide quick guidance for policy, one must sacrifice other aims.

It is enlightening to consider the nature of the position most directly opposed to my view of the conflicting aims of science. The opposed position is that the same scientific products individually further all aims of science—at least accurate representation, explanation, and prediction, and possibly also including more human-centric aims like providing grounds for policymaking. If this were the case, a single, best model would simultaneously offer the best causal representation, the best explanation, and the best predictions. I suspect that this kind of view is motivated by an implicit commitment to the idea that scientific products are true. If that were generally the case, then one could expect individual scientific products to play all these roles. Representations that are true in all important regards would provide a sufficient causal representation, be explanatorily unimpeachable, and ground accurate predictions. But science does not aim for truth, or so I have argued here. The aims of science are accordingly in tension.

This view of opposition among scientific aims also conflicts, though less directly, with views that relate confirmation and prediction to explanation. One example of such a view is inference to the best explanation; another is Douglas's (2009a) argument that focusing on the aim of prediction sheds light on approaches to explanation. As Cartwright (1983) has shown, predictive models often fail to be explanatory. Additionally, we hold on to explanations, like those provided by evolutionary game theory models of cooperative traits, even when their predictions fail [cit. omitted]. Moreover, we do not expect our explanations to be true, since as I discussed above, the representations that provide them are not true. Emphasizing the separate pursuit of these different scientific aims also undermines the appeal to multiple aims simultaneously—often prediction and explanation—that commonly goes unnoticed in philosophy of science. Consider, as just one example, that Odenbaugh and Alexandrova (2011) articulate their focus as the confirmation of "empirical hypotheses that later figure in explanations of particular... phenomena" (p.758). They see the confirmation of hypotheses—presumably by the accuracy of the predictions those hypotheses ground—to be preparatory for explanatory work. In contrast, I suggest that these aims of science are in fact opposed.

The conception of the relationship among scientific aims that I have developed here helps make sense of a feature of science that would otherwise be puzzling. There is often a proliferation of different approaches to studying the same phenomenon, approaches that apparently are in conflict with one another. For instance, Longino (2013) surveys an astounding variety of approaches to the study of human behavior, and in particular, human sexuality and aggression. She takes a pluralist stance, according to which each approach provides knowledge, and in her view, this knowledge is incommensurable across approaches. In Longino's view, multiple, incommensurate items of knowledge are possible because each approach differently defines the exact phenomenon under investigation and differently parses the space of causal influences. The variety of scientific aims and of specific deployments of individual aims makes sense of this practice. If accurate representation of, e.g., the full suite of causes for human aggression were the uniform aim, then different research results must be reconcilable (even if they could not be represented in a single model), or else one or more of them must be wrong. But with different aims of representation, prediction, policy guidance, etc., incommensurable findings that are in some sense about the same phenomenon may equally be successful science. This also accounts for the persistence of deep disagreements about fundamental principles within otherwise functional fields of research, e.g. as observed in population biology [cit. omitted]. Each approach may be successful given its specific aims, and no approach succeeds in addressing all the relevant aims.

Faced with a variety of scientific aims and the possibility of multiple deployments of a single aim, one might wonder what determines which aims are served by individual scientific projects. If there are many different, potentially applicable scientific aims, by what standard do we judge the success of individual scientific products? Some generalizations can be made about what determines the applicable aims. First, as outlined above, the particular research focus directly influences the features of a phenomenon that are central to understanding the phenomenon. The research program also influences the broader determination of the importance of prediction, explanation, action, etc. Second, a related influence is the features of the practitioners of science. We have seen that the aims of science reflect the particularities of humans, especially our cognitive limitations. The applicable aims are also influenced by the features of researchers themselves. Third, sometimes features of the target audience are also relevant. One example of this influence is the significance of whether the scientific product is intended for other researchers or policy-makers. When it comes to determining the aim by which a particular scientific product should be judged, the applicable aim can often be ascertained by charitable interpretation of the research conducted, including its setup, the conclusions drawn, and its research and social contexts. There may nonetheless be ambiguity in what aim is pursued and, thus, by what standard the work should be judged. Even more common is a mismatch between the pursued aim and conclusions drawn from the research by other researchers or, especially, by the popular media. Longino's (2013) study of the sciences of human aggression and sexuality nicely illustrates this as well.

3.3 Cautions for Reading Metaphysics Off Our Science

The rise of naturalistic philosophy, viz., philosophy that accords some or even total authority to empirical sciences, has been accompanied by questions about the relationship between science and metaphysics in particular. Attempts to naturalize metaphysics aim to demonstrate how our best scientific findings should inform our metaphysical commitments. In contrast, I have already defended two positions that drive a wedge between scientific practice and specific metaphysical implications. In Chapter Two I argued both that science does not uncover laws of nature and that an account of how the concept of causation is employed in science does not give insight into the metaphysics of causation. In Chapter Four I defend a similar conclusion for levels of organization. Here I address directly the matter of the metaphysical significance of our best scientific findings.

A few examples of approaching metaphysics with an eye toward science are (Dupré, 1993; Ladyman and Ross, 2007) and essays collected in (Ross et al., 2013); I engage with these examples below. These are but a few of many recent attempts to draw a variety of connections between science and metaphysics, and much effort has gone into articulating and analyzing versions of naturalized metaphysics. (Indeed, (Ross et al., 2013) is a nice survey of some of this work.) Let me be clear at the outset of this discussion that I do not pretend to settle the issue of the proper relationship between science and metaphysics here. My remarks may well not apply to some approaches to a naturalized metaphysics, and they may be consistent with other approaches. I do, however, believe that caution is in order when one examines science for metaphysical import, and that some attempts to naturalize metaphysics have been insufficiently cautious. Several of the ideas I advocate in this book can be used to motivate greater caution.

In this chapter I have defended the ideas that there is a variety of scientific aims, which are best pursued separately, and that the epistemic aim of science is not truth but understanding. Each of these views curtails science's metaphysical import. First, if science is not in the business of producing truths, then it is poorly positioned to uncover metaphysical truths, or even directly inform them. Granted, as discussed above, some broad truths emerge on the basis of our best science. However, such knowledge is not generally of the kind that supports metaphysical inferences. The truths that science provides are particular, viz., limited in scope and about rather concrete affairs, or they are partial, viz., have exceptions or elements of untruth. Nor is understanding, the ultimate epistemic achievement of science, well suited to ground metaphysical conclusions. Understanding is yoked to particularities of our epistemic position and human psychology, two limitations that a successful metaphysics must escape. Second, that science has a variety of aims, best pursued separately, is similarly problematic. This undermines the idea that scientific products are converging—or will at some future point in time converge—and provide a unified account of the world. A diverse array of tools for piecemeal prediction, explanation, policy guidance, etc., has little metaphysical import.

There is also a broader argument to be had against science's direct metaphysical import. As discussed in Chapter One, science is a human endeavor. It is a tool, or more accurately several tools, designed by humans to further our cognitive ends and our ability to exert influence on our world. If I am right that this indelibly shapes the character of science, then science is simply not objective in the right way to play a metaphysical role. Surely any metaphysics aspires to escape the particularities of human concerns and cognitive abilities. I address the topic of how human features and values enter into science and the resulting limits of science's objectivity more fully in Chapter Six. All of this suggests that the process and products of science, including scientific understanding, should not be used as a window into reality, at least not of the sort metaphysicians seek.

This view will become clearer by examining particular ways in which science has been used to motivate metaphysical conclusions. Dupré (1993) suggests that, even though metaphysical presuppositions ground the enterprise of science, scientific findings can and should be used to corroborate or undermine those assumptions. Metaphysical positions on the chopping block include essentialism about natural kinds, reductionism regarding scientific theories as well as regarding entities, and determinism. In this book I defend an account of science that is friendly, and in certain regards even similar, to Dupré's view of science. But the implications I see for metaphysics are different in kind. In the end, Dupré claims that his conception of scientific disunity implies the failure of all the metaphysical theses at issue: determinism, essentialism about kinds, and reductionism. In contrast, I think one can consistently recognize the features of science Dupré attends to and yet still embrace those metaphysical positions.

Consider, first, natural kinds. Attention to science might demonstrate that essentialism about natural kinds is not reflected in scientific practice. And then, Dupré is free to find inspiration in that for a radically altered conception of metaphysical kinds. But that thesis is not read off the science, for science is not in the business of telling us such things. The case of reductionism is more nuanced. In my view, Dupré is right that scientific practice undermines reductionism about scientific theories (see Chapter Four), but this is a thesis about the enterprise of science itself.

The metaphysical question of the reduction of entities that he also addresses, in contrast, is an issue distinct from scientific practice. Here a position might be informed by our best scientific findings, such as the failure of science to uncover reduction relationships among many entities (see Potochnik, 2010). But one might instead reasonably maintain metaphysical reductionism regarding entities. In that case, one might attribute the failure of science to hand us reduction relationships to science's many competing goals. Perhaps when reduction relationships are not uncovered, this is because that is not the scientific aim, or at least not the most important aim. The state of our science thus does not arbitrate for us the question of metaphysical reductionism.

Ladyman and Ross (2007) emphasize science's ability to get at the "objective character of the world." Their criticism of classical metaphysics is, in part, based on the idea that when common sense and science conflict, commonsense ideas must be jettisoned in favor of science. On their view, fundamental physics is also privileged over all other fields of science. Ladyman and Ross endorse an exceptionless principle to this effect; they say that a hypothesis simply conflicting with findings in fundamental physics is sufficient reason to reject that hypothesis. As a result, they hold that quantum physics has demonstrated that there are no things, only structures.

First let us consider whether scientific findings always warrant jettisoning commonsense ideas. Science is refined everyday reasoning, but as demonstrated in this chapter, its refinement actually ill-suits it for metaphysical work. To consider an example relevant to Ladyman and Ross's project, quantum physics is not reconciled with the theories of relativity, and I do not know of a reason to expect it to be in the future. Each of these scientific achievements provides an understanding of a certain domain—one domain certain phenomena at very large scales, the other certain phenomena at infinitesimal scales.[3] Multiplicity and disunity tend to persist among our scientific approaches and findings. This was illustrated above with the examples of various approaches to studying human aggression, as well as the variety of behavioral ecology approaches to accounting for cooperative behavior. And if assumptions or direct claims that figure into different scientific achievements are not reconcilable with one another, there are no grounds for concluding that those assumptions and claims should be reconcilable with extra-scientific beliefs. Moreover, the ways in which science's

[3]The idea that physics is fully general is inaccurate of the actual content of physics; theories and models there are just as limited in scope and accuracy as elsewhere in science.

practices and products are human-centric prevents them from being objective in the strong sense needed to ground metaphysical conclusions.

Melnyk (2013) is critical of Ladyman and Ross's idea that science provides access to objective reality. Nonetheless, he agrees that when the results of scientific methods and everyday methods conflict, we should always prefer the former (p.84). But this is also wrong. Science extends the reach of human understanding considerably, but it cannot escape its basis in our experiences. Those experiences must ultimately ground our concepts and our scientific understanding. Of course, this is not to suggest that quantum mechanics should be rejected because it belies what common sense would lead us to expect. But I do suggest that concluding on the basis of quantum physics that—at *any* scale, on any construal—there are no objects is overly hasty. This is not an element of quantum physics, but an extension of it to include the scale of middle-sized objects. It is familiarity with those very objects, and the scientific and non-scientific generalizations that can be made about them, that serve as our epistemic departure for all of science. Extending scientific results in certain ways amounts to an overextension; I suggest this is one of those ways.

A brief departure from the main focus here will, I think, prove helpful. The relationship I have motivated between science and commonsense reasoning is reminiscent of an idea central to Otto Neurath's conception of science and, in particular, the role of so-called protocol sentences. In Neurath's view, protocol sentences—the ultimate epistemic basis for all of science—are third-person reports of everyday observations, using ordinary language. That language is ambiguous and imprecise; Neurath called such terms *Ballungen*, which has been variously translated as "congestions," "conglomerations," and "clusters." The ambiguity and imprecision serves a purpose: these stabilize language-use across individuals, cultures, and time-periods. In contrast, scientific terms are precise, but they are also theory-driven. Neurath points out that "the terms of science must adapt themselves much more to the new theories than a cluster" (Neurath, 1936), reprinted in (Neurath, 1983, p.149). And so, according to Neurath,

> Our whole life consists in two opposite movements: in the one we tend to acquire always new concepts and to modify those that tradition has left us; but in the other we are obliged to take the traditional statements as the basis for our departure (p.150).

This is, akin to, and indeed related to, the epistemic relationship between common sense and science. The tools of science have much extended the reach of human understanding, and they have rightly led us to reject many commonsense beliefs that at one time appeared unassailable. But certain commonsense beliefs get the whole endeavor up and running, and even the best results of our scientific enterprise will always be more tentative.

I also take issue with a second position Ladyman and Ross employ to reach their metaphysical conclusions. In my view, physics should not be accorded a special status relative to the other sciences. This is a claim about the relationship among the fields of science, and so what we find in science is properly used to decide the claim's well-foundedness. But what science shows us does not favor Ladyman and Ross's conclusion. The epistemic position of our best fundamental physics is similar to, or perhaps worse than, our best findings in other fields. There is no reason to expect our investigations of microphysical phenomena to be epistemically privileged. Indeed, those investigations face a number of epistemic difficulties that other fields do not, such as the inaccessibility of direct observation and the tremendous amount of equipment generally required for experimentation. So there is no reason for our best physical theories to override findings from other fields.

Consider that William Thomson, later Lord Kelvin, and Darwin famously disagreed regarding the age of the Earth. In the first edition of the *Origin of Species* (1859), Darwin estimated from geological evidence that the Earth was of an age sufficient for gradual evolution by natural selection. However, based on thermodynamics, Kelvin estimated that the Sun was 30 million years old. This suggested that the Earth had actually not been habitable for a sufficient period of time for such evolutionary change. Darwin was sufficiently convinced by this that he removed discussion of timescales from the later editions of the Origin, but Kelvin was wrong, for he did not appreciate the role of fusion in the Sun's production of energy. The well-founded, carefully applied physics turned out to be wrong, and the geology and speculative, new biology turned out to be right. Granted, certain claims in physics are more epistemically secure than certain claims in other sciences. But the reverse is also true. If physics is not more trustworthy than other fields of investigation, then this undermines those projects, like Ladyman and Ross's, that would base their metaphysical conclusions entirely on the products of fundamental physics. In biology, geology, and many other properly scientific fields, references to particular objects abound.

At root, the position I advocate here is simply that extreme caution is warranted when drawing metaphysical conclusions from the practices or products of science. One cannot in general simply read metaphysical implications directly off scientific findings. This is due to the nature of science—its human-centric design and its limited connection to truth. An extension of scientific findings does not necessarily trump commonsense judgments. And, physics is not special among the sciences. Notice that this does not bar every attempt at a naturalized metaphysics nor a "scientific" metaphysics. That I advocate a wedge between scientific findings and metaphysical implications does not mean that I think analytic metaphysics should ignore the scientific enterprise and continue business as usual. One might very well go about constructing a metaphysics that is consistent with and informed by our best science (as well as everyday reasoning), and the view would almost certainly benefit from this connection. This would be analogous to the eventuation of commonsense-style claims from scientific research that are straightforwardly true or false as I outlined in the previous section. But any such metaphysics will not be read off our science, not even our fundamental physics.

Bibliography

(2012).

(2013).

(2013).

Akçay, Erol, Jeremy van Cleve, Marcus W. Feldman, and Joan Roughgarden (2009), "A theory for the evolution of other-regard integrating proximate and ultimate perspectives", *Proceedings of the National Academy of Sciences* 106: 19061–19066.

Axelrod, Robert, and William D. Hamilton (1981), "The evolution of cooperation", *Science* 211: 1390–1396.

Bateman, A.J. (1948), "Intra-Sexual Selection in Drosophila", *Heredity* 2: 349–368.

Batterman, Robert W. (2002), *The Devil in the Details*, Oxford: Oxford University Press.

——— (2009), "Idealization and modeling", *Synthese* 169: 427–446.

Bechtel, William, and Robert C. Richardson (1993), *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*, Princeton: Princeton University Press.

Bromberger, Silvain (1966), "Why-Questions", in R. Colodny, ed., *Mind and Cosmos*, Pittsburgh: University of Pittsburgh Press, 86–111.

Carnap, Rudolf (1928), *The Logical Structure of the World*, Berkeley: University of California Press.

Cartwright, Nancy (1983), *How the Laws of Physics Lie*, Oxford: Oxford University Press.

——— (1989), *Nature's Capacities and Their Measurement*, Oxford: Clarendon Press.

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

——— (2007), *Hunting Causes and Using Them: Approaches in Philosophy and Economics*, Cambridge University Press.

Chakravartty, Anjan (2013), "Scientific Realism", in Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/entries/scientific-realism/, Summer 2013 ed.

Chao, Hsiang-Ke, Szu-Ting Chen, and Roberta L. Millstein, eds. (2013), *Mechanism and Causality in Biology and Economics*, Springer.

Cohen, L. Jonathan (1992), *An Essay on Belief and Acceptance*, New York: Clarendon Press.

Craver, Carl F., and William Bechtel (2007), "Top-Down Causation without Top-Down Causes", *Biology and Philosophy* 22: 547–563.

Darwin, Charles (1859), *On The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*, John Murray, 1st ed.

——— (1871), *The Descent of Man and Selection in Relation to Sex*, London: John Murray.

Dizadji-Bahmani, Foad, Roman Frigg, and Stephan Hartmann (2010), "Who's afraid of Nagelian reduction?", *Erkenntnis* 73: 393–412.

Douglas, Heather (2009a), "Reintroducing Prediction to Explanation", *Philosophy of Science* 76: 444–463.

——— (2009b), *Science, Policy, and the Value-Free Ideal*, University of Pittsburgh Press.

Dowe, P. (2000), *Physical causation*, Cambridge University Press.

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

Earman, John, Clark Glymour, and Sandra Mitchell, eds. (2003), *Ceteris Paribus Laws*, Springer.

Elgin, Catherine Z. (2004), "True Enough", *Philosophical Issues* 14: 113–131.

——— (2010), "Keeping Things in Perspective", *Philosophical Studies* 150: 439–447.

Elliott, Kevin C. (2011), *Is a Little Pollution Good for You?: Incorporating Societal Values in Environmental Research*, Oxford University Press.

——— (2013), "Douglas on Values: From Indirect Roles to Multiple Goals", *Studies in History and Philosophy of Science* 44: 375–383.

Elliott, Kevin C., and David Willmes (2013), "Cognitive Attitudes and Values in Science", *Philosophy of Science* 80.

Epstein, Brian (2012), "Agent-Based Modeling and the Fallacies of Individualism", in Paul Humphreys and Cyrille Imbert, eds., *Models, Simulations, and Representations*, Routledge, chap. 6, 115–144.

Eshel, Ilan, and Marcus W. Feldman (2001), "Optimality and Evolutionary Stability under Short-Term and Long-Term Selection", in Steven Hecht Orzack and Elliott Sober, eds., *Adaptationism and Optimality*, Cambridge Studies in Philosophy and Biology, Cambridge: Cambridge University Press, chap. 4, 114–160.

Fehr, Carla, and Kathryn S. Plaisance, eds. (2010), *Socially Relevant Philosophy of Science*, A Special Issue of *Synthese*, vol. 177.

Feibleman, James K. (1954), "Theory of Integrative Levels", *The British Journal for the Philosophy of Science* 5: 59–66.

Feynman, Richard (1967), *The Character of Physical Law*, Cambridge, MA: MIT Press.

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

Garfinkel, Alan (1981), *Forms of Explanation: Rethinking the Questions in Social Theory*, New Haven: Yale University Press.

Giere, Ronald N. (1988), *Explaining Science; A Cognitive Approach*, Chicago: University of Chicago Press.

——— (1999), *Science Without Laws*, University of Chicago Press.

Godfrey-Smith, Peter (2006), "The strategy of model-based science", *Biology and Philosophy* 21: 725–740.

Goss-Custard, J.D. (1977), "Predator responses and prey mortality in the redshank *Tringa totanus* (L.) and a preferred prey *Corophium volutator* (Pallas)", *Journal of Animal Ecology* 46: 21–36.

Gould, Stephen Jay, and R.C. Lewontin (1979), "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme", *Proceedings of the Royal Society of London, Series B* 205: 581–598.

Griffiths, Paul, and Karola Stotz (2013), *Genetics and Philosophy: An Introduction*, Cambridge Introductions to Philosophy and Biology, Cambridge University Press.

Hartl, Daniel L., and Andrew G. Clark (1997), *Principles of Population Genetics*, Sunderland: Sinauer Associates, Inc., third ed.

Hempel, Carl (1966), *Philosophy of Natural Science*, Englewood Cliffs: Prentice-Hall.

Hempel, Carl, and Paul Oppenheim (1948), "Studies in the Logic of Explanation", *Philosophy of Science* 15: 135–175.

Hesse, Mary (1966), *Models and Analogies in Science*, Notre Dame: University of Notre Dame Press.

Horgan, Terence E. (1982), "Supervenience and Microphysics", *Pacific Philosophical Quarterly* 63: 29–43.

Hrdy, Sarah Blaffer (1986), "Empathy, Polyandry, and the Myth of the Coy Female", in R. Bleier, ed., *Feminist Approaches to Science*, The Athene Series, Pergamon Press, 119–146.

Hull, David L. (1978), "A Matter of Individuality", *Philosophy of Science* 45: 335–360.

Isaac, Alistair M.C. (2013), "Modeling Without Representation", *Synthese* 190: 3611–3623.

Jackson, Frank, and Philip Pettit (1992), "In Defense of Explanatory Ecumenism", *Economics and Philosophy* 8: 1–21.

Jagers op Akkerhuis, Gerard A. J. M. (2008), "Analysing Hierarchy in the Organization of Biological and Physical Systems", *Biological Reviews* 83: 1–12.

Kennedy, Ashley Graham (2012), "A Non Representationalist View of Model Explanation", *Studies in History and Philosophy of Science* 43: 326–332.

Kerr, Benjamin, and Peter Godfrey-Smith (2002), "Individualist and multi-level perspectives on selection in structured populations", *Biology and Philosophy* 17: 477–517.

Kim, Jaegwon (1999), "Making Sense of Emergence", *Philosophical Studies* 95: 3–36.

——— (2002), "The Layered Model: Metaphysical Considerations", *Philosophical Explorations* 5: 2–20.

Kitano, Hiroaki (2002), "Systems Biology: A Brief Overview", *Science* 295: 1662.

Kitcher, Philip (1984), "1953 and All That: A Tale of Two Sciences", *Philosophical Review* 93: 335–373.

Knight, Jonathan (2002), "Sexual Stereotypes", *Nature* 415: 254–256.

Kourany, Janet (2010), *Philosophy of Science After Feminism*, Oxford University Press.

Ladyman, James, and Don Ross (2007), *Every Thing Must Go*, Oxford University Press.

Lande, Russell, and George F. Barrowclough (1987), “Effective Population Size, Genetic Variation, and Their Use in Population Management”, in Michael E. Soulé, ed., *Viable Populations for Conservation*, Cambridge University Press.

Levins, Richard (1966), “The strategy of model building in population biology”, *American Scientist* 54: 421–431.

Lewis, David (2000), “Causation as Influence”, *Journal of Philosophy* 97: 182–197.

Lidicker, Jr., William Z. (2008), “Levels of Organization in Biology: On the Nature and Nomenclature of Ecology’s Fourth Level”, *Biological Reviews* 83: 71–78.

Lloyd, Elisabeth A. (1988), *The structure and confirmation of evolutionary theory*, Princeton: Princeton University Press.

Longino, Helen E. (1990), *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton: Princeton University Press.

——— (2001), *The Fate of Knowledge*, Princeton: Princeton University Press.

——— (2013), *Studying Human Behavior: How Scientists Investigate Aggression and Sexuality*, Chicago: University of Chicago Press.

MacArthur, Robert H. (1968), “The Theory of the Niche”, in Richard C. Lewontin, ed., *Population Biology and Evolution*, Syracuse: Syracuse University Press, 159–176.

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

Matthewson, John, and Michael Weisberg (2009), “The structure of tradeoffs in model building”, *Synthese* 170.

May, R.M. (1976), “Models for Two Interacting Populations”, in R.M. May, ed., *Theoretical ecology: Principles and applications*, Philadelphia: W B Saunders Co, 49–70.

Maynard Smith, John (1982), *Evolution and the Theory of Games*, Cambridge: Cambridge University Press.

McMullin, Ernan (1985), "Galilean Idealization", *Studies in History and Philosophy of Science* 16: 247–273.

McShea, Daniel W. (1991), "Complexity and Evolution: What Everybody Knows", *Biology and Philosophy* 6: 303–324.

Melnyk, Andrew (2013), "Can Metaphysics Be Naturalized? And if So, How?", in *Scientific Metaphysics*, Oxford University Press.

Milam, Erika, Roberta Millstein, Angela Potochnik, and Joan Roughgarden (2011), "Sex and Sensibility: The Role of Social Selection, A review symposium of Roughgarden's *The Genial Gene*", *Metascience* 20: 253–277.

Mitchell, Sandra D. (2003), *Biological Complexity and Integrative Pluralism*, Cambridge Studies in Philosophy and Biology, Cambridge: Cambridge University Press.

——— (2012a), "Emergence: Logical, Functional and Dynamical", *Synthese* 185: 171–186.

——— (2012b), *Unsimple Truths: Science, Complexity, and Policy*, Chicago: University of Chicago Press.

Molles, Manuel C. (2002), *Ecology: Concepts and Applications*, McGraw-Hill.

Nagel, Ernest (1961), *The Structure of Science*, London: Routledge and Kegen Paul.

Neurath, Otto (1936), "Encyclopedia as 'Model"', in Robert S. Cohen and Marie Neurath, eds., *Philosophical Papers 1913-1946*, Dordrecht: D. Reidel Publishing Company, *Vienna Circle Collection*, vol. 16.

——— (1983), *Philosophical Papers 1913-1946*, *Vienna Circle Collection*, vol. 16, Dordrecht: D. Reidel Publishing Company.

Ney, Alyssa (2009), "Physical Causation and Difference-Making", *The British Journal for the Philosophy of Science* 60: 737–764.

Odenbaugh, Jay (2003), "Complex systems, trade-offs, and theoretical population biology: Richard Levins' 'The strategy of model building in population biology' revisited", *Philosophy of Science* 70: 1496–1507.

Odenbaugh, Jay, and Anna Alexandrova (2011), "Buyer Beware: Robustness Analyses in Economics and Biology", *Biology and Philosophy* 26: 757–771.

Okasha, Samir (2006), *Evolution and the Levels of Selection*, Oxford University Press.

O'Neill, R. V., D. L. DeAngelis, J. B. Waide, and T. F. H. Allen (1986), *A Hierarchical Concept of Ecosystems*, Princeton: Princeton University Press.

O'Neill, Robert V. (1979), "Transmutations across Hierarchical Levels", *Systems Analysis of Ecosystems* : 59–78.

Oppenheim, Paul, and Hilary Putnam (1958), "Unity of Science as a Working Hypothesis", in Herbert Feigl, Michael Scriven, and Grover Maxwell, eds., *Minnesota Studies in the Philosophy of Science*, Minneapolis: University of Minnesota Press, vol. 2, 3–36.

Orzack, Steven Hecht, and Elliott Sober (1993), "A critical assessment of Levins's 'The strategy of model building in population biology' (1966)", *Quarterly Review of Biology* 68: 533–546.

——— (1994), "Optimality models and the test of adaptationism", *The American Naturalist* 143: 361–380.

Oyama, Susan, Russell D. Gray, and Paul E. Griffiths, eds. (2001), *Cycles of Contingency: Developmental Systems and Evolution*, Life and Mind: Philosophical Issues in Biology and Psychology, Bradford.

Parker, Wendy S. (2011), "When Climate Models Agree: The Significance of Robust Model Predictions", *Philosophy of Science* 78: 579–600.

Potochnik, Angela (2009), “Optimality modeling in a suboptimal world”, *Biology and Philosophy* 24: 183–197.

——— (2010), “Levels of Explanation Reconceived”, *Philosophy of Science* 77: 59–72.

——— (2012), “Modeling Social and Evolutionary Games”, *Studies in History and Philosophy of Biological and Biomedical Sciences* 43: 202–208.

——— (2013a), “Defusing Ideological Defenses in Biology”, *BioScience* 63: 118–123.

——— (2013b), “Helen Longino, *Studying Human Behavior: How Scientists Investigate Aggression and Sexuality*”, *Notre Dame Philosophical Reviews* .

Potochnik, Angela, ed. (forthcoming), *Socially Engaged Philosophy of Science*, A Special Issue of *Erkenntnis*.

Potochnik, Angela, and Brian McGill (2012), “The Limitations of Hierarchical Organization”, *Philosophy of Science* 79: 120–140.

Putnam, Hilary (1975), *Philosophy and our Mental Life*, Cambridge: Cambridge University Press, *Philosophical Papers*, vol. 2, chap. 14, 291–303.

Regt, Henk W. De, and Dennis Dieks (2005), “A contextual approach to scientific understanding”, *Synthese* 144.

Rice, Collin (forthcoming), “Moving Beyond Causes: Optimality Models and Scientific Explanation”, *Nous* .

Ricklefs, Robert E. (2008), “Disintegration of the Ecological Community”, *The American Naturalist* 172: 741–750.

Rohwer, Yasha, and Collin Rice (2013), “Hypothetical Pattern Idealization and Explanatory Models”, *Philosophy of Science* 80: 334–355.

Rosenberg, Alexander (1994), *Instrumental Biology or the Disunity of Science*, Chicago: University of Chicago Press.

Ross, Don, James Ladyman, and Harold Kincaid, eds. (2013), *Scientific Metaphysics*, Oxford University Press.

Rueger, Alexander, and Patrick McGivern (2010), “Hierarchies and Levels of Reality”, *Synthese* 176: 379–397.

Sadava, David, H. Craig Heller, Gordon H. Orians, William K. Purves, and David Hillis (2008), *Life: The Science of Biology*, Sinauer Associates, Inc., eighth ed.

Salmon, Wesley (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton: Princeton University Press.

Schoener, Thomas W. (1986), “Mechanistic Approaches to Community Ecology: A New Reductionism”, *American Zoologist* 26: 81–106.

Sober, Elliott (1983), “Equilibrium Explanation”, *Philosophical Studies* 43: 201–210.

Sober, Elliott, and David Sloan Wilson (1998), *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Cambridge: Harvard University Press.

Solomon, Miriam (1992), “Scientific Rationality and Human Reasoning”, *Philosophy of Science* 59: 439–454.

Sterelny, Kim (1996), “Explanatory pluralism in evolutionary biology”, *Biology and Philosophy* 11: 193–214.

Strevens, Michael (2006), *Bigger than Chaos: Understanding Complexity through Probability*, Cambridge: Harvard University Press.

——— (2009), *Depth: An Account of Scientific Explanation*, Cambridge: Harvard University Press.

Suppe, Frederick (1977), *The Structure of Scientific Theories*, Urbana: University of Illinois Press, 2nd ed.

Suppes, Patrick (2002), *Representation and Invariance of Scientific Structures*, Stanford: CSLI Publications.

Trivers, Robert L. (1971), "The Evolution of Reciprocal Altruism", *The Quarterly Review of Biology* 46: 35–57.

van Fraassen, Bas C. (1980), *The Scientific Image*, Oxford: Clarendon Press.

Waters, C. Kenneth (1990), "Why the Anti-Reductionist Consensus Won't Survive: The Case of Classical Mendelian Genetics", *Philosophy of Science Association* 1: 125–39.

Weisberg, Michael (2006), "Robustness Analysis", *Philosophy of Science* 73: 730–742.

——— (2007a), "Three Kinds of Idealization", *The Journal of Philosophy* 104: 639–659.

——— (2007b), "Who is a Modeler?", *The British Journal for the Philosophy of Science* 58: 207–233.

——— (2013), *Simulation and Similarity: Using Models to Understand the World*, Oxford University Press.

Welshon, Rex (2002), "Emergence, Supervenience, and Realization", *Philosophical Studies* 108: 39–51.

West, Geoffrey B., and James H. Brown (2005), "The Origin of Allometric Scaling Laws in Biology from Genomes to Ecosystems: Towards a Quantitative Unifying Theory of Biological Structure and Organization", *Journal of Experimental Biology* : 1575–1592.

Wilson, Edward O., and William H. Bossert (1971), *A Primer of Population Biology*, Sunderland: Sinauer Associates, Inc.

Wimsatt, William C. (1987), "False Models as Means to Truer Theories", in N. Nitecki and A. Hoffman, eds., *Neutral Models in Biology*, Oxford: Oxford University Press, 23–55.

——— (2007), *Re-Engineering Philosophy for Limited Beings*, Cambridge: Harvard University Press.

Winsberg, Eric B. (2010), *Science in the Age of Computer Simulation*, The University of Chicago Press.

Woodward, James (2003), *Making Things Happen: A Theory of Causal Explanation*, Oxford: Oxford University Press.

——— (2007), "Causation With a Human Face", in Huw Price and Richard Corry, eds., *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, Oxford University Press, 66–105.

Worden, Lee, and Simon A. Levin (2007), "Evolutionary escape from the prisoner's dilemma", *Journal of Theoretical Biology* 245: 411–422.